World Scientific Series in 20th Century Physics – Vol. 27

Selected Papers of RICHARD FEYNMAN

With Commentary

Editor

Laurie M. Brown

World Scientific

Selected Papers of RICHARD FEYNMAN

With Commentary

World Scientific Series in 20th Century Physics

Published

Vol. 1 Gauge Theories — Past and Future
edited by R. Akhoury, B. de Wit, and P. van Nieuwenhuizen

Vol. 2 Scientific Highlights in Memory of Léon Van Hove
edited by F. Nicodemi

Vol. 3 Selected Papers, with Commentary, of T. H. R. Skyrme
edited by G. E. Brown

Vol. 4 Salamfestschrift
edited by A. Ali, J. Ellis and S. Randjbar-Daemi

Vol. 5 Selected Papers of Abdus Salam (with Commentary)
edited by A. Ali, C. Isham, T. Kibble and Riazuddin

Vol. 6 Research on Particle Imaging Detectors
edited by G. Charpak

Vol. 7 A Career in Theoretical Physics
edited by P. W. Anderson

Vol. 8 Lepton Physics at CERN and Frascati
edited by N. Cabibbo

Vol. 9 Quantum Mechanics, High Energy Physics and Accelerators — Selected Papers of J. S. Bell (with Commentary)
edited by M. Bell, K. Gottfried and M. Veltman

Vol. 10 How We Learn; How We Remember: Toward an Understanding of Brain and Neural Systems — Selected Papers of Leon N. Cooper
edited by L. N. Cooper

Vol. 11 30 Years of the Landau Institute — Selected Papers
edited by I. M. Khalatnikov and V. P. Mineev

Vol. 12 Sir Nevill Mott — 65 Years in Physics
edited by N. Mott and A. S. Alexandrov

Vol. 13 Broken Symmetry — Selected Papers of Y. Nambu
edited by T. Eguchi and K. Nishijima

Vol. 14 Reflections on Experimental Science
edited by M. L. Perl

Vol. 15 Encounters in Magnetic Resonances — Selected Papers of Nicolaas Bloembergen (with Commentary)
edited by N. Bloembergen

Vol. 16 Encounters in Nonlinear Optics — Selected Papers of Nicolaas Bloembergen (with Commentary)
edited by N. Bloembergen

Vol. 17 The Collected Works of Lars Onsager (with Commentary)
edited by P. C. Hemmer, H. Holden and S. K. Ratkje

Vol. 18 Selected Works of Hans A. Bethe (with Commentary)
edited by Hans A. Bethe

Vol. 19 Selected Scientific Papers of Sir Rudolf Peierls (with Commentary)
edited by R. H. Dalitz and R. Peierls

Vol. 20 The Origin of the Third Family
edited by O. Barnabei, L. Maiani, R. A. Ricci and F. R. Monaco

Vol. 21 Spectroscopy with Coherent Radiation — Selected Papers of Norman F. Ramsey (with Commentary)
edited by N. F. Ramsey

Vol. 22 A Quest for Symmetry — Selected Works of Bunji Sakita
edited by K. Kikkawa, M. Virasoro and S. R. Wadia

Vol. 23 Selected Papers of Kun Huang (with Commentary)
edited by B.-F. Zhu

Vol. 24 Subnuclear Physics — The First 50 Years: Highlights from Erice to ELN
A. Zichichi; edited by O. Barnabei, P. Pupillo and F. Roversi Monaco

Vol. 25 The Creation of Quantum Chromodynamics and the Effective Energy
V. N. Gribov, G. 't Hooft, G. Veneziano and V. F. Weisskopf; edited by L. N. Lipatov

Vol. 26 A Quantum Legacy — Seminal Papers of Julian Schwinger
edited by K. A. Milton

Forthcoming

The Legacy of Léon Van Hove
edited by A. Giovannini

In Conclusion — A Collection of Summary Talks in High Energy Physics
edited by J. D. Bjorken

Selected Papers of Kenneth G. Wilson
edited by M. E. Peskin and B. Baaquie

Selected Works of Emil Wolf (with Commentary)
edited by E. Wolf

World Scientific Series in 20th Century Physics – Vol. 27

Selected Papers of RICHARD FEYNMAN

With Commentary

Editor

Laurie M. Brown

Northwestern University, USA

World Scientific

Singapore • New Jersey • London • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: Suite 202, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

The editor and the publisher would like to thank the following for their assistance and their permission to reproduce the articles found in this volume:

Academic Press (*Ann. Phys.*), American Physical Society (*Phys. Rev.*, *Phys. Rev. Lett.*, *Rev. Mod. Phys.*), Elsevier Science Publishers B. V. (*Nucl. Phys.*), Gordon and Breach, Kluwer Academic/Plenum Publisher (*Found. Phys.*), Polish Academy of Sciences (*Acta Phys. Pol.*), W. H. Freeman & Co.

First published 2000
Reprinted 2003

SELECTED PAPERS OF RICHARD FEYNMAN

ISBN 981-02-4130-5
ISBN 981-02-4131-3 (pbk)

Printed in Singapore by World Scientific Printers

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: Suite 202, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

The editor and the publisher would like to thank the following for their assistance and their permission to reproduce the articles found in this volume:

Academic Press (*Ann. Phys.*), American Physical Society (*Phys. Rev.*, *Phys. Rev. Lett.*, *Rev. Mod. Phys.*), Elsevier Science Publishers B. V. (*Nucl. Phys.*), Gordon and Breach, Kluwer Academic/Plenum Publisher (*Found. Phys.*), Polish Academy of Sciences (*Acta Phys. Pol.*), W. H. Freeman & Co.

First published 2000
Reprinted 2003

SELECTED PAPERS OF RICHARD FEYNMAN

ISBN 981-02-4130-5
ISBN 981-02-4131-3 (pbk)

Printed in Singapore by World Scientific Printers

CONTENTS

PREFACE

As its title indicates, this volume contains a selection of Feynman's important scientific papers together with short comments. Most of the papers contain pure research, but among them are scattered some articles that are largely pedagogical, such as published lectures that Feynman gave at advanced physics workshops and summer schools. As the editor I chose the papers and also provided the comments, except as indicated in the text. Such a selection cannot avoid arbitrariness, and I apologize to those who feel that their favorites may have been unjustly omitted.

In the course of preparation, I have consulted some physicists, historians, and others, whom I would like to thank: Tian Yu Cao, Michael Cohen, Don Ellis, Joan Feynman, and Robert Michaelson. Carl Iddings and Frank Vernon, Jr. sent me valuable information concerning their collaborations with Feynman which are included in Part III. Danny Hillis helped to orient me with regard to the papers on computers. I am especially indebted to Alexander Fetter for writing the commentary which appears with the liquid helium papers. I am also greatly obliged for the hospitality of Judith Goodstein and the staff at the Caltech Archives who assisted me when I was reading Feynman's unpublished documents, and provided the bibliography of Feynman's writings at the end of this volume. Finally, I wish to acknowledge the great help of the editors at World Scientific.

RICHARD PHILLIPS FEYNMAN

One of the century's outstanding scientists, Richard P. Feynman was born on 11 May 1918 at Far Rockaway, Queens, a part of New York City, where he received his education through high school. He then attended the Massachusetts Institute of Technology; exhibiting superior mathematical skills, he took advanced physics courses and earned a B.S. degree. According to S.S. Schweber, "By the time Feynman finished his undergraduate studies at MIT in 1939 he had mastered many of the fields of theoretical physics." Feynman went on to do graduate school at Princeton University, where his principal mentor was Professor John A. Wheeler, under whose sponsorship he wrote a doctoral thesis entitled "The Principle of Least Action in Quantum Mechanics"; he was awarded a Ph.D. in 1942.

From Princeton Feynman moved to Los Alamos, New Mexico, to work on the wartime development of the nuclear bomb. There he became an important member, and the youngest group leader, of the theoretical physics division headed by Hans Bethe of Cornell University. After the war Bethe persuaded Feynman to accept a faculty position at Cornell, and it was in Ithaca, New York, that Feynman developed his diagrammatic methods and accomplished the work that earned him the Nobel Prize in Physics. In 1950 he became a professor at the California Institute of Technology, being appointed the Richard Chase Tolman Professor of Physics in 1959. He remained in Pasadena until his death at the age of sixty-nine, on 15 February 1988.

Feynman had three marriages. The first was in 1942 to Arline Greenbaum, who was incurably ill at the time of the marriage and died in 1945. His second marriage, to Mary Louise Bell in 1952, ended in divorce. In 1960 he married Gweneth Howarth, with whom he had a son, Carl, and an adopted daughter, Michelle.

Early in his career, Feynman became well-known to the world physics community for his brilliant research, his outstanding teaching, and his flamboyant personality. After sharing the Nobel Prize in 1965 for his work on renormalized quantum electrodynamics (QED) with Julian Seymour Schwinger and Sin-itiro Tomanaga, he also became an important establishment figure, receiving many invitations to lecture throughout the world (most of which he politely declined). However, he became a notable *public* figure only near the end of his life, when, in January 1986, President Reagan appointed him to a presidential commission to investigate the cause of the explosion of the space shuttle *Challenger*. Appearing in the televised hearing of the investigative committee, Feynman performed a simple experiment with a glass of ice water and a piece of the shuttle's failed O ring, in order to demonstrate the immediate cause of the disaster. His presentation made a deep impression on millions of television viewers, and, coupled with two best-selling books of his reminiscences ([112] and [121]*), other television appearances, and adulation in the press, Feynman became something of a cult figure, especially after his death. A few of the biographies and other books dealing with him and his work are listed after this account.

Besides the Nobel Prize, Feynman received the Albert Einstein Award (1954), the Ernest Orlando Lawrence Award for Physics (1962), the Oersted Medal (1972), and the Niels Bohr International Gold Medal (1973). He was a Member of the Brazilian Academy of Sciences and a Foreign Fellow of the Royal Society of London (1965). He was elected to the National

*Numbers in square brackets refer to the bibliography at the end of this volume.

Academy of Sciences (USA), but later resigned, and often declared that he was totally uninterested in receiving "honors."

See:
James Gleick, *Genius.* New York, 1992.
Laurie M. Brown and John S. Rigden (eds.), *Most of the Good Stuff.* New York, 1993.
Jagdish Mehra, *The Beat of a Different Drum.* Oxford, 1994.
Silvan S. Schweber, *QED and the Men Who Made It.* Princeton, 1994.
Christopher Sykes (ed.), *No Ordinary Genius.* New York, 1994.
David Goodstein and Judith Goodstein, *Feynman's Lost Lecture: The Motion of Planets Around the Sun.* New York, 1996.
John Gribbin and Mary Gribbin, *Richard Feynman: A Life in Science.* New York, 1997.

I. Quantum Chemistry

In his final year as an undergraduate at the Massachusetts Institute of Technology, Feynman published with one of his teachers, Manuel S. Vallarta, a Letter to the Editor of the *Physical Review* on cosmic rays [1]. He also completed a senior dissertation under John C. Slater entitled "Forces and Stresses in Molecules" and published a shortened version, "Forces in Molecules," as an article in the *Physical Review*. The latter contained a result — a general quantum-mechanical theorem — that has played an important role in theoretical chemistry and condensed matter physics and is frequently cited as the Hellmann–Feynman theorem.[1] According to Feynman's abstract, "The force on a nucleus in an atomic system is shown to be just the classical electrostatic force that would be exerted on this nucleus by other nuclei and by the electrons' charge distribution." Quantum mechanics is used to calculate the charge distribution as the absolute square of the Schrödinger wave function. The importance of the forces on the atomic nuclei for molecular geometry, the theory of chemical binding, and for crystal structure is evident.

Selected Paper

[2] Forces in molecules, *Phys. Rev.* **56** (1939): 340–343.

[1]The German quantum chemist H. Hellmann published the theorem in a textbook, *Einführing in die Quantenchemie* (1937, Franz Deuticke, Leipzig), but the reference was unknown to Feynman and Slater. For the history of the H–F theorem see J.I. Musher, *Am. J. Phys.* **34** (1966): 267–268, and J.C. Slater, *Solid State and Molecular Theory* (1975, Wiley, New York), pp. 193–199. For applications see B.M. Deb (ed.), *The Force Concept in Chemistry* (1981, Van Nostrand Reinhold, New York).

AUGUST 15, 1939 PHYSICAL REVIEW VOLUME 56

Forces in Molecules

R. P. FEYNMAN
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received June 22, 1939)

Formulas have been developed to calculate the forces in a molecular system directly, rather than indirectly through the agency of energy. This permits an independent calculation of the slope of the curves of energy *vs.* position of the nuclei, and may thus increase the accuracy, or decrease the labor involved in the calculation of these curves. The force on a nucleus in an atomic system is shown to be just the classical electrostatic force that would be exerted on this nucleus by other nuclei and by the electrons' charge distribution. Qualitative implications of this are discussed.

MANY of the problems of molecular structure are concerned essentially with forces. The stiffness of valence bonds, the distortions in geometry due to the various repulsions and attractions between atoms, the tendency of valence bonds to occur at certain definite angles with each other, are some examples of the kind of problem in which the idea of force is paramount.

Usually these problems have been considered through the agency of energy, and its changes with changing configuration of the molecule. The reason for this indirect attack through energy, rather than the more qualitatively illuminating one, by considerations of force, is perhaps twofold. First it is probably thought that force is a quantity that is not easily described or calculated by wave mechanics, while energy is, and second, the first molecular problem to be solved is the analysis of band spectra, strictly a problem of energy as such. It is the purpose of this paper to show that forces are almost as easy to calculate as energies are, and that the equations are quite as easy to interpret. In fact, all forces on atomic nuclei in a molecule can be considered as purely classical attractions involving Coulomb's law. The electron cloud distribution is prevented from collapsing by obeying Schrödinger's equation. In these considerations the nuclei are considered as mass points held fixed in position.

A usual method of calculating interatomic forces runs somewhat as follows.

For a given, fixed configuration of the nuclei, the energy of the entire system (electrons and nuclei) is calculated. This is done by the variation method or other perturbation schemes. This entire process is repeated for a new nuclear position, and the new value of energy calculated. Proceeding in this way, a plot of energy *vs.* position is obtained. The force on a nucleus is of course the slope of this curve.

The following method is one designed to obtain the forces at a given configuration, when only the configuration is known. It does not require the calculations at neighboring configurations. That is, it permits a calculation of the slope of the energy curve as well as its value, for any particular configuration. It is to be emphasized that this allows a considerable saving of labor of calculations. To obtain force under the usual scheme the energy needs to be calculated for two or more different and neighboring configurations. Each point requires the calculation of the wave functions for the entire system. In this new method, only one configuration, the one in question, need have its wave functions computed in detail. Thus the labor is considerably reduced. Because it permits one to get an independent value of the slope of the energy curve, the method might increase the accuracy in the calculation of these curves, being especially helpful in locating the normal separation, or position of zero force.

In the following it is to be understood that the nuclei of the atoms in the molecule, or other atomic system, are to be held fixed in position, as point charges, and the force required to be applied to the nuclei to hold them is to be calculated. This will lead to two possible definitions of force in the nonsteady state, for then the energy is not a definite quantity, and the slope of the energy curve shares this indefinite-

ness. It will be shown that these two possible definitions are exactly equivalent in the steady-state case, and, of course, no ambiguity should arise there.

Let λ be one of any number of parameters which specify nuclear positions. For example, λ might be the x component of the position of one of the nuclei. A force f_λ is to be associated with λ in such a way that $f_\lambda d\lambda$ measures the virtual work done in displacing the nuclei through $d\lambda$. This will define the force only when the molecule is in a steady state, of energy U, for then we can say $f_\lambda = -\partial U/\partial\lambda$. In the non-steady-state case we have no sure guide to a definition of force. For example, if $\bar{U} = \int \psi^* H \psi dv$ be the average energy of the system of wave function ψ and Hamiltonian H, we might define

$$f_\lambda' = -\partial(\bar{U})/\partial\lambda'. \tag{1}$$

Or again, we might take f_λ to be the average of $-\partial H/\partial\lambda$ or

$$f_\lambda = -\left\langle \frac{\partial H}{\partial\lambda} \right\rangle_{Av} = -\int \psi^* \frac{\partial H}{\partial\lambda} \psi dv. \tag{2}$$

We shall prove that under steady-state conditions, both these definitions of force become exactly equivalent, and equal to $-\partial U/\partial\lambda$, the slope of the energy curve. Since (2) is simpler than (1) we can define force by (2) in general. In particular, it gives a simple expression for the slope of the energy curve.

Thus we shall prove, when $H\psi = U\psi$ and $\int \psi\psi^* dv = 1$ that,

$$\frac{\partial U}{\partial\lambda} = \int \psi^* \frac{\partial H}{\partial\lambda} \psi dv.$$

Now

$$U = \int \psi^* H \psi dv,$$

whence,

$$\frac{\partial U}{\partial\lambda} = \int \psi^* \frac{\partial H}{\partial\lambda} \psi dv + \int \frac{\partial\psi^*}{\partial\lambda} H\psi dv + \int \psi^* H \frac{\partial\psi}{\partial\lambda} dv.$$

Since H is a self-adjoint operator,

$$\int \psi^* H \frac{\partial\psi}{\partial\lambda} dv = \int \frac{\partial\psi}{\partial\lambda} H\psi^* dv.$$

But $H\psi = U\psi$ and $H\psi^* = U\psi^*$ so that we can write,

$$\frac{\partial U}{\partial\lambda} = \int \psi^* \frac{\partial H}{\partial\lambda} \psi dv + U\int \frac{\partial\psi^*}{\partial\lambda} \psi dv + U\int \frac{\partial\psi}{\partial\lambda} \psi^* dv.$$

These last two terms cancel each other since their sum is,

$$U\frac{\partial}{\partial\lambda}\int \psi^*\psi dv = U\frac{\partial}{\partial\lambda}(1) = 0.$$

Whence

$$\frac{\partial U}{\partial\lambda} = \int \psi^* \frac{\partial H}{\partial\lambda} \psi dv$$

in the steady state. This much is true, regardless of the nature of H, (whether for spin, or nuclear forces, etc.). In the special case of atomic systems when $H = T + V$ where T is the kinetic energy operator, and V the potential, since $\partial H/\partial\lambda = \partial V/\partial\lambda$ we can write

$$f_\lambda' = f_\lambda = -\frac{\partial U}{\partial\lambda} = -\int \psi^* \frac{\partial V}{\partial\lambda} \psi dv. \tag{3}$$

The actual calculation of forces in a real molecule by means of this theorem is not impractical. The $\int \psi^*\psi(\partial V/\partial\lambda)dv$ is not too different from $\int \psi^*\psi V dv$, which must be calculated if the energy is to be found at all in the variational method. Although the theorem (3) is the most practical for actual calculations, it can be modified to get a clearer qualitative picture of what it means. Suppose, for example, the system for which ψ is the wave function contains several nuclei, and let the coordinates of one of these nuclei, α, be X^α, Y^α, Z^α or X_μ^α where $\mu = 1, 2, 3$, mean X, Y, Z. If we take our λ parameter to be one of these coordinates, the resultant force on the nucleus α in the μ direction will be given directly by

$$f_\mu^\alpha = -\int \psi\psi^* (\partial V/\partial X_\mu^\alpha) dv$$

from (3).

Now V is made up of three parts, the interaction of all nuclei with each other ($V_{\alpha\beta}$), of each nucleus with an electron ($V_{\beta i}$), and of each

electron with every other (V_{ij}); or

$$V=\sum_{\alpha,\beta} V_{\alpha\beta}+\sum_{\beta,i} V_{\beta i}+\sum_{i,j} V_{ij}.$$

Suppose $x_\mu{}^i$ are the coordinates of electron i, and as before $X_\mu{}^\alpha$ those of nucleus α of charge q_α. Then $V_{\beta i}=q_\beta e/R_{\beta i}$, where

$$R_{\beta i}{}^2=\sum_{\mu=1}^{3}(X_\mu{}^\beta-x_\mu{}^i)^2.$$

So we see that

$$\frac{\partial V_{\beta i}}{\partial X_\mu{}^\alpha}=-\delta_\beta{}^\alpha\frac{\partial V_{\beta i}}{\partial x_\mu{}^i} \quad \text{and that} \quad \frac{\partial V_{ij}}{\partial X_\mu{}^\alpha}=0.$$

Then (3) leads to

$$f_\mu{}^\alpha=+\int\psi\psi^*\sum_i\frac{\partial V_{\alpha i}}{\partial x_\mu{}^i}dv-\sum_\beta\int\frac{\partial V_{\alpha\beta}}{\partial X_\mu{}^\alpha}\psi\psi^*dv$$

$$=\int\sum_i\frac{\partial V_{\alpha i}}{\partial x_\mu{}^i}\left[\int i\int\psi\psi^*dv\right]dv-\sum_\beta\frac{\partial V_{\alpha\beta}}{\partial X_\mu{}^\alpha} \quad (4)$$

since $\partial V_{\alpha i}/\partial x_\mu{}^i$ does not involve any electron coordinate except those of electron i. $\int i\int\cdots dv$ means the integral over the coordinates of all electrons except those of electron i. The last term has been reduced since $\partial V_{\alpha\beta}/\partial x_\mu{}^\alpha$ does not involve the electron coordinates, and is constant as far as integration over these coordinates goes. This term gives ordinary Coulomb electrostatic repulsion between the nuclei and need not be considered further. Now $e\int i\int\psi\psi^*dv$ is just the charge density distribution $\rho_i(x)$ due to electron i, where e is the charge on one electron. The electric field $E_\mu{}^\alpha(x^i)$ at any point x^i due to the nucleus α is $(1/e)\partial V_{\alpha i}/\partial x_\mu{}^i$, so that (4) may be written

$$f_\mu{}^\alpha=\int[\sum_i\rho_i(x)]E_\mu{}^\alpha(x)dv-\sum_\beta\frac{\partial V_{\alpha\beta}}{\partial X_\mu{}^\alpha}.$$

The $3N$ space for N electrons has been reduced to a 3 space. This can be done since $E_\mu{}^\alpha(x^i)$ depends only on x^i and is the same function of x^i no matter which i we pick. This implies the following conclusion:

The force on any nucleus (considered fixed) in any system of nuclei and electrons is just the classical electrostatic attraction exerted on the nucleus in question by the other nuclei and by the electron charge density distribution for all electrons,

$$\rho(x)=\sum_i\rho_i(x).$$

It is possible to simplify this still further. Suppose we construct an electric field vector F such that

$$\nabla\cdot F=-4\pi\rho(x); \quad \nabla\times F=0.$$

Now from the derivation of $E_\mu{}^\alpha$ we know that it arises from the charge q_α on nucleus α, so that $\nabla\cdot E^\alpha=0$ except at the charge α where its integral equals q_α. Further,

$$-\sum_\beta\frac{\partial V_{\alpha\beta}}{\partial X_\mu{}^\alpha}=q_\alpha[\sum_\beta E_\mu{}^\beta]_{\text{at } x^\alpha}.$$

Then

$$f_\mu{}^\alpha=-\frac{1}{4\pi}\int(\nabla\cdot F)E_\mu{}^\alpha dv-\sum_\beta\frac{\partial V_{\alpha\beta}}{\partial x_\mu{}^\alpha}$$

$$=+\frac{1}{4\pi}\int F_\mu(\nabla\cdot E_\mu{}^\alpha)dv-\sum_\beta\frac{\partial V_{\alpha\beta}}{\partial x_\mu{}^\alpha}$$

$$=q_\alpha[F_\mu]_{\text{at } x^\alpha}+q_\alpha[\sum_\beta E_\mu{}^\beta]_{\text{at } x^\alpha} \quad (5)$$

the transformation of the integral being accomplished by integrating by parts. Or finally, the force on a nucleus is the charge on that nucleus times the electric field there due to all the electrons, plus the fields from the other nuclei. This field is calculated classically from the charge distribution of each electron and from the nuclei.

It now becomes quite clear why the strongest and most important attractive forces arise when there is a concentration of charge between two nuclei. The nuclei on each side of the concentrated charge are each strongly attracted to it. Thus they are, in effect, attracted toward each other. In a H_2 molecule, for example, the antisymmetrical wave function, because it must be zero exactly between the two H atoms, cannot concentrate charge between them. The symmetrical solution, however, can easily permit charge concentration between the nuclei, and hence it is only the solution which is symmetrical that leads to strong attraction, and the formation of a molecule, as is well known. It is

clearly seen that concentrations of charge between atoms lead to strong attractive forces, and hence, are properly called valence bonds.

Van der Waals' forces can also be interpreted as arising from charge distributions with higher concentration between the nuclei. The Schrödinger perturbation theory for two interacting atoms at a separation R, large compared to the radii of the atoms, leads to the result that the charge distribution of each is distorted from central symmetry, a dipole moment of order $1/R^7$ being induced in each atom. The negative charge distribution of each atom has its center of gravity moved slightly toward the other. It is not the interaction of these dipoles which leads to van der Waals' force, but rather the attraction of each nucleus for the distorted charge distribution of its *own* electrons that gives the attractive $1/R^7$ force.

The author wishes to express his gratitude to Professor J. C. Slater who, by his advice and helpful suggestions, aided greatly in this work. He would also like to thank Dr. W. C. Herring for the latter's excellent criticisms.

II. Classical and Quantum Electrodynamics

While Feynman made many original and imaginative contributions to theoretical physics, it may well be that his place in the history of science will be largely based on his approach to renormalizing quantum electrodynamics (QED), and especially on the tools that he invented to accomplish that goal, such as path integrals, the operator calculus, and the famous Feynman diagrams. Eventually QED may be replaced by a finite theory, rather than the present divergent, though renormalizable, one. (QED is already incorporated in the unified electroweak theory, one of the two parts of the Standard Model.) Feynman himself never regarded renormalized QED as complete, frequently pointing out its limitations and suggesting that it was merely what we now call an "effective field theory." But even if QED proves to be transitory, the theoretical methods that Feynman developed are permanently embedded in mathematical physics, and have been widely applied in areas far beyond their original domain.

Of our nine selected papers that deal with electrodynamics, two are in the nature of reviews, one being his 1961 report to the Solvay Conference [45], included for its lively originality. The other is Feynman's Nobel Lecture, which is placed first in this section on electrodynamics for a special reason. That is not because the Prize itself has great scientific significance. (He even thought of refusing it, had that been practical; its award to Feynman honors the Prize as much as its recipient.) Rather, paper [73] occupies the leading position here because it provides a far more valuable and eloquent commentary on this group of papers than could be produced in any other way. It outlines the significant steps, including the less successful ones, by which Feynman recognized and worked his way through the problem situation, from classical to quantum electrodynamics, to find the solution, and it tells about the physicists with whom he interacted. In its context as a Nobel Lecture it is a surprisingly "human" story; at the least, its style would surprise us if it were told by anyone other than Feynman.

The separation of the various papers into neat groups is rather arbitrary, and I have chosen to place papers [7], [14], and [15], which also contain derivations of QED, in a separate Section III which emphasizes the methodological innovations, because they have a wide range of applications in other fields as well. This applies especially to [7], which describes the path integral method.[1]

II.A *Classical and Quantum Electrodynamics — The Space–Time View*

Selected Paper

[73] The development of the space–time view of quantum electrodynamics. In: *Les Prix Nobel 1965.* Stockholm, 1965: Imprimerie Royale P.A. Norstedt & Soner: 172–191.

[1]This is the subject of Feynman's doctoral thesis, written at Princeton University, under the sponsorship of John Wheeler. We would have included this dissertation, but were unable to get permission to do so.

RICHARD P. FEYNMAN

The development of the space-time view of quantum electrodynamics

Nobel Lecture, December 11, 1965

We have a habit in writing articles published in scientific journals to make the work as finished as possible, to cover all the tracks, to not worry about the blind alleys or to describe how you had the wrong idea first, and so on. So there isn't any place to publish, in a dignified manner, what you actually did in order to get to do the work, although, there has been in these days, some interest in this kind of thing. Since winning the prize is a personal thing, I thought I could be excused in this particular situation, if I were to talk personally about my relationship to quantum electrodynamics, rather than to discuss the subject itself in a refined and finished fashion. Furthermore, since there are three people who have won the prize in physics, if they are all going to be talking about quantum electrodynamics itself, one might become bored with the subject. So, what I would like to tell you about today are the sequence of events, really the sequence of ideas, which occurred, and by which I finally came out the other end with an unsolved problem for which I ultimately received a prize.

I realize that a truly scientific paper would be of greater value, but such a paper I could publish in regular journals. So, I shall use this Nobel Lecture as an opportunity to do something of less value, but which I cannot do elsewhere. I ask your indulgence in another manner. I shall include details of anecdotes which are of no value either scientifically, nor for understanding the development of ideas. They are included only to make the lecture more entertaining.

I worked on this problem about eight years until the final publication in 1947. The beginning of the thing was at the Massachusetts Institute of Technology, when I was an undergraduate student reading about the known physics, learning slowly about all these things that people were worrying about, and realizing ultimately that the fundamental problem of the day was that the quantum theory of electricity and magnetism was not completely satisfactory. This I gathered from books like those of Heitler and Dirac. I was inspired by the remarks in these books; not by the parts in which everything was proved and demonstrated carefully and calculated, because I couldn't

understand those very well. At the young age what I could understand were the remarks about the fact that this doesn't make any sense, and the last sentence of the book of Dirac I can still remember, « It seems that some essentially new physical ideas are here needed. » So, I had this as a challenge and an inspiration. I also had a personal feeling, that since they didn't get a satisfactory answer to the problem I wanted to solve, I don't have to pay a lot of attention to what they did do.

I did gather from my readings, however, that two things were the source of the difficulties with the quantum electrodynamical theories. The first was an infinite energy of interaction of the electron with itself. And this difficulty existed even in the classical theory. The other difficulty came from some infinites which had to do with the infinite numbers of degrees of freedom in the field. As I understood it at the time(as nearly as I can remember) this was simply the difficulty that if you quantized the harmonic oscillators of the field (say in a box) each oscillator has a ground state energy of $(1/2)\hbar\omega$ and there is an infinite number of modes in a box of every increasing frequency ω, and therefore there is an infinite energy in the box. I now realize that that wasn't a completely correct statement of the central problem; it can be removed simply by changing the zero from which energy is measured. At any rate, I believed that the difficulty arose somehow from a combination of the electron acting on itself and the infinite number of degrees of freedom of the field.

Well, it seemed to me quite evident that the idea that a particle acts on itself, that the electrical force acts on the same particle that generates it, is not a necessary one-it is a sort of a silly one, as a matter of fact. And, so I suggested to myself, that electrons cannot act on themselves, they can only act on other electrons. That means there is no field at all. You see, if all charges contribute to making a single common field, and if that common field acts back on all the charges, then each charge must act back on itself. Well, that was where the mistake was, there was no field. It was just that when you shook one charge, another would shake later. There was a direct interaction between charges, albeit with a delay. The law of force connecting the motion of one charge with another would just involve a delay. Shake this one, that one shakes later. The sun atom shakes; my eye electron shakes eight minutes later, because of a direct interaction across.

Now, this has the attractive feature that it solves both problems at once. First, I can say immediately, I don't let the electron act on itself, I just let this act on that, hence, no self-energy! Secondly, there is not an infinite number of degrees of freedom in the field. There is no field at all; or if you insist on

thinking in terms of ideas like that of a field, this field is always completely determined by the action of the particles which produce it. You shake this particle, it shakes that one, but if you want to think in a field way, the field, if it's there, would be entirely determined by the matter which generates it, and therefore, the field does not have any *independent* degrees of freedom and the infinities from the degrees offreedom would then be removed. As a matter of fact, when we look out anywhere and see light, we can always « see » some matter as the source of the light. We don't just see light (except recently some radio reception has been found with no apparent material source).

You see then that my general plan was to first solve the classical problem, to get rid of the infinite self-energies in the classical theory, and to hope that when I made a quantum theory of it, everything would just be fine.

That was the beginning, and the idea seemed so obvious to me and so elegant that I fell deeply in love with it. And, like falling in love with a woman, it is only possible if you do not know much about her, so you cannot see her faults. The faults will become apparent later, but after the love is strong enough to hold you to her. So, I was held to this theory, in spite of all difficulties, by my youthful enthusiasm.

Then I went to graduate school and somewhere along the line I learned what was wrong with the idea that an electron does not act on itself. When you accelerate an electron it radiates energy and you have to do extra work to account for that energy. The extra force against which this work is done is called the force of radiation resistance. The origin of this extra force was identified in those days, following Lorentz, as the action of the electron itself The first term of this action, of the electron on itself, gave a kind of inertia (not quite relativistically satisfactory). But that inertia-like term was infinite for a point-charge. Yet the next term in the sequence gave an energy loss rate, which for a point-charge agrees exactly with the rate you get by calculating how much energy is radiated. So, the force of radiation resistance, which is absolutely necessary for the conservation of energy would disappear if I said that a charge could not act on itself.

So, I learned in the interim when I went to graduate school the glaringly obvious fault of my own theory. But, I was still in love with the original theory, and was still thinking that with it lay the solution to the difficulties of quantum electrodynamics. So, I continued to try on and off to save it somehow. I must have some action develop on a given electron when I accelerate it to account for radiation resistance. But, if I let electrons only act on other electrons the only possible source for this action is another electron in the

world. So, one day, when I was working for Professor Wheeler and could no longer solve the problem that he had given me, I thought about this again and I calculated the following. Suppose I have two charges-I shake the first charge, which I think of as a source and this makes the second one shake, but the second one shaking produces an effect back on the source. And so, I calculated how much that effect back on the first charge was, hoping it might add up the force of radiation resistance. It didn't come out right, of course, but I went to Professor Wheeler and told him my ideas. He said, -yes, but the answer you get for the problem with the two charges that you just mentioned will, unfortunately, depend upon the charge and the mass of the second charge and will vary inversely as the square of the distance R, between the charges, while the force ofradiation resistance depends on none of these things. I thought, surely, he had computed it himself, but now having become a professor, I know that one can be wise enough to see immediately what some graduate student takes several weeks to develop. He also pointed out something that also bothered me, that if we had a situation with many charges all around the original source at roughly uniform density and if we added the effect of all the surrounding charges the inverse R square would be compensated by the R^2 in the volume element and we would get a result proportional to the thickness of the layer, which would go to infinity. That is, one would have an infinite total effect back at the source. And, finally he said to me, and you forgot something else, when you accelerate the first charge, the second acts later, and then the reaction back here at the source would be still later. In other words, the action occurs at the wrong time. I suddenly realized what a stupid fellow I am, for what I had described and calculated was just ordinary reflected light, not radiation reaction.

But, as I was stupid, so was Professor Wheeler that much more clever. For he then went on to give a lecture as though he had worked this all out before and was completely prepared, but he had not, he worked it out as he went along. First, he said, let us suppose that the return action by the charges in the absorber reaches the source by advanced waves as well as by the ordinary retarded waves of reflected light; so that the law ofinteraction acts backward in time, as well as forward in time. I was enough of a physicist at that time not to say, « Oh, no, how could that be? » For today all physicists know from studying Einstein and Bohr, that sometimes an idea which looks completely paradoxical at first, if analyzed to completion in all detail and in experimental situations, may, in fact, not be paradoxical. So, it did not bother me any more

than it bothered Professor Wheeler to use advance waves for the back reaction -a solution of Maxwell's equations, which previously had not been physically used.

Professor Wheeler used advanced waves to get the reaction back at the right time and then he suggested this : If there were lots of electrons in the absorber, there would be an index of refraction n, so, the retarded waves coming from the source would have their wave lengths slightly modified in going through the absorber. Now, if we shall assume that the advanced waves come back from the absorber without an index-why? I don't know, let's assume they come back without an index-then, there will be a gradual shifting in phase between the return and the original signal so that we would only have to figure that the contributions act as if they come from only a finite thickness, that of the first wave zone. (More specifically, up to that depth where the phase in the medium is shifted appreciably from what it would be in vacuum, a thickness proportional to $\lambda/(n-1)$.) Now, the less the number of electrons in here, the less each contributes, but the thicker will be the layer that effectively contributes because with less electrons, the index differs less from 1. The higher the charges of these electrons, the more each contribute, but the thinner the effective layer, because the index would be higher. And when we estimated it, (calculated without being careful to keep the correct numerical factor) sure enough, it came out that the action back at the source was completely independent of the properties of the charges that were in the surrounding absorber. Further, it was of just the right character to represent radiation resistance, but we were unable to see if it was just exactly the right size. He sent me home with orders to figure out exactly how much advanced and how much retarded wave we need to get the thing to come out numerically right, and after that, figure out what happens to the advanced effects that you would expect if you put a test charge here close to the source? For if all charges generate advanced, as well as retarded effects, why would that test not be affected by the advanced waves from the source?

I found that you get the right answer if you use half-advanced and half-retarded as the field generated by each charge. That is, one is to use the solution of Maxwell's equation which is symmetrical in time and that the reason we got no advanced effects at a point close to the source in spite of the fact that the source was producing an advanced field is this. Suppose the source s surrounded by a spherical absorbing wall ten light seconds away, and that the test charge is one second to the right of the source. Then the source is as much

as eleven seconds away from some parts of the wall and only nine seconds away from other parts. The source acting at time $t = 0$ induces motions in the wall at time + 10. Advanced effects from this can act on the test charge as early as eleven seconds earlier, or at $t = -1$. This is just at the time that the direct advanced waves from the source should reach the test charge, and it turns out the two effects are exactly equal and opposite and cancel out! At the later time + 1 effects on the test charge from the source and from the walls are again equal, but this time are of the same sign and add to convert the half-retarded wave of the source to full retarded strength.

Thus, it became clear that there was the possibility that if we assume all actions are via half-advanced and half-retarded solutions of Maxwell's equations and assume that all sources are surrounded by material absorbing all the the light which is emitted, then we could account for radiation resistance as a direct action of the charges of the absorber acting back by advanced waves on the source.

Many months were devoted to checking all these points. I worked to show that everything is independent of the shape of the container, and so on, that the laws are exactly right, and that the advanced effects really cancel in every case. We always tried to increase the efficiency of our demonstrations, and to see with more and more clarity why it works. I won't bore you by going through the details of this. Because of our using advanced waves, we also had many apparent paradoxes, which we gradually reduced one by one, and saw that there was in fact no logical difficulty with the theory. It was perfectly satisfactory.

We also found that we could reformulate this thing in another way, and that is by a principle of least action. Since my original plan was to describe everything directly in terms of particle motions, it was my desire to represent this new theory without saying anything about fields. It turned out that we found a form for an action directly involving the motions of the charges only, which upon variation would give the equations of motion of these charges. The expression for this action A is

$$A = \Sigma_i m_i \int \left(\dot{X}^i_\mu \dot{X}^i_\mu \right)^{\frac{1}{2}} \mathrm{d}\alpha_i + \tfrac{1}{2} \sum_{\substack{ij \\ i \neq j}} e_i e_j \int\int \delta(I_{ij}^2)\, \dot{X}^i_\mu(\alpha_i)\, \dot{X}^j_\mu(\alpha_j)\, \mathrm{d}\alpha_i\, \mathrm{d}\alpha_j \quad (1)$$

where

$$I_{ij}^2 = [X^i_\mu(\alpha_i) - \mathrm{Xi},(\alpha_j)][X^i_\mu(\alpha_i) - X^j_\mu(\alpha_j)]$$

where $X^i_\mu(\alpha_i)$ is the four-vector position of the i^{th} particle as a function of

some parameter α_i, $\dot{X}^i_\mu(\alpha_i)$ is dX^i_μ (a) / $d\alpha_i$. The first term is the integral of proper time, the ordinary action of relativistic mechanics of free particles of mass m_i. (We sum in the usual way on the repeated index μ.) The second term represents the electrical interaction of the charges. It is summed over each pair of charges (the factor 1/2 is to count each pair once, the term $i=j$ is omitted to avoid self- action) .The interaction is a double integral over a delta function of the square of space- time interval I^2 between two points on the paths. Thus, interaction occurs only when this interval vanishes, that is, along light cones.

The fact that the interaction is exactly one- half advanced and half- retarded meant that we could write such a principle of least action, whereas interaction *via* retarded waves alone cannot be written in such a way.

So, all of classical electrodynamics was contained in this very simple form. It looked good, and therefore, it was undoubtedly true, at least to the beginner. It automatically gave half- advanced and half-retarded effects and it was without fields. By omitting the term in the sum when i = j, *I* omit self-interaction and no longer have any infinite self-energy. This then was the hoped-for solution to the problem of ridding classical electrodynamics of the infinities.

It turns out, of course, that you can reinstate fields if you wish to, but you have to keep track of the field produced by each particle separately. This is because to find the right field to act on a given particle, you must exclude the field that it creates itself. A single universal field to which all contribute will not do. This idea had been suggested earlier by Frenkel and so we called these Frenkel fields. This theory which allowed only particles to act on each other was equivalent to Frenkel's fields using half- advanced and half-retarded solutions.

There were several suggestions for interesting modifications of electrodynamics. We discussed lots of them, but I shall report on only one. It was to replace this delta function in the interaction by another function, say, $f(I^2_{ij})$, which is not infinitely sharp. Instead of having the action occur only when the interval between the two charges is exactly zero, we would replace the delta function of I^2 by a narrow peaked thing. Let's say that $f(Z)$ is large only near $Z=$ o width of order a^2. Interactions will now occur when $T^2 - R^2$ is of order a^2 roughly where T is the time difference and R is the separation of the charges. This might look like it disagrees with experience, but if a is some small distance, like 10^{-13} cm, it says that the time delay Tin action is roughly $\sqrt{R^2 \pm a^2}$ or approximately,-if R is much larger than a, $T = R \pm a^2/2R$. This means that the deviation of time T from the ideal theoretical time R of Maxwell, gets smaller and smaller, the further the pieces are apart. Therefore, all theories

involving in analyzing generators, motors, etc., in fact, all of the tests of electrodynamics that were available in Maxwell's time, would be adequately satisfied if a were 10^{-13} cm. If R is of the order of a centimeter this deviation in T is only 10^{-26} parts. So, it was possible, also, to change the theory in a simple manner and to still agree with all observations of classical electrodynamics. You have no clue of precisely what function to put in for f, but it was an interesting possibility to keep in mind when developing quantum electrodynamics.

It also occurred to us that if we did that (replace δ by f) we could not reinstate the term $i = j$ in the sum because this would now represent in a relativistically invariant fashion a finite action of a charge on itself. In fact, it was possible to prove that if we did do such a thing, the main effect of the self-action (for not too rapid accelerations) would be to produce a modification of the mass. In fact, there need be no mass m_j term, all the mechanical mass could be electromagnetic self-action. So, if you would like, we could also have another theory with a still simpler expression for the action A. In expression (i) only the second term is kept, the sum extended over all i and j, and some function f replaces δ. Such a simple form could represent all of classical electrodynamics, which aside from gravitation is essentially all of classical physics.

Although it may sound confusing, I am describing several different alternative theories at once. The important thing to note is that at this time we had all these in mind as different possibilities. There were several possible solutions of the difficulty of classical electrodynamics, any one of which might serve as a good starting point to the solution of the difficulties of quantum electrodynamics.

I would also like to emphasize that by this time I was becoming used to a physical point of view different from the more customary point of view. In the customary view, things are discussed as a function of time in very great detail. For example, you have the field at this moment, a differential equation gives you the field at the next moment and so on; a method, which I shall call the Hamilton method, the time differential method. We have, instead (in (i) say) a thing that describes the character of the path throughout all of space and time. The behavior of nature is determined by saying her whole space-time path has a certain character. For an action like (i) the equations obtained by variation $(\mathrm{of} X^i{}_\mu(\alpha_i))$ are no longer at all easy to get back into Hamiltonian form. If you wish to use as variables only the coordinates of particles, then you can talk about the property of the paths- but the path of one particle at a given time is affected by the path of another at a different time. If you try to

describe, therefore, things differentially, telling what the present conditions of the particles are, and how these present conditions will affect the future- you see, it is impossible with particles alone, because something the particle did in the past is going to affect the future.

Therefore, you need a lot of bookkeeping variables to keep track of what the particle did in the past. These are called field variables. You will, also, have to tell what the field is at this present moment, if you are to be able to see later what is going to happen. From the overall space- time view of the least action principle, the field disappears as nothing but bookkeeping variables insisted on by the Hamiltonian method.

As a by-product of this same view, I received a telephone call one day at the graduate college at Princeton from Professor Wheeler, in which he said, « Feynman, I know why all electrons have the same charge and the same mass » « Why? » « Because, they are all the same electron! » And, then he explained on the telephone, « suppose that the world lines which we were ordinarily considering before in time and space-instead of only going up in time were a tremendous knot, and then, when we cut through the knot, by the plane corresponding to a fixed time, we would see many, many world lines and that would represent many electrons, except for one thing. If in one section this is an ordinary electron world line, in the section in which it reversed itself and is coming back from the future we have the wrong sign to the proper time - to the proper four velocities - and that's equivalent to changing the sign of the charge, and, therefore, that part of a path would act like a positron. » « But, Professor », I said, « there aren't as many positrons as electrons. » « Well, maybe they are hidden in the protons or something », he said. I did not take the idea that all the electrons were the same one from him as seriously as I took the observation that positrons could simply be represented as electrons going from the future to the past in a back section of their world lines. That, I stole !

To summarize, when I was done with this, as a physicist I had gained two things. One, I knew many different ways of formulating classical electrodynamics, with many different mathematical forms. I got to know how to express the subject every which way. Second, I had a point ofview-the overall space- time point of view-and a disrespect for the Hamiltonian method of describing physics.

I would like to interrupt here to make a remark. The fact that electrodynamics can be written in so many ways-the differential equations of Maxwell, various minimum principles with fields, minimum principles without fields,

all different kinds of ways,was something I knew, but I have never understood. It always seems odd to me that the fundamental laws of physics, when discovered, can appear in so many different forms that are not apparently identical at first, but, with a little mathematical fiddling you can show the relationship. An example of that is the Schrödinger equation and the Heisenberg formulation of quantum mechanics. I don't know why this is -it remains a mystery, but it was something I learned from experience. There is always another way to say the same thing that doesn't look at all like the way you said it before. I don't know what the reason for this is. I think it is somehow a representation of the simplicity of nature. A thing like the inverse square law is just right to be represented by the solution of Poisson's equation, which, therefore, is a very different way to say the same thing that doesn't look at all like the way you said it before. I don't know what it means, that nature chooses these curious forms, but maybe that is a way of defining simplicity. Perhaps a thing is simple if you can describe it fully in several different ways without immediately knowing that you are describing the same thing.

I was now convinced that since we had solved the problem of classical electrodynamics (and completely in accordance with my program from M. I.T., only direct interaction between particles, in a way that made fields unnecessary) that everything was definitely going to be all right. I was convinced that all I had to do was make a quantum theory analogous to the classical one and everything would be solved.

So, the problem is only to make a quantum theory, which has as its classical analog, this expression (i). Now, there is no unique way to make a quantum theory from classical mechanics, although all the textbooks make believe there is. What they would tell you to do, was find the momentum variables and replace themby(&/i)(a/ax), but I couldn't find a momentum variable, as there wasn't any.

The character of quantum mechanics of the day was to write things in the famous Hamiltonian way - in the form of a differential equation, which described how the wave function changes from instant to instant, and in terms of an operator, H. If the classical physics could be reduced to a Hamiltonian form, everything was all right. Now, least action does not imply a Hamiltonian form if the action is a function of anything more than positions and velocities at the same moment. If the action is of the form of the integral of a function, (usually called the Lagrangian) of the velocities and positions at the same time

$$S = \int L(\dot{x}, x)\, dt \qquad (2)$$

then you can start with the Lagrangian and then create a Hamiltonian and work out the quantum mechanics, more or lessuniquely. But this thing (I) involves the key variables, positions, at two different times and therefore, it was not obvious what to do to make the quantum-mechanical analogue.

I tried - I would struggle in various ways. One of them was this; if I had harmonic oscillators interacting with a delay in time, I could work out what the normal modes were and guess that the quantum theory of the normal modes was the same as for simple oscillators and kind of work my way back in terms of the original variables. I succeeded in doing that, but I hoped then to generalize to other than a harmonic oscillator, but I learned to my regret something, which many people have learned. The harmonic oscillator is too simple; very often you can work out what it should do in quantum theory without getting much of a clue as to how to generalize your results to other systems.

So that didn't help me very much, but when I was struggling with this problem, I went to a beer party in the Nassau Tavern in Princeton. There was a gentleman, newly arrived from Europe (Herbert Jehle) who came and sat next to me. Europeans are much more serious than we are in America because they think that a good place to discuss intellectual matters is a beer party. So, he sat by me and asked, « what are you doing » and so on, and I said, « I'm drinking beer. » Then I realized that he wanted to know what work I was doing and I told him I was struggling with this problem, and I simply turned to him and said, ((listen, do you know any way of doing quantum mechanics, starting withaction - where the action integral comes into the quantum mechanics? » « No », he said, « but Dirac has a paper in which the Lagrangian, at least, comes into quantum mechanics. I will show it to you tomorrow. »

Next day we went to the Princeton Library, they have little rooms on the side to discuss things, and he showed me this paper. What Dirac said was the following : There is in quantum mechanics a very important quantity which carries the wave function from one time to another, besides the differential equation but equivalent to it, a kind of a kernal, which we might call $K(x', x)$, which carries the wave function $\psi(x)$ known at time t, to the wave function $\psi(x')$ at time, $t+\varepsilon$. Dirac points out that this function K was *analogous* to the quantity in classical mechanics that you would calculate if you took the exponential of $i\varepsilon$, multiplied by the Lagrangian $L(\dot{x}, x)$ imagining that these two positions x, x' corresponded t and $t+\varepsilon$. In other words,

$$K(x', x) \text{ is analogous to } e^{i\varepsilon L(\frac{x'-x}{\varepsilon}, x)/\hbar}$$

Professor Jehle showed me this, I read it, he explained it to me, and I said, « what does he mean, they are analogous; what does that mean, *analogous?* What is the use of that? » He said, « you Americans ! You always want to find a use for everything! » I said, that I thought that Dirac must mean that they were equal. « No », he explained, « he doesn't mean they are equal. » « Well », I said, « let's see what happens if we make them equal. »

So I simply put them equal, taking the simplest example where the Lagrangian is $^1/_2 Mx^2 - V(x)$ but soon found I had to put a constant of proportionality A in, suitably adjusted. When I substituted $Ae^{i\varepsilon L/\hbar}$ for K to get

$$\psi(x', t+\varepsilon) = \int A \exp\left[\frac{i\varepsilon}{\hbar} L\left(\frac{x'-x}{\varepsilon}, x\right)\right] \psi(x, t)\, dx \qquad (3)$$

and just calculated things out by Taylor series expansion, out came the Schrödinger equation. So, I turned to Professor Jehle, not really understanding, and said, « well, you see Professor Dirac meant that they were proportional. » Professor Jehle's eyes were bugging out-he had taken out a little notebook and was rapidly copying it down from the blackboard, and said, « no, no,this is an important discovery. You Americans are always trying to find out how something can be used. That's a good way to discover things! » So, I thought I was finding out what Dirac meant, but, as a matter of fact, had made the discovery that what Dirac thought was analogous, was, in fact, equal. I had then, at least, the connection between the Lagrangian and quantum mechanics, but still with wave functions and infinitesimal times.

It must have been a day or so later when I was lying in bed thinking about these things, that I imagined what would happen if I wanted to calculate the wave function at a finite interval later.

I would put one of these factors $e^{i\varepsilon L}$ in here, and that would give me the wave functions the next moment, $t+\varepsilon$ and then I could substitute that back into (3) to get another factor of $e^{i\varepsilon L}$ and give me the wave function the next moment, $t + 2\varepsilon$, and so on and so on. In that way I found myself thinking of a large number of integrals, one after the other in sequence. In the integrand was the product of the exponentials, which, of course, was the exponential of the sum of terms like εL. Now, L is the Lagrangian and ε is like the time interval dt, so that if you took a sum of such terms, that's exactly like an integral. That's like Riemann's formula for the integral $\int L dt$, you just take the value at each point and add them together. We are to take the limit as ε - o, of course. Therefore, the connection between the wave function of one instant and the wave function of another instant a finite time later could be obtained by an infinite number of integrals, (because ε goes to zero, of course) of ex-

ponential $(iS/\hbar)$ where S is the action expression (2). At last, I had succeeded in representing quantum mechanics directly in terms of the action S.

This led later on to the idea of the amplitude for a path; that for each possible way that the particle can go from one point to another in space-time, there's an amplitude. That amplitude is e to the $i/\hbar$ times the action for the path. Amplitudes from various paths superpose by addition. This then is another, a third way, of describing quantum mechanics, which looks quite different than that of Schrödinger or Heisenberg, but which is equivalent to them.

Now immediately after making a few checks on this thing, what I wanted to do, of course, was to substitute the action (1) for the other (2). The first trouble was that I could not get the thing to work with the relativistic case of spin one-half. However, although I could deal with the matter only non-relativistically, I could deal with the light or the photon interactions perfectly well by just putting the interaction terms of (I) into any action, replacing the mass terms by the non-relativistic $(M\dot{x}^2/2)dt$. When the action has a delay, as it now had, and involved more than one time, I had to lose the idea of a wave function. That is, I could no longer describe the program as; given the amplitude for all positions at a certain time to compute the amplitude at another time. However, that didn't cause very much trouble. It just meant developing a new idea. Instead of wave functions we could talk about this; that if a source of a certain kind emits a particle, and a detector is there to receive it, we can give the amplitude that the source will emit and the detector receive. We do this without specifying the exact instant that the source emits or the exact instant that any detector receives, without trying to specify the state of anything at any particular time in between, but by just finding the amplitude for the complete experiment. And, then we could discuss how that amplitude would change if you had a scattering sample in between, as you rotated and changed angles, and so on, without really having any wave functions.

It was also possible to discover what the old concepts of energy and momentum would mean with this generalized action. And, so I believed that I had a quantum theory of classical electrodynamics-or rather of this new classical electrodynamics described by action (1). I made a number of checks. If I took the Frenkel field point of view, which you remember was more differential, I could convert it directly to quantum mechanics in a more conventional way. The only problem was how to specify in quantum mechanics the classical boundary conditions to use only half-advanced and half-retarded solutions. By some ingenuity in defining what that meant, I found that the quantum

mechanics with Frenkel fields, plus a special boundary condition, gave me back this action, (1) in the new form of quantum mechanics with a delay. So, various things indicated that there wasn't any doubt I had everything straightened out.

It was also easy to guess how to modify the electrodynamics, if anybody ever wanted to modify it. I just changed the delta to an f, just as I would for the classical case. So, it was very easy, a simple thing. To describe the old retarded theory without explicit mention of fields I would have to write probabilities, not just amplitudes. I would have to square my amplitudes and that would involve double path integrals in which there are two S's and so forth. Yet, as I worked out many of these things and studied different forms and different boundary conditions. I got a kind of funny feeling that things weren't exactly right. I could not clearly identify the difficulty and in one of the short periods during which I imagined I had laid it to rest, I published a thesis and received my Ph.D.

During the war, I didn't have time to work on these things very extensively, but wandered about on buses and so forth, with little pieces of paper, and struggled to work on it and discovered indeed that there was something wrong, something terribly wrong. I found that if one generalized the action from the nice Langrangian forms (2) to these forms (1) then the quantities which I defined as energy, and so on, would be complex. The energy values of stationary states wouldn't be real and probabilities of events wouldn't add up to 100%. That is, if you took the probability that this would happen and that would happen -everything you could think of would happen, it would not add up to one.

Another problem on which I struggled very hard, was to represent relativistic electrons with this new quantum mechanics. I wanted to do a unique and different way-and not just by copying the operators of Dirac into some kind of an expression and using some kind of Dirac algebra instead of ordinary complex numbers. I was very much encouraged by the fact that in one space dimension, I did find a way of giving an amplitude to every path by limiting myself to paths, which only went back and forth at the speed of light. The amplitude was simple (is) to a power equal to the number ofvelocity reversals where I have divided the time into steps ε and I am allowed to reverse velocity only at such a time. This gives (as ε approaches zero) Dirac's equation in two dimensions-one dimension of space and one of time ($\hbar = M = c = 1$).

Dirac's wave function has four components in four dimensions, but in this case, it has only two components and this rule for the amplitude of a path

automatically generates the need for two components. Because if this is the formula for the amplitudes of path, it will not do you any good to know the total amplitude of all paths, which come into a given point to find the amplitude to reach the next point. This is because for the next time, if it came in from the right, there is no new factor $i\epsilon$ if it goes out to the right, whereas, if it came in from the left there was a new factor $i\epsilon$. So, to continue this same information forward to the next moment, it was not sufficient information to know the total amplitude to arrive, but you had to know the amplitude to arrive from the right and the amplitude to arrive to the left, independently. If you did, however, you could then compute both of those again independently and thus you had to carry two amplitudes to form a differential equation (first order in time).

And, so I dreamed that if I were clever, I would find a formula for the amplitude of a path that was beautiful and simple for three dimensions of space and one of time, which would be equivalent to the Dirac equation, and for which the four components, matrices, and all those other mathematical funny things would come out as a simple consequence-I have never succeeded in that either. But, I did want to mention some of the unsuccessful things on which I spent almost as much effort, as on the things that did work.

To summarize the situation a few years after the way, I would say, I had much experience with quantum electrodynamics, at least in the knowledge of many different ways of formulating it, in terms of path integrals of actions and in other forms. One of the important by-products, for example, of much experience in these simple forms, was that it was easy to see how to combine together what was in those days called the longitudinal and transverse fields, and in general, to see clearly the relativistic invariance of the theory. Because of the need to do things differentially there had been, in the standard quantum electrodynamics, a complete split of the field into two parts, one of which is called the longitudinal part and the other mediated by the photons, or transverse waves. The longitudinal part was described by a Coulomb potential acting instantaneously in the Schrödinger equation, while the transverse part had entirely different description in terms of quantization of the transverse waves. This separation depended upon the relativistic tilt of your axes in space-time. People moving at different velocities would separate the same field into longitudinal and transverse fields in a different way. Furthermore, the entire formulation ofquantum mechanics insisting, as it did, on the wave function at a given time, was hard to analyze relativistically. Somebody else in a different coordinate system would calculate the succession of events in terms of wave

functions on differently cut slices of space- time, and with a different separation of longitudinal and transverse parts. The Hamiltonian theory did not look relativistically invariant, although, of course, it was. One of the great advantages of the overall point of view, was that you could see the relativistic invariance right away-or as Schwinger would say- the covariance was manifest. I had the advantage, therefore, of having a manifestedly covariant form for quantum electrodynamics with suggestions for modifications and so on. I had the disadvantage that if I took it too seriously-I mean, if I took it seriously at all in this form,-1 got into trouble with these complex energies and the failure of adding probabilities to one and so on. I was unsuccessfully struggling with that.

Then Lamb did his experiment, measuring the separation of the $2S\frac{1}{2}$ and $2P\frac{1}{2}$ levels of hydrogen, finding it to be about 1000 megacycles of frequency difference. Professor Bethe, with whom I was then associated at Cornell, is a man who has this characteristic : If there's a good experimental number you've got to figure it out from theory. So, he forced the quantum electrodynamics of the day to give him an answer to the separation of these two levels. He pointed out that the self-energy of an electron itself is infinite, so that the calculated energy of a bound electron should also come out infinite. But, when you calculated the separation of the two energy levels in terms of the corrected mass instead of the old mass, it would turn out, he thought, that the theory would give convergent finite answers. He made an estimate of the splitting that way and found out that it was still divergent, but he guessed that was probably due to the fact that he used an unrelativistic theory of the matter. Assuming it would be convergent if relativistically treated, he estimated he would get about a thousand megacycles for the Lamb-shift, and thus, made the most important discovery in the history of the theory of quantum electrodynamics. He worked this out on the train from Ithaca, New York to Schenectady and telephoned me excitedly from Schenectady to tell me the result, which I don't remember fully appreciating at the time.

Returning to Cornell, he gave a lecture on the subject, which I attended. He explained that it gets very confusing to figure out exactly which infinite term corresponds to what in trying to make the correction for the infinite change in mass. If there were any modifications whatever, he said, even though not physically correct, (that is not necessarily the way nature actually works) but any mod&cation whatever at high frequencies, which would make this correction finite, then there would be no problem at all to figuring out how to keep track of everything. You just calculate the finite mass correc-

tion Δm to the electron mass m_0, substitute the numerical values of $m_0 + \Delta m$ for m in the results for any other problem and all these ambiguities would be resolved. If, in addition, this method were relativistically invariant, then we would be absolutely sure how to do it without destroying relativistically invariant.

After the lecture, I went up to him and told him, « I can do that for you, I'll bring it in for you tomorrow. » I guess I knew every way to modify quantum electrodynamics known to man, at the time. So, I went in next day, and explained what would correspond to the modification of the delta-function to f and asked him to explain to me how you calculate the self-energy of an electron, for instance, so we can figure out if it's finite.

I want you to see an interesting point. I did not take the advice of Professor Jehle to find out how it was useful. I never used all that machinery which I had cooked up to solve a single relativistic problem. I hadn't even calculated the self-energy of an electron up to that moment, and was studying the difficulties with the conservation of probability, and so on, without actually doing anything, except discussing the general properties of the theory.

But now I went to Professor Bethe, who explained to me on the blackboard, as we worked together, how to calculate the self-energy of an electron. Up to that time when you did the integrals they had been logarithmically divergent. I told him how to make the relativistically invariant modifications that I thought would make everything all right. We set up the integral which then diverged at the sixth power of the frequency instead of logarithmically!

So, I went back to my room and worried about this thing and went around in circles trying to figure out what was wrong because I was sure physically everything had to come out finite, I couldn't understand how it came out infinite. I became more and more interested and finally realized I had to learn how to make a calculation. So, ultimately, I taught myself how to calculate the self-energy of an electron working my patient way through the terrible confusion of those days of negative energy states and holes and longitudinal contributions and so on. When I finally found out how to do it and did it with the modifications I wanted to suggest, it turned out that it was nicely convergent and finite, just as I had expected. Professor Bethe and I have never been able to discover what we did wrong on that blackboard two months before, but apparently we just went off somewhere and we have never been able to figure out where. It turned out, that what I had proposed, if we had carried it out without making a mistake would have been all right and would have given a finite correction. Anyway, it forced me to go back over all this and to

convince myself physically that nothing can go wrong. At any rate, the correction to mass was now finite, proportional to ln (mu /$\hbar$) where a is the width of that function f which was substituted for δ. If you wanted an unmodified electrodynamics, you would have to take a equal to zero, getting an infinite mass correction. But, that wasn't the point. Keeping a finite, I simply followed the program outlined by Professor Bethe and showed how to calculate all the various things, the scatterings of electrons from atoms without radiation, the shifts of levels and so forth, calculating everything in terms of the experimen - tal mass, and noting that the results as Bethe suggested, were not sensitive to a in this form and even had a definite limit as $a \rightarrow$ o.

The rest of my work was simply to improve the techniques then available for calculations, making diagrams to help analyze perturbation theory quicker. Most of this was first worked out by guessing-you see, I didn't have the relativistic theory of matter. For example, it seemed to me obvious that the velocities in non-relativistic formulas have to be replaced by Dirac's matrix α or in the more relativistic forms by the operators γ_μ. I just took my guesses from the forms that I had worked out using path integrals for non-relativistic matter, but relativistic light. It was easy to develop rules of what to substitute to get the relativistic case. I was very surprised to discover that it was not known at that time, that every one of the formulas that had been worked out so patiently by separating longitudinal and transverse waves could be obtained from the formula for the transverse waves alone, if instead of summing over only the two perpendicular polarization directions you would sum over all four possible directions of polarization. It was so obvious from the action (i) that I thought it was general knowledge and would do it all the time. I would get into arguments with people, because I didn't realize they didn't know that; but, it turned out that all their patient work with the longitudinal waves was always equivalent to just extending the sum on the two transverse directions of polarization over all four directions. This was one of the amusing advantages of the method. In addition, I included diagrams for the various terms of the perturbation series, improved notations to be used, worked out easy ways to evaluate integrals, which occurred in these problems, and so on, and made a kind of handbook on how to do quantum electrodynamics.

But one step of importance that was physically new was involved with the negative energy sea of Dirac, which caused me so much logical difficulty. I got so confused that I remembered Wheeler's old idea about the positron being, maybe, the electron going backward in time. Therefore, in the time depen-

dent perturbation theory that was usual for getting self-energy, I simply supposed that for a while we could go backward in the time, and looked at what terms I got by running the time variables backward. They were the same as the terms that other people got when they did the problem a more complicated way, using holes in the sea, except, possibly, for some signs. These, I, at first, determined empirically by inventing and trying some rules.

I have tried to explain that all the improvements of relativistic theory were at first more or less straightforward, semi-empirical shenanigans. Each time I would discover something, however, I would go back and I would check it so many ways, compare it to every problem that had been done previously in electrodynamics (and later, in weak coupling meson theory) to see if it would always agree, and so on, until I was absolutely convinced of the truth of the various rules and regulations which I concocted to simplify all the work.

During this time, people had been developing meson theory, a subject I had not studied in any detail. I became interested in the possible application of my methods to perturbation calculations in meson theory. But, what was meson theory? All I knew was that meson theory was something analogous to electrodynamics, except that particles corresponding to the photon had a mass. It was easy to guess the δ– function in (I), which was a solution of d'Alembertian equals zero, was to be changed to the corresponding solution of d'Alembertian equals m^2. Next, there were different kind of mesons-the one in closest analogy to photons, coupled via $\gamma_\mu\gamma_\mu$, are called vector mesons- there were also scalar mesons. Well, maybe that corresponds to putting unity in place of the γ_μ, I would here then speak of « pseudo vector coupling » and I would guess what that probably was. I didn't have the knowledge to understand the way these were defined in the conventional papers because they were expressed at that time in terms of creation and annihilation operators, and so on, which, I had not successfully learned. I remember that when someone had started to teach me about creation and annihilation operators, that this operator creates an electron, I said, « how do you create an electron? It disagrees with the conservation of charge », and in that way, I blocked my mind from learning a very practical scheme of calculation. Therefore, I had to find as many opportunities as possible to test whether I guessed right as to what the various theories were.

One day a dispute arose at a Physical Society meeting as to the correctness of a calculation by Slotnick of the interaction of an electron with *a* neutron using pseudo scalar theory with pseudo vector coupling and also, pseudo scalar theory with pseudo scalar coupling. He had found that the answers were not

the same, in fact, by one theory, the result was divergent, although convergent with the other. Some people believed that the two theories must give the same answer for the problem. This was a welcome opportunity to test my guesses as to whether I really did understand what these two couplings were. So, I went home, and during the evening I worked out the electron neutron scattering for the pseudo scalar and pseudo vector coupling, saw they were not equal and subtracted them, and worked out the difference in detail. The next day at the meeting, I saw Slotnick and said, « Slotnick, I worked it out last night, I wanted to see if I got the same answers you do. I got a different answer for each coupling-but, I would like to check in detail with you because I want to make sure of my methods. » And, he said, « what do you mean you worked it out last night, it took me six months ! » And, when wecompared the answers he looked at mine and he asked, « what is that Q in there, that variable Q? » (I had expressions like (tan -'Q) /Q etc.). I said, « that's the momentum transferred by the electron, the electron deflected by different angles. » « Oh », he said, « no, I only have the limiting value as Q approaches zero; the forward scattering. » Well, it was easy enough to just substitute Q equals zero in my form and I then got the same answers as he did. But, it took him six months to do the case of zero momentum transfer, whereas, during one evening I had done the finite and arbitrary momentum transfer. That was a thrilling moment for me, like receiving the Nobel Prize, because that convinced me, at last, I did have some kind of method and technique and understood how to do something that other people did not know how to do. That was my moment of triumph in which I realized I really had succeeded in working out something worthwhile.

At this stage, I was urged to publish this because everybody said it looks like an easy way to make calculations, and wanted to know how to do it. I had to publishit, missing two things; one was proof of every statement in a mathematically conventional sense. Often, even in a physicist's sense, I did not have a demonstration of how to get all of these rules and equations from conventional electrodynamics. But, I did know from experience, from fooling around, that everything was, in fact, equivalent to the regular electrodynamics and had partial proofs of many pieces, although, I never really sat down, like Euclid did for the geometers of Greece, and made sure that you could get it all from a single simple set of axioms. As a result, the work was criticized, I don't know whether favorably or unfavorably, and the « method » was called the aintuitive method)). For those who do not realize it, however, I should like to emphasize that there is a lot of work involved in using this <<intuitive

method>> successfully. Because no simple clear proof of the formula or idea presents itself, it is necessary to do an unusually great amount of checking and rechecking for consistency and correctness in terms of what is known, by comparing to other analogous examples, limiting cases, etc. In the face of the lack of direct mathematical demonstration, one must be careful and thorough to make sure of the point, and one should make a perpetual attempt to demonstrate as much of the formula as possible. Nevertheless, a very great deal more truth can become known than can be proven.

It must be clearly understood that in all this work, I was representing the conventional electrodynamics with retarded interaction, and not my half-advanced and half-retarded theory corresponding to (1). I merely use (1) to guess at forms. And, one of the forms I guessed at corresponded to changing δ to a function f of width a^2, so that I could calculate finite results for all of the problems. This brings me to the second thing that was missing when I published the paper, an unresolved difficulty. With δ replaced by f the calculations would give results which were not « unitary », that is, for which the sum of the probabilities of all alternatives was not unity. The deviation from unity was very small, in practice, if a was very small. In the limit that I took a very tiny, it might not make any difference. And, so the process of the renormalization could be made, you could calculate everything in terms of the experimental mass and then take the limit and the apparent difficulty that the unitary is violated temporarily seems to disappear. I was unable to demonstrate that, as a matter of fact, it does.

It is lucky that I did not wait to straighten out that point, for as far as I know, nobody has yet been able to resolve this question. Experience with meson theories with stronger couplings and with strongly coupled vector photons, although not proving anything, convinces me that if the coupling were stronger, or if you went to a higher order (137th order of perturbation theory for electrodynamics), this difficulty would remain in the limit and there would be real trouble. That is, I believe there is really no satisfactory quantum electrodynamics, but I'm not sure. And, I believe, that one of the reasons for the slowness of present-day progress in understanding the strong interactions is that there isn't any relativistic theoretical model, from which you can really calculate everything. Although, it is usually said, that the difficulty lies in the fact that strong interactions are too hard to calculate, I believe, it is really because strong interactions in field theory have no solution, have no sense-they're either infinite, or, if you try to modify them, the modification destroys the unitarity. I don't think we have a completely satisfactory relativistic quan-

turn- mechanical model, even one that doesn't agree with nature, but, at least, agrees with the logic that the sum of probability of all alternatives has to be 100%. Therefore, I think that the renormalization theory is simply a way to sweep the difficulties of the divergences of electrodynamics under the rug. I am, of course, not sure of that.

This completes the story of the development of the space-time view of quantum electrodynamics. I wonder if anything can be learned from it. I doubt it. It is most striking that most of the ideas developed in the course of this research were not ultimately used in the final result. For example, the half-advanced and half-retarded potential was not finally used, the action expression (1) was not used, the idea that charges do not act on themselves was abandoned. The path-integral formulation of quantum mechanics was useful for guessing at final expressions and at formulating the general theory of electrodynamics in new ways-although, strictly it was not absolutely necessary. The same goes for the idea of the positron being a backward moving electron, it was very convenient, but not strictly necessary for the theory because it is exactly equivalent to the negative energy sea point of view.

We are struck by the very large number of different physical viewpoints and widely different mathematical formulations that are all equivalent to one another. The method used here, ofreasoning in physical terms, therefore, appears to be extremely inefficient. On looking back over the work, I can only feel a kind of regret for the enormous amount of physical reasoning and mathematically re-expression which ends by merely re-expressing what was previously known, although in a form which is much more efficient for the calculation of specific problems. Would it not have been much easier to simply work entirely in the mathematical framework to elaborate a more efficient expression? This would certainly seem to be the case, but it must be remarked that although the problem actually solved was only such a reformulation, the problem originally tackled was the (possibly still unsolved) problem of avoidante of the inifinities of the usual theory. Therefore, a new theory was sought, not just a modification of the old. Although the quest was unsuccessful, we should look at the question of the value of physical ideas in developing a new theory.

Many different physical ideas can describe the same physical reality. Thus, classical electrodynamics can be described by a field view, or an action at a distance view, etc. Originally, Maxwell filled space with idler wheels, and Faraday with fields lines, but somehow the Maxwell equations themselves are

pristine and independent of the elaboration of words attempting a physical description. The only true physical description is that describing the experimental meaning of the quantities in the equation-or better, the way the equations are to be used in describing experimental observations. This being the case perhaps the best way to proceed is to try to guess equations, and disregard physical models or descriptions. For example, McCullough guessed the correct equations for light propagation in a crystal long before his colleagues using elastic models could make head or tail of the phenomena, or again, Dirac obtained his equation for the description of the electron by an almost purely mathematical proposition. A simple physical view by which all the contents of this equation can be seen is still lacking.

Therefore, I think equation guessing might be the best method to proceed to obtain the laws for the part of physics which is presently unknown. Yet, when I was much younger, I tried this equation guessing and I have seen many students try this, but it is very easy to go off in wildly incorrect and impossible directions. I think the problem is not to find the best or most efficient method to proceed to a discovery, but to find any method at all. Physical reasoning does help some people to generate suggestions as to how the unknown may be related to the known. Theories of the known, which are described by different physical ideas may be equivalent in all their predictions and are hence scientifically indistinguishable. However, they are not psychologically identical when trying to move from that base into the unknown. For different views suggest different kinds of modifications which might be made and hence are not equivalent in the hypotheses one generates from them in ones attempt to understand what is not yet understood. I, therefore, think that a good theoretical physicist today might find it useful to have a wide range of physical viewpoints and mathematical expressions of the same theory (for example, of quantum electrodynamics) available to him. This may be asking too much of one man. Then new students should as a class have this. If every individual student follows the same current fashion in expressing and thinking about electrodynamics or field theory, then the variety of hypotheses being generated to understand strong interactions, say, is limited. Perhaps rightly so, for possibly the chance is high that the truth lies in the fashionable direction. But, on the off-chance that it is in another direction-a direction obvious from an unfashionable view of field theory-who will find it? Only someone who has sacrificed himself by teaching himself quantum electrodynamics from a peculiar and unusual point of view; one that he may have to invent for himself. I say sacrificed himself because he most likely will get

nothing from it, because the truth may lie in another direction, perhaps even the fashionable one.

But, if my own experience is any guide, the sacrifice is really not great because if the peculiar viewpoint taken is truly experimentally equivalent to the usual in the realm of the known there is always a range of applications and problems in this realm for which the special viewpoint gives one a special power and clarity of thought, which is valuable in itself. Furthermore, in the search for new laws, you always have the psychological excitement of feeling that possible nobody has yet thought of the crazy possibility you are looking at right now.

So what happened to the old theory that I fell in love with as a youth? Well, I would say it's become an old lady, that has very little attractive left in her and the young today will not have their hearts pound when they look at her anymore. But, we can say the best we can for any old woman, that she has been a very good mother and she has given birth to some very good children. And, I thank the Swedish Academy of Sciences for complimenting one of them. Thank you.

II.B *Action-at-a-Distance Classical Electrodynamics*

While still an undergraduate at MIT, Feynman became aware of the so-called divergence problems of QED. That is, physical quantities which should have been calculable by the theory, such as the self-mass of the electron (the effect of the action of the electron's own electromagnetic field on its mass), were predicted to have an absurd result: infinity. Feynman knew that a similar result was predicted classically; namely, the energy contained in the Coulomb field of a point charge is theoretically infinite. As he said in his Nobel Lecture, his "general plan was to first solve the classical problem, to get rid of the infinite self-energies and to hope that when I made a quantum theory of it, everything would just be fine." The idea which he embraced ("fell deeply in love with") was to replace the field itself by "delayed action-at-a-distance." In this view the electron would act only on *other charges*, not on itself, and the field would be only a useful invention for representing that delayed interaction. He abandoned this idea after an accurate experimental value obtained for the Lamb shift in hydrogen in the late 1940s showed the presence of an effect called "vacuum polarization," which could only be obtained by using the full field concept.[1]

Papers [4] and [10] are, respectively, Part 3 and Part 2 of a three-part paper projected by John Archibald Wheeler and Richard Phillips Feynman. Part 1, never published (and probably never written), was to have been a careful study of the classical limit of the quantum theory of radiation. Paper [4] introduces the absorber theory, according to which half of the electromagnetic field propagates *before* the electron emitting it accelerates (advanced) and half as it accelerates (retarded). The advanced field is assumed to be absorbed in distant matter, where it would reradiate and arrive at the accelerating electron at the right time and in the right amount to produce the "radiation reaction" that is needed to reduce the radiating electron's kinetic energy by the amount of energy that it is radiates. There are no *observable* "advanced effects." That was the solution to the problem of lack of energy conservation that would result (as Wheeler had pointed out to Feynman) if the electron's radiated field did not act back upon the electron. Edward Kerner has remarked that the "'complete absorption' in the electromagnetic universe is a kind of electrodynamic Mach's principle accounting marvelously for the appearance of the Lorentz–Dirac force of radiation damping, and for the appearance of retarded interactions on the local scene."[2]

Paper [10], "Classical Electrodynamics in Terms of Direct Interparticle Interaction," was published in 1949 while Feynman was deeply involved in his work on *quantum* electrodynamics. Like [4] it is a scholarly paper, published in the *Reviews of Modern Physics*. Remarking on it in an interview, Feynman said, "That was written by Wheeler, and was done essentially independently. We worked together."[3] By this he meant that the contents were worked out jointly with Wheeler, who did the actual writing. The paper continues the critique of classical electrodynamics begun in [4], based upon this idea: The field of a charge is determined by its motion; its field is only sensed by its action on other charges, whose motions act back upon the first charge. Thus it should be possible to eliminate the field and to discuss directly

[1] In a letter to John Wheeler in 1951, Feynman wrote, "I wish to deny the correctness of the assumption that" electrons act only on other electrons, citing two pieces of evidence, one being the Lamb shift. He concluded the letter thus: "So I think we guessed wrong in 1941. Do you agree?" (Feynman to Wheeler, May 4, 1951).
[2] E.H. Kerner, ed., *The Theory of Action-at-a-Distance in Relativistic Particle Dynamics*. New York, 1972, pp. viii–ix.
[3] Interview of Feynman by Charles Weiner, June 27, 1966, p. 39.

how the motion of one charge affects the motion of another. This can be done by writing the relativistic expression for the principle of least action, which determines the equations of motion of the charges (Fokker's action principle). The last requires the use of half-advanced and half-retarded four-vector potentials, and this leads to a discussion on the "paradox of advanced effects."

Paper [8] is included in this section because it is a further step in Feynman's plan to modify classical electrodynamics as a forerunner to attacking the problems of QED. Again it uses the action-at-a-distance approach, half-advanced and half-retarded interaction, and the Fokker action principle, although Feynman points out that his modification of the classical "pointlike" interaction could also be applied to the conventional electrodynamics. However, the latter makes use of the Hamiltonian method that singles out the time as a preferred variable, making it difficult to construct a relativistic theory, which is more symmetrical in time and space. Indeed, using the Hamiltonian requires keeping track of an infinity of variables on a plane of constant time in space–time (or on a spacelike surface). That would be at least as complicated as the field concept that Feynman is trying to eliminate.

The solution invoked by H.A. Lorentz to the classical electron self-energy problem at the turn of the century was to give the electron a finite size. In [8], Feynman introduces an equivalent relativistic "cut-off" that spreads out the interaction, conventionally occurring only on the light-cone ("pointlike" interaction), over a small timelike interval. This finite interval can be chosen as small as one wishes, and it would in principle be possible to determine it experimentally at high energy. The paper also discusses the least action solution to the problem of an electron striking a barrier and penetrating it, either directly, or indirectly by a process involving the production of a virtual positron electron pair. It introduces the forerunner of the Feynman diagrams, containing an electron moving "backward in time" to represent the positron!

Selected Papers

[4] With J.A. Wheeler. Interaction with the absorber as the mechanism of radiation, *Rev. Mod. Phys.* **17** (1945): 157–181.

[10] With J.A. Wheeler. Classical electrodynamics in terms of direct interparticle action. *Rev. Mod. Phys.* **21** (1949): 425-433.

[8] A relativistic cut-off for classical electrodynamics. *Phys. Rev.* **74** (1948): 939–946.

REVIEWS OF MODERN PHYSICS VOLUME 17, NUMBERS 2 AND 3 APRIL-JULY, 1945

Interaction with the Absorber as the Mechanism of Radiation†*

JOHN ARCHIBALD WHEELER** AND RICHARD PHILLIPS FEYNMAN***
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

"We must, therefore, be prepared to find that further advance into this region will require a still more extensive renunciation of features which we are accustomed to demand of the space time mode of description."—Niels Bohr[1]

PAST FAILURE OF ACTION AT A DISTANCE TO ACCOUNT FOR THE MECHANISM OF RADIATION

IT was the 19th of March in 1845 when Gauss described the conception of an action at a distance propagated with a finite velocity, the natural generalization to electrodynamics of the view of force so fruitfully applied by Newton and his followers. In the century between then and now what obstacle has discouraged the general use of this conception in the study of nature?

The difficulty has not been that of giving to the idea of propagated action at a distance a suitable embodiment of electromagnetic equations. This problem, to be true, remained unsolved to Gauss and his successors for three quarters of the century. But the formulation then developed by Schwarzschild and Fokker, described and amplified in another article,[2] demonstrated that the conception of Gauss is at the same time mathematically self consistent, in agreement with experience on static and current electricity, and in complete harmony with Maxwell's equations.

To find the real obstacle to acceptance of the tool of Newton and Gauss for the analysis of forces, we have to go beyond the bounds of steady-state electromagnetism to the phenomena of emission and propagation of energy. No branch of science has done more than radiation physics to favor the evolution of present concepts of field or more to pose difficulties for the idea of action at a distance. The difficulties have been twofold—to obtain a satisfactory account of the field generated by an accelerated charge at a

* A preliminary account of the considerations which appear in this paper was presented by us at the Cambridge meeting of the American Physical Society, February 21, 1941, Phys. Rev. 59, 683 (1941)

** On leave of absence from Princeton University.

*** Now a member of the faculty of Cornell University, but on leave of absence from that institution.

† *Introductory Note.*—In commemoration of the sixtieth birthday of Niels Bohr it had been hoped to present a critique of classical field theory which has been in preparation since before the war by the writer and his former student, R. P. Feynman. The accompanying joint article, representing the third part of the survey, is however the only section now finished. The war has postponed completion of the other parts. As reference to them is made in the present section, it may be useful to outline the plan of the survey.

The motive of the analysis is to clear the present quantum theory of interacting particles of those of its difficulties which have a purely classical origin. The method of approach is to define as closely as one can within the bounds of classical theory the proper use of the field concept in the description of nature. Division I is intended first to recall the possibility of idealizing to the case of arbitrarily small quantum effects, a possibility which is offered by the freedom of choice in the present quantum theory for the dimensionless ratio (quantum of angular momentum)(velocity of light)/(electronic charge)^2; then however to recognize the possible limitations placed on this analysis by the relatively large value, 137, of the ratio in question in nature; and finally to present a general summary of the conclusions drawn from the more technical parts of the survey. The plan of the second article is a derivation and resumé of the theory of action at a distance of Schwarzschild and Fokker, to prepare this theory as a tool to analyze the field concept. From the correlation of the two points of view, one comes to Frenkel's solution of the problem of self-energy in the classical field theory and to new expressions for the energy of electromagnetic interaction in the theory of action at a distance. The third division, which is published herewith, is an analysis of the mechanism of radiation believed to complete the last tie between action at a distance and field theory and to remove the obstacle which has so far prevented the use of both points of view as complementary tools in the description of nature. It is the plan of a subsequent division to discuss the problems which arise when the fields are regarded as subordinate entities with no degrees of freedom of their own. An infinite number of degrees of freedom are found to be attributed to the particles themselves by the theory of propagated action at a distance. However, it appears that the additional modes of motion are divergent and have on this account to be excluded by a general principle of selection. Acceptance of this principle leads to the conclusion that the union of action at a distance and field theory constitutes the natural and self-consistent generalization of Newtonian mechanics to the four-dimensional space of Lorentz and Einstein.—J. A. W.

[1] Niels Bohr, *Atomic Theory and the Description of Nature* (Cambridge University Press, Teddington, England, 1934).

[2] Unpublished, see Introductory Note.

remote point and to understand the source of the force experienced by the charge itself as a result of its motion:

(a) An accelerated charge generates a field given, according to the formulation of Schwarzschild and Fokker, by half the usual retarded solution of Maxwell's equations, plus half the advanced solution. From the presence of the advanced field in the expression for the electric vector, it follows that a distant test body will experience a premonitory force well before the source itself has commenced to move. To avoid a conclusion so opposed to experience Ritz[3] and Tetrode[4] proposed to abandon the symmetry in time of the elementary law of force. However, it was then necessary to give up the possibility to derive the equations of motion and all the electromagnetic forces consistently from a single unified principle of least action like that of Fokker. More important, the sacrifice made to alleviate one difficulty of the theory of action at a distance did not help to solve the other, the problem of the origin of the force of radiative reaction.

(b) Experience indicates that an accelerated charge suffers a force of damping which is simultaneous with the moment of acceleration. However, the theory of action at a distance predicts that an accelerated charge in otherwise charge-free space will experience no electric force. To exclude the acceleration and thus to avoid the issue does not appear reasonable. Uncharged particles can be present and can accelerate the charge via gravitational forces. It seems just as difficult to explain the reactive force when other charged particles are present. They will indeed be set into motion and will act back on the source. However, if these elementary interactions have the purely retarded character assumed by Ritz, and also by Frenkel,[5] the reaction will arrive at the accelerated particle too late and will have the wrong magnitude[6] to produce the damping phenomenon. On the other hand, interactions symmetrical between past and future—the half-retarded, half-advanced fields of the unified theory of action at a distance—have so far appeared to be equally incapable of accounting for the observed force of radiative reaction, with its definitely irreversible character.

It is clear why the viewpoint of Newton and Gauss has not been generally applied in recent times; it has so far failed to give a satisfactory account of the mechanism of radiation.

The failure of action at a distance cannot pass unnoticed by field theory. The two points of view, according to the thesis of the present critique, are not independent, but mutually complementary. Consequently field theory, too, faces in the radiation problem a significant issue: does this theory give an explanation for the observed force of radiative reaction which can be translated into the particle mechanics of Schwarzschild and Fokker, or does it likewise fail to provide a complete picture of the mechanism of radiation?

In attacking the radiation problem our first move, following the above reasoning, is to review the status of the reaction force in existing classical field theory. No more intelligible clue is found to the physical origin of the force in this theory than in the theory of action at a distance. Stopped on this approach, we take up a suggestion made long ago by Tetrode that the act of radiation should have some connection with the presence of an absorber. We develop this idea into the thesis that the force of radiative reaction arises from the action on the source owing to the half-advanced fields of the particles of the absorber; or, more briefly, that radiation is a matter as much of statistical mechanics as of pure electrodynamics. We find that this thesis leads to a quantitative solution of the radiation problem. Finally we examine some of the implications of this thesis for the conception of causality.

THE STATUS OF RADIATIVE REACTION IN FIELD THEORY

A charged particle on being accelerated sends out electromagnetic energy and itself loses energy. This loss is interpreted as caused by a force acting on the particle given in magnitude and direction by the expression

$$\frac{2\,(\text{charge})^2\,(\text{time rate of change of acceleration})}{3\,(\text{velocity of light})^3},$$

when the particle is moving slowly, and by a more complicated expression when its speed is appreciable relative to the velocity of light. The existence of this force of radiative reaction is well attested: (a) by the electrical potential required to drive a wireless antenna; (b) by the loss of energy experienced by a charged particle which has been deflected, and therefore accelerated, in its passage near an atomic nucleus; and (c) by the cooling suffered by a glowing body.

The origin of the force of radiative reaction has not been nearly so clear as its existence.

[3] W. Ritz, Ann. d. Chem. et d. Physique 13, 145 (1908).
[4] H. Tetrode, Zeits. f. Physik 10, 317 (1922).
[5] J. Frenkel, Zeits. f. Physik 32, 518 (1925).
[6] J. L. Synge, Proc. Roy. Soc. London A177, 118 (1940).

Lorentz[7] considers the charged particle to have a finite size and attributes the force in question to the retarded action of one part of the particle on another. His expression for the force appears as a series in powers of the radius of the particle. The first term in the series gives the expression already mentioned. Otherwise, the derivation leads to difficulties:

(a) All higher terms depend explicitly upon the structure assumed for the entity. These dubious terms enter in a more and more important way into the calculated law of radiative reaction as the frequency of oscillation of the particle is raised, and the period approaches the time required for light to cross the system.

(b) Non-electric forces are required to hold together the charge distribution, according to Poincaré,[8] for to neglect such forces is to violate the relativistic relation between mass and energy. A composite system of this kind would possess an infinite number of internal degrees of freedom of oscillation. No consistent model has been found for the Lorentz electron in either classical or quantum mechanics.

Briefly, Lorentz attempts to propose a physical mechanism behind the radiative reaction, but arrives at a mathematically incomplete expression for this force.

Dirac,[9] in contrast, advances no explanation for the origin of the radiative damping, but supplies a well-defined and relativistically invariant prescription to calculate its magnitude:

Let the motion of the particle be given. Calculate the field produced by the particle from Maxwell's equations, with the boundary condition that at large distances from the particle this field shall contain only outgoing waves. In addition to the so-defined retarded field of the particle, calculate its advanced field, the sole change being the existence of only convergent waves at large distances. Define half the difference between retarded and advanced fields as the radiation field (half the quantity denoted as radiation field by Dirac). This field is everywhere finite. Evaluate it at the position of the particle and multiply by the magnitude of the charge to obtain the force of radiative reaction.

Dirac's prescription is appealing. (a) It is well-defined. (b) The calculated force reduces for slowly moving particles to the simple expression which was given above and which has been well-tested at non-relativistic velocities. (c) The calculation treats the elementary charge as being localized at a mathematical point, a picture which is not only physically reasonable but also translatable into quantum mechanics. (d) The elements of the prescription involve no more than standard electromagnetic theory plus the assumption that the radiation field, as above defined, is the source of the force.

The physical origin of Dirac's radiation field is nevertheless not clear. (a) This field is defined for times before as well as after the moment of acceleration of the particle. (b) The field has no singularity at the position of the particle and by Maxwell's equations must, therefore, be attributed either to sources other than the charge itself or to radiation coming in from an infinite distance.

We accept as reasonable Dirac's results. His concept of radiation field, however, we cannot adopt as an assumption subject to no further analysis. To do so would be to add to field theory a principle incapable of translation into the language of action at a distance.

To carry the analysis further requires us to find a new idea. We go back to a suggestion once made by Tetrode.[10] He proposed to abandon the conception of electromagnetic radiation as an elementary process and to interpret it as a consequence of an interaction between a source and an absorber. In his words,

"The sun would not radiate if it were alone in space and no other bodies could absorb its radiation. . . . If for example I observed through my telescope yesterday evening that star which let us say is 100 light years away, then not only did I know that the light which it allowed to reach my eye was emitted 100 years ago, but also the star or individual atoms of it knew already 100 years ago that I, who then did not even exist, would view it yesterday evening at such and such a time. . . . One might accordingly adopt the opinion that the amount of material in the universe determines the rate of emission. Still this is not necessarily so, for two competing absorption centers

[7] H. A. Lorentz (1892), republished in his *Collected Papers*, Vol. II, pp. 281 and 343. See also his treatise *The Theory of Electrons* (Leipzig, 1909), pp. 49 and 253.

[8] H. Poincaré, Rend. Palermo 21, 165 (1906).

[9] P. A. M. Dirac, Proc. Roy. Soc. London A167, 148 (1938).

[10] H. Tetrode, Zeits. f. Physik 10, 317 (1922). When we gave a preliminary account of the considerations which appear in this paper (Cambridge meeting of the American Physical Society, February 21, 1941, Phys. Rev. 59, 683 (1941)) we had not seen Tetrode's paper. We are indebted to Professor Einstein for bringing to our attention the ideas of Tetrode and also of Ritz, who is cited in this article. An idea similar to that of Tetrode was subsequently proposed by G. N. Lewis, Nat. Acad. Sci. Proc. 12, 22 (1926): "I am going to make the . . . assumption that an atom never emits light except to another atom, and to claim that it is as absurd to think of light emitted by one atom regardless of the existence of a receiving atom as it would be to think of an atom absorbing light without the existence of light to be absorbed. I propose to eliminate the idea of mere emission of light and substitute the idea of *transmission*, or a process of exchange of energy between two definite atoms or molecules." Lewis went nearly as far as it is possible to go without explicitly recognizing the importance of other absorbing matter in the system, a point touched upon by Tetrode, and shown below to be essential for the existence of the normal radiative mechanism.

will not collaborate but will presumably interfere with each other. If only the amount of matter is great enough and is distributed to some extent in all directions, further additions to it may well be without influence."

Tetrode's idea that the absorber may be an essential element in the mechanism of radiation has been neglected, perhaps partly because it appears to conflict with customary notions of causality, and partly also because of his mistaken belief that the new point of view could by itself explain quantum phenomena. In this connection he assumed that the interaction between charged particles should be described by forces more complicated than those given by electromagnetic theory. Finally, as Tetrode remarks, "on the last pages we have let our conjectures go rather far beyond what has mathematically been proven."

ABSORBER RESPONSE AS THE MECHANISM OF RADIATIVE REACTION

We take up the proposal of Tetrode that the absorber may be an essential element in the mechanism of radiation. Using the language of the theory of action at a distance, we give the idea the following definite formulation:

(1) An accelerated point charge in otherwise charge-free space does not radiate electromagnetic energy.

(2) The fields which act on a given particle arise only from other particles.

(3) These fields are represented by one-half the retarded plus one-half the advanced Liénard-Wiechert solutions of Maxwell's equations. This law of force is symmetric with respect to past and future. In connection with this assumption we may recall an inconclusive but illuminating discussion carried on by Ritz and Einstein in 1909, in which "Ritz treats the limitation to retarded potentials as one of the foundations of the second law of thermodynamics, while Einstein believes that the irreversibility of radiation depends exclusively on considerations of probability."[11] Tetrode, himself, like Ritz, was willing to assume elementary interactions which were not symmetric in time. However, complete reversibility is assumed here because it is an essential element in a unified theory of action at a distance. In proceeding on the basis of this symmetrical law of interaction, we shall be testing not only Tetrode's idea of absorber reaction, but also Einstein's view that the one-sidedness of the force of radiative reaction is a purely statistical phenomenon. This point leads to our final assumption:

(4) Sufficiently many particles are present to absorb completely the radiation given off by the source.

On the basis of these assumptions we shall consider as the source of radiation an accelerated charge located in the absorbing system. A disturbance travels outward from the source. By it each particle of the absorber is set in motion and caused to generate a field, half-advanced and half-retarded. The sum of the advanced effects of all particles of the absorber, evaluated in the neighborhood of the source, gives a field which we find to have the following properties:

(1) It is independent of the properties of the absorbing medium.

(2) It is completely determined by the motion of the source.

(3) It exerts on the source a force which is finite, is simultaneous with the moment of acceleration, and is just sufficient in magnitude and direction to take away from the source the energy which later shows up in the surrounding particles.

(4) It is equal in magnitude to one-half the retarded field minus one-half the advanced field generated by the accelerated charge. In other words, the absorber is the physical origin of Dirac's radiation field.

(5) This field combines with the half-retarded, half-advanced field of the source to give for the total disturbance the full retarded field which accords with experience.

It will be sufficient to establish these results in order to have both in field theory and in the theory of action at a distance a solution of the problem of radiation, including an explanation of the force of radiative damping.

We shall present four derivations of the reaction of radiation on the source of successively increasing generality. In the first we consider an absorber in which the particles are far from one another. We assume without proof that the disturbance which passes through the medium is the full retarded field of experience. In the second derivation we examine the field of the absorber in the neighborhood of the source and find it just such as to compensate the advanced field of the accelerated charge and to give a retarded field of the previously assumed magnitude. In this case we have allowed the medium to have arbitrary density. The third derivation —in contrast to the first two, where the source was taken to be at rest or moving only slowly— considers the case of motion with arbitrary velocity and leads to the same relativistic expression which Dirac has given for the force of radiative reaction. All three treatments proceed by adding up the fields owing to the individual particles of the absorber. A fourth derivation uses a much more general approach, assuming only that the medium is a complete absorber.

[11] W. Ritz and A. Einstein, Physik. Zeits. **10**, 323 (1909); see also W. Ritz, Ann. d. Chemie et d. Physique **13**, 145 (1908).

THE RADIATIVE REACTION: DERIVATION I

For a first analysis of the mechanism of radiative reaction, we shall simplify as much as possible the properties of the absorber:

(a) it is taken to be composed of free-charged particles;

(b) these corpuscles are at rest or are moving only slowly with respect to the particle which we treat as the source;

(c) the charged entities are well separated from one another;

(d) the particles occupy space to distances sufficiently great to bring about essentially complete absorption of radiation from the source.

We begin by considering the reaction set up between the source and a typical charge in the absorber when the particle of the source receives an acceleration $\mathfrak{A}$, by collision with a third particle or otherwise. The source has a charge $+e$ and, therefore, sends out an electromagnetic disturbance. This effect traverses the distance r_k to the particle of the absorber, reaching it at a time (r_k/c) seconds later than the instant of acceleration. For the electric field acting on the absorber at this place and time, we adopt the usual retarded solution of Maxwell's equation, in conformity with experience, but without any attempt in this first derivation of the force of radiative reaction to reconcile such an assumption with the half-retarded, half-advanced field of the theory of action at a distance. At the distances in which we are interested, the retarded field of the source reduces to the well-known expression,

$$-(e\mathfrak{A}/r_kc^2)\sin(\mathfrak{A}, r_k), \qquad (1)$$

together with a term of electrostatic origin. This second term falls off inversely as the square of the distance and may, therefore, be neglected. The electric vector lies in the plane defined by the directions of $\mathfrak{A}$ and r_k, is perpendicular to r_k, and is considered positive when its component along the direction of $\mathfrak{A}$ is positive.

The typical particle of the absorber has a charge e_k and mass m_k. It will experience in the electric field of the disturbance an acceleration, $\mathfrak{A}_k$, equal to (e_k/m_k) times expression (1). Its motion will generate a field which will be half-advanced and half-retarded. The advanced part of this field will exert on the source a force simultaneous with the original acceleration. The component of this reactive force along the direction of the acceleration will be

$$-e(e_k\mathfrak{A}_k/2r_kc^2)\sin(\mathfrak{A}, r_k)$$
$$=(\mathfrak{A}e^2/2c^4)(e_k^2/m_kr_k^2)\sin^2(\mathfrak{A}, r_k). \qquad (2)$$

From expression (2) for the reactive force due to one particle of the absorber, we can evaluate the total effect due to many particles, present to the number N per unit of volume. The number of particles in a spherical shell of thickness dr_k will be $4\pi Nr_k^2dr_k$. For the particles in this shell the average value of the geometrical factor $\sin^2(\mathfrak{A}, r_k)$ will be $(2/3)$. Consequently we obtain for total force of reaction the integral of the expression

$$(2\mathfrak{A}e^2/3c^3)(2\pi Ne_k^2/m_kc)dr_k. \qquad (3)$$

The force (3) gives an account of the phenomenon of radiative reaction which is not in accord with experience:

(1) The force acts on the source in phase with its acceleration; or in other words, it is proportional to the acceleration itself rather than to the time rate of change of acceleration.

(2) The reaction depends upon the nature of the absorbing particles.

(3) The force appears at first sight to grow without limit as the number of particles or the thickness of the absorber is indefinitely increased.

Nevertheless, proper addition of the effects due to all the particles of a complete absorber, with due allowance for their phase relations, does lead, as we shall see, to a reasonable expression for the reaction on the source.

There exists a phase lag between outgoing disturbance and returning reaction which we have not taken into account. The advanced force acting on the source due to the motion of a typical particle of the absorber is an elementary interaction between two charges, propagated with the speed of light in vacuum. On the other hand, the disturbance which travels outward from the source and determines the motion of the particle in question is made up not only of the proper field of the originally accelerated charge, but also of the secondary fields generated in the material of the absorber. The elementary interactions are of course propagated with the speed of light; but the combined disturbance travels, as is well known from the theory of the refractive index, at a different speed,

$$c/(\text{refractive index}) = c/n.$$

In order to speak of the change in velocity of the disturbance, or to treat the refractive index of the absorber in a well-defined way, it will be necessary to consider a single Fourier component of the acceleration. The connection between acceleration and reactive force being a linear one, it will be legitimate to decompose the acceleration into parts of this kind, and later to recompose the corresponding Fourier components of the radiative reaction. We shall, therefore, suppose for the moment that the primary acceleration varies with time according to the formula

$$\mathfrak{A}=\mathfrak{A}_0 \exp(-i\omega t), \tag{4}$$

where ω represents the circular frequency of the motion. A disturbance of this frequency will experience in a medium of low density a refractive index given by the familiar formula,

$$n=1-2\pi N e_k^2/m_k\omega^2. \tag{5}$$

Thus the radiative reaction which reaches the source from a depth r_k in the absorber will lag in phase behind the acceleration by the angle

$$\omega(r_k/c-nr_k/c)=(2\pi N e_k^2/m_k c\omega). \tag{6}$$

We apply this phase correction to the contribution (3) of absorber particles in the range r_k to r_k+dr_k, and sum over all depths in the medium to obtain the total reactive force,

$$(2e^2/3c^3)\mathfrak{A}\int_0^\infty (2\pi N e_k^2/m_k c)dr_k \times\exp(-ir_k 2\pi N e_k^2/m_k c\omega). \tag{7}$$

This integral will converge at the upper limit when we allow for the existence of a small but finite coefficient of absorption in the medium. Or in the language of physical optics, so familiar from the writings of R. W. Wood, we can say that we have to determine the combined effect of a number of wave zones, alternately in and out of phase with the acceleration. The resultant force is 90° out of phase with the acceleration and is equal in magnitude to the arithmetic sum of the contributions from depths up to a point where the phase lag is one radian:

$$\begin{aligned}(\text{total reaction}) &= (2e^2/3c^3)(-i\omega\mathfrak{A}) \\ &= (2e^2/3c^3)(d\mathfrak{A}/dt).\end{aligned} \tag{8}$$

This result, derived by considering only a single Fourier component of the acceleration, no longer contains explicit reference to the frequency of that component. Consequently expression (8) applies whatever is the dependence of acceleration upon time, so long as the velocities of all particles remain non-relativistic. In this respect we have a quite general derivation of the law of radiative reaction generally accepted as correct for a slowly moving particle subjected to an arbitrary acceleration.

We conclude that the force of radiative reaction arises, not from the direct action of a particle upon itself, but from the advanced action upon this charge caused by the future motion of the particles of the absorber.

RADIATIVE REACTION: DERIVATION II

In the above treatment we considered first the retarded electromagnetic disturbance traveling outward in the absorbing medium; second, the motion of the particles of the medium due to this disturbance; third, the advanced part of the elementary fields produced by these motions; fourth, the sum of these fields at the position of the source. The same chain of reasoning will allow us to sum the elementary advanced fields of the particles of the absorbing medium at points in the neighborhood of the source. We shall find that this field is just sufficient, when added to the half-advanced, half-retarded field of the source itself, to give the usual full strength purely retarded field which one is accustomed to attribute to a radiating source. Thus we shall justify the assumption made in the first derivation as to the strength of the outgoing disturbance. In order to make it clear that our reasoning is not circular, we shall represent the magnitude of the disturbance by a multiple, (?), of the usual full retarded field, and shall actually deduce the value unity for this at present undetermined factor.

We shall now evaluate the contribution of particles in the absorber to the electric field acting in the region roundabout the source. In order to simplify the geometrical considerations as much as possible, we shall visualize the source as located at the center of a spherical cavity of radius R in the medium. We shall take the distance, r, from the source to the point of evaluation of the field to be small in comparison with this radius. We shall however give up the

assumption that the particles of the absorber are necessarily free, or that they are far from one another. To make this generalization in our previous derivation, we shall express the acceleration of the typical particle of the absorber for a disturbance of circular frequency ω in the form

$$(\text{electric field of disturbance})\cdot(e_k/m_k)\cdot p(\omega). \quad (9)$$

Here $p(\omega)$ is in general a complex function of ω which approaches unity only in the case of weak binding or high frequencies. The factor $p(\omega)$, according to the theory of dispersion, determines the complex refractive index, $n-ik$, of the medium:

$$1-(n-ik)^2=(4\pi Ne_k^2/m_k\omega^2)p(\omega). \quad (10)$$

The advanced field produced by the absorber at the distance from the source will be given in amplitude and phase by the product of the following factors:

$\mathfrak{A}=\mathfrak{A}_0\exp(-i\omega t)$,
the acceleration of the source, here assumed for simplicity to be periodic, although this periodicity will drop out of the final result.

$-(e/r_kc^2)\sin(\mathfrak{A}, r_k)$,
the factor by which the acceleration must be multiplied to obtain the strength of the full retarded electric field in vacuum at a great distance, r_k, from an accelerated particle of charge e.

(?),
factor as yet undetermined, which allows for the possibility that the disturbance which is propagated outward from the particle, and which is in general due only partly to the source itself, may differ in strength from the usual full retarded field. For an isolated charge in otherwise charge-free space this factor is equal to $(\frac{1}{2})$. In the present case of a complete absorber we shall however later find for this factor the value unity. The product of the factors so far gives the strength of the electric field which would act on an isolated particle at the distance r_k.

$\exp(i\omega r_k/c)$,
the phase of the disturbance which would act on such an isolated particle.

$2(1+n-ik)^{-1}$,
factor by which the strength of the electric field of the disturbance inside the medium is reduced by reflection at the wall of the cavity, a factor taken over from electromagnetic theory.

$\exp(i\omega(n-ik-1)(r_k-R)/c)$,
factor allowing for the change in phase and amplitude of the disturbance produced by propagation to the depth (r_k-R) in the medium.

$(e_k/m_k)p(\omega)$
factor relating acceleration of absorber particle to electric field experienced by it.

$-(e_k/2r_kc^2)\sin(\mathfrak{A}, r_k)$,
factor to be multiplied by acceleration of absorber particle to give the magnitude of the component of the advanced electric field produced by the absorber in the neighborhood of, and parallel to the acceleration of, the source.

$\exp(-i\omega r_k/c)$,
factor allowing for the difference in phase between (a) the advanced field of the absorber as evaluated at the source itself and (b) the acceleration of the typical absorber particle.

$\exp(i\omega r\cos(r, r_k)/c)$,
correction to be applied to phase of absorber field at the source itself in order to evaluate this field at the distance, r, from the source. The product of the factors so far gives in magnitude and phase the advanced field at this point owing to a single particle of the absorber.

$Nr_k^2dr_kd\Omega$,
number of absorber particles in the element of solid angle $d\Omega$ and in the interval of distance dr_k.

We evaluate the product of the listed factors and sum over all particles of the absorber to evaluate the total advanced field of the absorber in the neighborhood of the source:

$$(?)(e/c^3)\mathfrak{A}_0\exp(-i\omega t)\int_\Omega \exp(i(\omega/c)r\cos(r, r_k))\times\sin^2(\mathfrak{A}, r_k)(d\Omega/4\pi)$$
$$\int_R^\infty (4\pi Ne_k^2/m_kc)p(\omega)(1+n-ik)^{-1}dr_k \times\exp(i\omega(n-ik-1)(r_k-R)/c). \quad (11)$$

The last integral is simplified by the relationship (10) between refractive index and physical properties of the medium to an expression

$$\int_R^\infty(\omega^2/c)(1-n+ik)dr_k \times\exp(i\omega(n-ik-1)(r_k-R)/c)=-i\omega, \quad (12)$$

completely independent of the properties of the absorber. Having thus summed over all particles lying in a given direction, we sum over different directions, using the relations

$$(d\Omega/4\pi)=(1/2)d\cos(r, r_k)(d\varphi/2\pi), \quad (13)$$

where φ is the dihedral angle between the $(r, \mathfrak{A})$ and (r, r_k) planes; and

$$\int\sin^2(\mathfrak{A}, r_k)(d\varphi/2\pi)$$
$$=(2/3)\int[1-P_2(\cos(\mathfrak{A}, r_k))]d\varphi/2\pi$$
$$=(2/3)[1-P_2(\cos(\mathfrak{A}, r))P_2(\cos(r_k, r))]. \quad (14)$$

Also we note the integrals

$$(1/2)\int_{-1}^{1}\exp(iu\cos\theta)d(\cos\theta)=F_0(u)$$
$$=u^{-1}\sin u\simeq\begin{cases}1 \text{ for small } u\\ (e^{iu}-e^{-iu})/2iu \text{ for all } u,\end{cases}\quad(15)$$

and

$$(1/2)\int_{-1}^{1}\exp(iu\cos\theta)P_2(\cos\theta)d(\cos\theta)=F_2(u)$$
$$=(u^{-1}-3u^{-3})\sin u+3u^{-2}\cos u$$
$$\simeq\begin{cases}-u^2/15 \text{ for small } u\\ (e^{iu}-e^{-iu})/2iu \text{ for large } u.\end{cases}\quad(16)$$

By use of these mathematical results, we find that the advanced field of the absorber has in the cavity an electric component parallel to the acceleration of the source which is given in magnitude and phase by the expression

$$(?)(2e/3c^3)(-i\omega\mathfrak{A}_0)\exp(-i\omega t)$$
$$[F_0(\omega r/c)-P_2(\cos(\mathfrak{A},r))F_2(\omega r/c)].\quad(17)$$

The radiation field so obtained reduces at the source itself to the form

$$(?)(2e/3c^3)(d\mathfrak{A}/dt),\quad(18)$$

and at distances a number of wave-lengths from the source goes over into the expression

$$-(?)(e\mathfrak{A}_0/2rc^2)\exp(i\omega r/c-i\omega t)$$
$$+(?)(e\mathfrak{A}_0/2rc^2)\exp(-i\omega r/c-i\omega t).\quad(19)$$

In words, formula (19) states that the advanced field of the absorber is equal in the neighborhood of the accelerated particle to the still undetermined factor (?), multiplied by the difference between half the retarded field (first term) and half the advanced field (second term) which one calculates for the source itself.

It is instructive to see how superposition of the advanced fields of a large number of particles can give the appearance of both retarded and advanced fields due to the source itself. The advanced field of a single charge of the absorber can be symbolized as a sphere which is converging towards the particle and which will collapse upon it at just the moment when it is disturbed by the source. But at the moment when the source particle itself was accelerated, the sphere in question had a substantial radius. One point on it touched, or nearly touched, the source. The shrinking sphere therefore appears to the source as a nearly plane wave which passes over it headed towards one of the particles of the absorber. When we consider the effect of all the absorbing charges, we have to visualize an array of approximately plane waves, all marching towards the source and passing over it in step. The resultant of these individual effects is a spherical wave, the envelope of the many nearly plane waves. The sphere converges, collapses on the source, and then pours out again as a divergent sphere. An observer in the neighborhood will gain the impression that this divergent wave originated from the source.

A test particle will be unable to make a separation between the two retarded fields, one properly owing to the source, the other really owing to the advanced field of the absorber. Thus we have for the disturbance diverging from the source the relation

$$\begin{pmatrix}\text{total disturbance}\\ \text{diverging from}\\ \text{source}\end{pmatrix}=\begin{pmatrix}\text{proper retarded}\\ \text{field of source}\\ \text{itself}\end{pmatrix}+\begin{pmatrix}\text{field apparently diverging from source,}\\ \text{actually composed of parts converging}\\ \text{on individual absorber particles}\end{pmatrix}.\quad(20)$$

We are now in a position to evaluate the undetermined factor, (?), in the expression we have used in the above analysis for the force acting on the typical particle of the absorber. We have only to express all three terms of Eq. (20) in units of the usual retarded solution of Maxwell's equations, a solution which asymptotically for large distances from the source gives for the electric field parallel to the acceleration the expression

$$-(e\mathfrak{A}/rc^2)\sin^2(\mathfrak{A},r).\quad(21)$$

To evaluate the third term in (20), we refer back to Eq. (19). Thus we find the algebraic equation

$$(?)=(1/2)+(?/2),\quad(22)$$

of which the solution is

$$(1)=(1/2)+(1/2).\quad(23)$$

From our derivation we find for the disturbance diverging from an accelerated charge the full retarded field required by experience.

Along the same lines we can find the strength

of the advanced field converging upon the source before the moment of acceleration:

$$\begin{pmatrix}\text{total disturbance}\\ \text{converging on}\\ \text{source}\end{pmatrix} = \begin{pmatrix}\text{proper advanced}\\ \text{field of source}\\ \text{itself}\end{pmatrix} + \begin{pmatrix}\text{field apparently convergent on source,}\\ \text{actually composed of parts convergent}\\ \text{on individual absorber particles}\end{pmatrix}. \tag{24}$$

At distances of several wave-lengths from the source, the two terms on the right possess simple mathematical expressions. Measured in terms of (21) as a unit of field strength, the right-hand side of (24) has the value $(1/2)-(1/2)=0$. We conclude that there is no net disturbance converging upon the source prior to the time of acceleration. The advanced field of the source is completely compensated by the advanced field of the absorber.

Our picture of the mechanism of radiation is seen to be self-consistent. Any particle on being accelerated generates a field which is half-advanced and half-retarded. From the source a disturbance travels outward into the surrounding absorbing medium and sets into motion all the constituent particles. They generate a field which is equal to half the retarded minus half the advanced field of the source. In this field we have the explanation of the radiation field assumed by Dirac. The radiation field combines with the field of the source itself to produce the usual retarded effects which we expect from observation, and such retarded effects only. The radiation field also acts on the source itself to produce the force of radiative reaction. What we have said of one particle holds for every particle in a completely absorbing medium. All advanced fields are concealed by interference. Their effects show up directly only in the force of radiative reaction. Otherwise we appear to have a system of particles acting on each other via purely retarded forces.

RADIATIVE REACTION: DERIVATION III

So far the source has been assumed to be at rest, or in slow motion, at the time of acceleration. The expression derived above for the force of radiative reaction is therefore limited in its applicability. To obtain the corresponding law of damping for a swift particle three possibilities suggest themselves, each calling for a mathematical technique quite different from that of the other two. The first and simplest procedure is to look at the particle from a frame of reference moving with nearly its own speed, apply in this frame the expression which we already have, and then transform back to the laboratory frame of reference. This application of the transformation of Lorentz is perfectly legitimate but not especially instructive.

A second method to calculate the force of reaction for a fast particle comes from Dirac. He makes the assumption that the damping arises from the action on the particle of a field equal to half the difference between the particle's own retarded and advanced fields, a conception which we have now interpreted in terms of the radiative reaction of the absorber. As each of the two fields is individually singular at the location of the charge, evaluation of the difference requires one to apply a limiting process which presents a certain mathematical difficulty, though in principle perfectly straightforward.

In connection with the limiting process of Dirac, it is interesting to refer back to the calculation of radiative reaction made by Lorentz on the model of an extended charge, every part of which exerted a retarded effect upon every other part. The elementary retarded field can be written in the form

$$\begin{pmatrix}\text{retarded}\\ \text{field}\end{pmatrix} = \left[\tfrac{1}{2}\begin{pmatrix}\text{retarded}\\ \text{field}\end{pmatrix} + \tfrac{1}{2}\begin{pmatrix}\text{advanced}\\ \text{field}\end{pmatrix}\right] + \left[\tfrac{1}{2}\begin{pmatrix}\text{retarded}\\ \text{field}\end{pmatrix} - \tfrac{1}{2}\begin{pmatrix}\text{advanced}\\ \text{field}\end{pmatrix}\right].$$

Here the first term is singular and is related to what Lorentz called the electromagnetic mass of the particle. The second part, on the other hand, is the only one asymmetrical in time and capable of contributing to the force of radiative reaction. In present terms, the procedure of Lorentz amounts to an ingenious means to determine the limiting value of Dirac's radiation field at the position of the source. Unfortunately the procedure is not convenient to apply to a rapidly moving extended charge because of the relativistic contraction of its spherical form.

The third procedure to evaluate the law of radiative damping for a swift particle is to calculate directly the reaction of the absorber on the source, along the lines of derivations I and II. This approach uses the expression for the field of a radiating charge, not at small distances, where it is singular, but at large distances, where it has a simple asymptotic form. We shall explore this type of derivation because of its direct relation to the absorber theory of radiation.

We idealize the absorber as before as a sphere of very large radius, r, centered on the point reached by the given particle at a chosen instant. At the surface of the absorber, those constituents of the field which drop away as $1/r^2$ will have become negligible in comparison with those which fall off as $1/r$. The typical particle on this surface experiences an electric field perpendicular to the direction of r. The magnitude of this field was represented in the case of a slowly moving source by an expression of the form, $-(e/rc^2)$(component of acceleration perpendicular to r), and is similarly representable in the present case by an expression of the type, $-(e/r)$(function of motion)$_\Omega$. Here the function in parenthesis, while more complicated than before, still depends only on the motion of the particle and the direction to the point of action. The influence of this disturbance causes the particles of the absorber to generate a field, the advanced part of which at the position of the source claims our attention. We consider the portion of this returned field arising from particles of the absorber which lie within an element of solid angle $d\Omega$. The electric component of this field is perpendicular to r and had for a slowly moving source the magnitude $(e/c^2)(-i\omega/c)$(component of acceleration perpendicular to r)$(d\Omega/4\pi)$ when the acceleration was a periodic function of time, and more generally was given by the derivative $(e/c^2)(d/cdt)$ $\times$(component of acceleration perpendicular to r) $\times(d\Omega/4\pi)$.

The relationship between returned field and original disturbance is a property only of the absorber and is independent of the state of motion of the source. Consequently, for the case of a particle moving at arbitrary velocity the returned electric field is perpendicular to r and equal in magnitude to

$$e(d/cdt)(\text{function of motion})_\Omega(d\Omega/4\pi).$$

What we have said of the electric field applies also to the magnetic field, because at great distances from an accelerated particle the two vectors have equal magnitudes and perpendicular directions. Thus we conclude that the reaction of the absorber on the source is described by a field, F_{mn}, which is directly related to the retarded field, R_{mn}, of the source at great distances, r, by the equation[12]

$$F_{mn} = -r\int(\partial R_{mn}/\partial x^4)(d\Omega/4\pi). \qquad (25)$$

The retarded field of the source particles, R_{mn}, in Eq. (25) is derived from the retarded potentials

$$A_m = 2e\int \dot{a}^m(\alpha)\delta(xa_\mu xa^\mu)d\alpha,$$

through the equation

$$R_{mn} = (\partial A_n/\partial x^m) - (\partial A_m/\partial x^n).$$

Here the integration over the proper cotime, α, goes only over that portion of the world line of particle a from which a retarded disturbance can reach the point of action, x^m. The significant value of α is connected with the coordinates x^m

[12] Here and below we use the following notation:

$x^1 = x_1$, $x^2 = x_2$, $x^3 = x_3$ — the three space coordinates of a typical point of evaluation of the field.

$x^4 = -x_4$, a quantity also having the dimensions of a length, and given by the product of the velocity of light and the time elapsed between a certain zero hour and the moment of observation.

a^m, similar space-time coordinates of a typical point on the world line of the ath particle.

Successive points along the world line are designated by the values of a parameter, α, the proper cotime, which has the dimensions of a length and is equal to the product of the velocity of light and the proper time. The difference, $d\alpha$, in proper cotime between two neighboring points has the same sign as the difference da^4, and is given in magnitude by the equation

$$(d\alpha)^2 = c^2(\text{time interval})^2 - (\text{space interval})^2 = -da_\mu da^\mu.$$

Derivatives with respect to α are denoted by dots. In comparing formulae given in this notation with those given elsewhere in the literature, it will be noted that some authors go from contravariant to covariant representation of a vector by reversing the sign of its space components and leaving its time component unaltered; also that dots are often used to indicate differentiation with respect to proper time, rather than proper cotime. In our notation the derivatives $\dot{a}^m$ are dimensionless quantities which satisfy the relation $\dot{a}_\mu\dot{a}^\mu = -1$. We use xa^m as an abbreviation for the vector, $x^m - a^m$. The usual scalar potential of the electromagnetic field is represented by the component A^4 of a four-vector, of which the other three parts, A^1, A^2, A^3, constitute the space components of the customary vector potential. The typical component of the field is given in the equation $F_{mn} = (\partial A_n/\partial x^m) - (\partial A_m/\partial x^n)$, where we have for the electric field $E_x = F_{14} = -F_{41}$, etc., and for the magnetic field $H_x = F_{23} = -F_{32}$, etc.

by the equation

$$xa_\mu xa^\mu = 0. \tag{26}$$

The integration yields

$$A_m = -e\dot{a}_m/(\dot{a}_\mu xa^\mu).$$

In differentiating this expression with respect to the coordinates of the point of observation, we have to allow for the associated change in the value of the proper cotime, given by the differential of (26),

$$(dx_\mu - \dot{a}_\mu d\alpha)(x^\mu - a^\mu) = 0,$$

or

$$d\alpha = (xa_\mu dx^\mu)/(\dot{a}_\mu xa^\mu). \tag{27}$$

Thus the retarded field of a is found to be given by the expression

$$R_{mn} = e(\dot{a}_\mu xa^\mu)^{-2}(\ddot{a}_m xa_n - \ddot{a}_n xa_m) + e(1+\ddot{a}_\mu xa^\mu)(\dot{a}_\mu xa^\mu)^{-3}(-\dot{a}_m xa_n + \dot{a}_n xa_m).$$

All terms in this expression fall off at large distances inversely as the first power of the separations xa, except for the terms arising from the unity in the factor $(1+\ddot{a}_\mu xa^\mu)$, which we may henceforth omit. For the same reason in differentiating the field with respect to x^4, we may treat all differences xa^m as constant. Thus we find in the limit of large distances

$$\begin{aligned}
&-r\int (\partial R_{mn}/\partial x^4)(d\Omega/4\pi) \\
&= r\int (-xa_4/\dot{a}_\mu xa^\mu)(d/d\alpha)R_{mn}(d\Omega/4\pi) \\
&= e\int \{3(\ddot{a}_\mu xa^\mu)(\dot{a}_\mu xa^\mu)^{-4}(xa_m\ddot{a}_n - xa_n\ddot{a}_m) \\
&\quad -(\dot{a}_\mu xa^\mu)^{-3}(xa_m\dddot{a}_n - xa_n\dddot{a}_m) \\
&\quad +(\dddot{a}_\mu xa^\mu)(\dot{a}_\mu xa^\mu)^{-4}(xa_m\dot{a}_n - xa_n\dot{a}_m) \\
&\quad -3(\ddot{a}_\mu xa^\mu)^2(\dot{a}_\mu xa^\mu)^{-5}(xa_m\dot{a}_n - xa_n\dot{a}_m)\} \\
&\qquad \times (r^2 d\Omega/4\pi).
\end{aligned} \tag{28}$$

As variables of integration it is convenient to use a colatitude θ and azimuthal angle φ, taking for polar axis the direction of the space component, $(\dot{a}^1, \dot{a}^2, \dot{a}^3)$, of the four-vector, $\dot{a}^m$. With this choice of variables the denominator of the typical term in the preceding expression is a power of the factor $(\dot{a}_\mu xa^\mu) = r(\dot{a}_4 + \dot{a}\cos\theta)$, where $\dot{a}_4{}^2 - \dot{a}^2 = 1$. The absence of the azimuthal angle from the denominator and the relatively simple form of the numerator makes it easy to carry out the integration over φ. The numerator of the typical term then reduces to a polynomial in $\cos\theta$. The integration over θ therefore leads only to algebraic functions of $\cos\theta$ to be evaluated at the two limits $\cos\theta = \pm 1$. The reduction of the resulting expressions to simple form requires rather long calculation. The final result for the field of radiative reaction at the location of the source is

$$\begin{aligned}
F_{mn} &= -r\int (\partial R_{mn}/\partial x^4)(d\Omega/4\pi) \\
&= (2e/3)(\dot{a}_m\ddot{a}_n - \ddot{a}_m\dot{a}_n).
\end{aligned} \tag{29}$$

This expression for the field of the absorbing particles agrees with that given by Dirac for half the difference of retarded and advanced fields due to the source itself, provided account is taken of the difference between the present notation and his.

If we define the force of radiative reaction through its contribution to the product of the mass of the particle by its acceleration, $mc^2\ddot{a}^m$, then we have for this force the expression

$$eF_{m\mu}\dot{a}^\mu = (2e^2/3)(\dot{a}_m\dddot{a}_\mu - \dddot{a}_m\dot{a}_\mu)\dot{a}^\mu. \tag{30}$$

In the case of a slowly moving particle the first space component of this force is readily evaluated by noting that (1) $\dot{a}_m$ is of the order of the ratio of the velocity of the corpuscle to the speed of light and is therefore negligible; (2) the quantity $-\dot{a}_\mu\dot{a}^\mu$ has the value unity; and (3) the derivative $\dddot{a}_m$ represents $(1/c^3)$ times the time rate of change of the given component of the acceleration, $\mathfrak{A}$. Consequently, the expression (30) reduces in the non-relativistic limit to the usual formula, $(2e^2/3c^3)(d\mathfrak{A}/dt)$, for the damping force.

From the properties of the retarded field at large distances from an accelerated particle in motion at an arbitrary velocity, we have obtained an expression for the force of radiative reaction previously derived by Dirac on the assumption that this force arises from half the difference of the advanced and retarded fields of the particle itself. It is, therefore, of interest to see that this equivalence can be demonstrated without going through the rather long calculations which are required on either method of derivation to obtain explicit expressions for the force of radiative reaction. To bring out the relationship between the two derivations, we go back to that expression for the retarded field of the source

which contains a delta function, and arrange the evaluation of Eq. (25) in such a way as always to keep a delta function in evidence. Thus we write the retarded field in the form

$$R_{mn}=2e\int[-\dot{a}_m(\partial/\partial x^n)+\dot{a}_n(\partial/\partial x^m)]\delta(xa_\mu xa^\mu)d\alpha.$$

In order to postpone the differentiation of the delta function, we adopt an expedient to transform the variable of differentiation. We consider in addition to the actual world line of the source, $a^m(\alpha)$, a displaced world line, a particle moving along which reaches at the proper cotime, α, the point $\bar{a}^m(\alpha)=a^m(\alpha)+D^m$, where the D^m are four numbers independent of α. We note that the derivative with respect to x^m of any function of the differences $x^k-\bar{a}^k$ is equal to the negative of the derivative of the same function with respect to D^m. Consequently we may write the expression for the retarded field in the form

$$R_{mn}=2e\int[\dot{a}_m(\partial/\partial D^n)-\dot{a}_n(\partial/\partial D^m)]\times\delta(x\bar{a}^\mu x\bar{a}_\mu)d\alpha, \quad (31)$$

where the result is to be evaluated in the limit when the displacements D^m go to zero.

We now insert expression (31) for the retarded field into the integral for the field returned by the absorber,

$$F_{mn}=r\int(\partial R_{mn}/\partial D^4)(d\Omega/4\pi),$$

and encounter the integral

$$F_{mn}=2er(\partial/\partial D^4)\int\int[\dot{a}_m(\partial/\partial D^n)-\dot{a}_n(\partial/\partial D^m)]\times\delta(x\bar{a}_\mu x\bar{a}^\mu)d\alpha(d\Omega/4\pi).$$

To bring out the meaning of this integral, we note that we want the radiative reaction on the source at a definite point, $a^m=a^m(\alpha^*)$, along its world line; that this point is at the center of a sphere of radius r; and that advanced disturbances from the particles on the inner surface of this sphere contribute to the force at this point only if they start at a cotime, x^4, equal to $r+a^4(\alpha^*)$. Consequently, x^4 has this fixed value as the integration over the surface of the sphere is carried out. Also during this integration we keep fixed the variable α and consequently hold constant $\bar{a}=\bar{a}(\alpha)$. Under these circumstances it is convenient to adopt for variable of integration the angle θ between the space directions $a\bar{a}$ and ax:

$$(\bar{a}x)^2=(a\bar{a})^2+(ax)^2-2(a\bar{a})(ax)\cos\theta.$$

Then we have

$$\int\delta(\bar{a}x_\mu\bar{a}x^\mu)(d\Omega/4\pi)=(1/2)\int_{-1}^{1}\delta(\bar{a}x_\mu\bar{a}x^\mu)d(\cos\theta)$$

$$=\int_{(\bar{a}x)=(ax)-(a\bar{a})}^{(\bar{a}x)=(ax)+(a\bar{a})}\delta(\bar{a}x_\mu\bar{a}x^\mu)\times d[(\bar{a}x)^2-(a\bar{a})^2-(ax)^2]/4(a\bar{a})(ax)$$

$$=\int\delta(\bar{a}x_\mu\bar{a}x^\mu)d(\bar{a}x_\mu\bar{a}x^\mu)/4(a\bar{a})r.$$

In this last expression the range of integration includes the point, $\bar{a}x_\mu\bar{a}x^\mu$, for which the delta function gives a contribution, only if there are some points on the surface of the sphere which can be reached simultaneously by two retarded waves which start out with $a^m(\alpha^*)$ and $\bar{a}^m(\alpha)$ as centers. This condition will be satisfied if and only if $\bar{a}(\alpha)$ lies between the forward and backward light cones drawn with $a(\alpha^*)$ as origin. Thus we have

$$F_{mn}=e(\partial/\partial D^4)\int d\alpha[\dot{a}_m(\partial/\partial D^n)-\dot{a}_n(\partial/\partial D^m)]\times\begin{cases}1/2(a\bar{a}) & \text{when } \bar{a}a_\mu\bar{a}a^\mu>0\\ 0 & \text{when } \bar{a}a_\mu\bar{a}a^\mu<0\end{cases}.$$

The differentiation with respect to D^4 of the discontinuous function in the last pair of brackets gives a function which has the character of a delta function except for a change in sign at one of the singularities. Specifically, writing

$$\delta(\bar{a}a_\mu\bar{a}a^\mu)=\delta_+-\delta_-,$$

where δ_+ is different from zero only when a retarded disturbance from $\bar{a}(\alpha)$ can reach the point $a(\alpha^*)$, and δ_- is different from zero only when an advanced disturbance from $\bar{a}(\alpha)$ can reach the point $a(\alpha^*)$, we have

$$(\partial/\partial D^4)\begin{cases}1/2(a\bar{a}) & \text{when } \bar{a}a_\mu\bar{a}a^\mu>0\\ 0 & \text{when } \bar{a}a_\mu\bar{a}a^\mu<0\end{cases}=\delta_+-\delta_-.$$

Then the field due to the absorber takes the form

$$F_{mn}=e\int[\dot{a}_m(\partial/\partial D^n)-\dot{a}_n(\partial/\partial D^m)](\delta_+-\delta_-)d\alpha$$

$$=e\int[\dot{a}_n(\partial/\partial a^m)-\dot{a}_m(\partial/\partial a^n)](\delta_+-\delta_-)d\alpha. \quad (32)$$

In other words, the reactive field at the point $a^m=a^m(\alpha^*)$ of the actual path is equal to half the retarded minus half the advanced field due to an equal charge moving on a world line of identical shape, all points of which are displaced by the amount D^m, this field evaluated in the limit $D^m\to 0$. This result establishes the connection between two different methods of evaluating the force of radiative reaction, one based on the properties of the retarded field of the source at great distances, the other containing half the difference of retarded and advanced fields at the location of the source itself.

THE RADIATIVE REACTION: DERIVATION IV

From the preceding applications of the absorber theory of radiation, it has become clear that such properties of the absorber as refractive index and density have no bearing on the magnitude of the force of radiative reaction. The only essential point is that the medium should be a complete absorber. We therefore expect that there should somehow be a means to take this point into account in a very general way.

In physical terms, complete absorption implies that a test charge placed anywhere outside the absorbing medium will experience no disturbance. In mathematical terms, using $F^{(k)}_{\rm ret}$ and $F^{(k)}_{\rm adv}$ to denote the retarded and advanced fields due to the kth particle, we have

$$\sum_k(\tfrac{1}{2}F^{(k)}_{\rm ret}+\tfrac{1}{2}F^{(k)}_{\rm adv})=0 \text{ (outside the absorber)}. \quad (33)$$

From the fact that this sum vanishes outside the absorber everywhere and at all times, it follows that each of the two sums also vanishes outside the absorber:

$$\sum_k F^{(k)}_{\rm ret}=0 \text{ (outside)} \quad (34)$$

and

$$\sum_k F^{(k)}_{\rm adv}=0 \text{ (outside)}. \quad (35)$$

Thus, the one sum, if it does not vanish, represents at large distances an outgoing wave, and the other represents a converging wave; but complete destructive interference between two such waves is impossible. Hence, if their sum vanishes, so does each field individually. From this conclusion it follows that the difference of the fields vanishes outside the absorber at all times:

$$\sum_k(\tfrac{1}{2}F^{(k)}_{\rm ret}-\tfrac{1}{2}F^{(k)}_{\rm adv})=0 \text{ (outside)}. \quad (36)$$

The field (36), in contrast to the fields (33)–(35), has no singularities within the absorber; it is a solution of Maxwell's equations for free space. Vanishing outside the absorber at all times, it must therefore forever be zero inside. The special property of a completely absorbing medium is expressed by the equation

$$\sum_k(F^{(k)}_{\rm ret}-F^{(k)}_{\rm adv})=0 \text{ (everywhere)}. \quad (37)$$

The consequences of Eq. (37) for the force on a typical particle are easily deduced. On the ath charge the entire field acting is given, according to the theory of action at a distance, by the sum

$$\sum_{k\neq a}(\tfrac{1}{2}F^{(k)}_{\rm ret}+\tfrac{1}{2}F^{(k)}_{\rm adv}). \quad (38)$$

This expression can be broken down into three parts:

$$\sum_{k\neq a}F^{(k)}_{\rm ret}+(\tfrac{1}{2}F^{(a)}_{\rm ret}-\tfrac{1}{2}F^{(a)}_{\rm adv})-\sum_{all\ k}(\tfrac{1}{2}F^{(k)}_{\rm ret}-\tfrac{1}{2}F^{(k)}_{\rm adv}). \quad (39)$$

Of these terms the third has just been shown to vanish for a complete absorber. The second gives rise to the phenomenon of radiative damping. In the case of non-relativistic velocities we have the result

$$e_a(\tfrac{1}{2}E^{(a)}_{\rm ret}-\tfrac{1}{2}E^{(a)}_{\rm adv})=(2e_a^2/3c^3)(d\mathfrak{V}_a/dt); \quad (40)$$

and in the case of swift particles we have for the force on the ath charge

$$e_a(\tfrac{1}{2}F^{(a)}_{n\alpha\ \rm ret}-\tfrac{1}{2}F^{(a)}_{n\alpha\ \rm adv})\dot{a}^\alpha.$$

This expression reduces, according to Dirac, to the form

$$(2e_a^2/3)(\dot{a}_n\ddot{a}_\alpha-\ddot{a}_n\dot{a}_\alpha)\dot{a}^\alpha, \quad (41)$$

in agreement with the reaction of the absorber as calculated in the preceding derivation. With this reactive term and the first term of (39), we arrive at the equation of motion of the typical particle in a completely absorbing medium

$$m_a\ddot{a}_n=e_a\sum_{k\neq a}F^{(k)}_{n\alpha\ \rm ret}\dot{a}^\alpha+(2e_a^2/3)(\dot{a}_n\ddot{a}_\alpha-\ddot{a}_n\dot{a}_\alpha)\dot{a}^\alpha. \quad (42)$$

In arriving at this equation we have shown that the half-advanced, half-retarded fields of the theory of action at a distance lead to a satisfactory account of the mechanism of radiative

reaction and to a description of the action of one particle on another in which no evidence of the advanced fields is apparent. We find in the case of an absorbing universe a complete equivalence between the theory of Schwarzschild and Fokker on the one hand and the usual formalism of electrodynamics on the other. This is what was to be proved.

THE IRREVERSIBILITY OF RADIATION

An oscillating charge surrounded by an absorbing medium loses energy. Why does radiation have this irreversible character even in a formulation of electrodynamics which is from the beginning symmetrical with respect to the interchange of past and future?

It might at first sight appear that the irreversibility is connected with the property of complete absorption. This is not the case. The expression (37) of the condition of absorption is perfectly symmetrical between advanced and retarded fields. We have only to reverse the roles of these two fields in the derivation following (37) in order to arrive at an equation of motion for the typical particle just as legitimate as (42), and in complete harmony with that equation:

$$m_a\ddot{a}_n = e_a \sum_{k\neq a} F^{(k)}_{n\alpha\ \mathrm{adv}}\dot{a}^\alpha - (2e_a^2/3)(\dot{a}_n\dddot{a}_\alpha - \ddot{a}_n\ddot{a}_\alpha)\dot{a}^\alpha. \quad (43)$$

In this equation, however, the force of radiative reaction appears with a sign just opposite to its usual one. Evidently the explanation of the one-sidedness of radiation is not purely a matter of electrodynamics.

We have to conclude with Einstein[11] that the irreversibility of the emission process is a phenomenon of statistical mechanics connected with the asymmetry of the initial conditions with respect to time. In our example the particles of the absorber were either at rest or in random motion before the time at which the impulse was given to the source. It follows that in the equation of motion (42) the sum, $\sum_{k\neq a} F^{(k)}_{n\alpha\ \mathrm{ret}}$, of the retarded fields of the adsorber particles had no particular effect on the acceleration of the source. Consequently the normal term of radiative damping dominates the picture. In the reverse formulation (43) of the equation of motion, the sum of the advanced fields of the absorber particles is not at all negligible, for they are put into motion by the source at just the right time to contribute to the sum $\sum_{k\neq a} F^{(k)}_{n\alpha\ \mathrm{adv}}$. This contribution, apart from the natural random effects of the changes of the absorber, has twice the magnitude of the usual damping term. The negative reactive force of (43) is therefore cancelled out, and a force of the expected sign and magnitude remains.

That it is solely the nature of the initial conditions which governs the direction of the radiation process can be seen by imagining a reversal of the direction of time in the preceding example. We have then a solution of the equations of motion just as consistent as the original solution. However, our interpretation of the solution is different. As the result of chaotic motion going on in the absorber, we see each one of the particles receiving at the proper moment just the right impulse to generate a disturbance which converges upon the source at the precise instant when it is accelerated. The source receives energy and the particles of the absorber are left with diminished velocity. No electrodynamic objection can be raised against this solution of the equations of motion. Small *a priori* probability of the given initial conditions provides our only basis on which to exclude such phenomena.

A comparison of radiation with heat conduction is illuminating. Both processes convert ordered into disordered motion although every elementary interaction involved is microscopically reversible.

Consider for the moment the question of the irreversibility of heat conduction, later to be put into relation with the problem of the one-sidedness of radiation. A portion of matter observed at the present moment to be warmer than its surroundings will cool off in the future with a probability overwhelmingly greater than the chance for it to grow hotter. About the past of the same portion of matter Boltzmann's *H*-theorem however also predicts an enormously greater likelihood that the body warmed up to its present state rather than cooled down to it. In other words, we are asked to understand the present temperature of the body as the result of a simple statistical fluctuation in the distribution of energy through the entire system. This de-

TABLE I. Decomposition of the symmetric fields of the theory of action at a distance into the fields of the retarded field theory.

Total field acting on ath particle in theory of action at a distance; here decomposed into:	$\sum_{k\neq a}(\frac{1}{2}F^{(k)}_{\text{ret}}+\frac{1}{2}F^{(k)}_{\text{adv}})$
(1) Retarded fields of usual formulation of electrodynamics.	$\sum_{k\neq a} F^{(k)}_{\text{ret}}$
(2) A field completely determined by the motion of the particle itself; denoted as the "radiation field" by Dirac; accounts for the normal force of radiative reaction.	$[\frac{1}{2}F^{(a)}_{\text{ret}}-\frac{1}{2}F^{(a)}_{\text{adv}}]_{mn}$ $=(2e_a/3)(\dot{a}_m\ddot{a}_n-\ddot{a}_m\dot{a}_n)$
(3) A residual field, with singularities at none of the particles, but completely determined by the motion of the particles; identified by us with Dirac's "incident field."	$\sum_{\text{all particles}}(\frac{1}{2}F^{(k)}_{\text{adv}}-\frac{1}{2}F^{(k)}_{\text{ret}})$ $\equiv F_{\text{inc}}$

duction is based on the premise that the system was isolated before observation. However, common experience tells us that the given portion of matter probably acquired its abnormal temperature, not via an internal statistical fluctuation, but because it had earlier not been isolated from the outside.

For the radiative analogy of this example of heat conduction, conceive a charged particle bound to a position of equilibrium by a quasi-elastic force. Furthermore suppose its energy at the moment of observation is large in comparison with the agitation of the surrounding absorber particles. There is then an overwhelming probability that the oscillator will lose energy to the absorber at a rate in close accord with the law of radiative damping. What can be said of the particle prior to the moment of acceleration? In an ideal absorbing system completely free of special disturbances, there is an equally overwhelming chance that the energy of the charge was then increasing at a rate given approximately by the inverse of the law of radiative damping. In this case as in heat conduction the abnormally high energy of the object is to be interpreted as the result of a statistical fluctuation. However, that the sun at some past age acquired its energy by such a fluctuation no one now would seriously propose. Obviously the universe is a special system with respect to the origin of which probability considerations cannot freely be applied.

We conclude that radiation and radiative damping come under the head, not of pure electrodynamics, but of statistical mechanics. The conventional expression for the force of radiative reaction, like those for frictional resistance and viscous drag, represents a statistical average only. Application of this concept is not required in such an instance as the case of complete thermodynamical equilibrium, where the relative fluctuations of the actual forces about the conventional values are substantial. The concept of radiative damping is of real value only when we deal with the conversion of organized into disorganized energy, as in wireless transmission or light production.

COMPLETE AND INCOMPLETE ABSORPTION

In the picture of radiation which we have built on the foundation of Tetrode's suggestion, the absorber plays a role of hitherto unsuspected importance. On this account we should investigate not only how much the mechanism depends upon the completeness of the interception, but also the question what should be said of the absorption in the case of the actual universe.

In discussing the case of incomplete interception, we require a convenient means to take into account the initial conditions which so clearly control the irreversibility of the force of radiative reaction. For this reason we shall break down the half-retarded, half-advanced fields of the theory of action at a distance into three parts as shown in Table I. With this decomposition of the field, we arrive at a description of the behavior of a system of particles which is entirely equivalent to the theory of action at a distance but which in the equation of motion,

$$m\ddot{a}_m=e_a\sum_{k\neq a}F^{(k)}_{m\alpha\ \text{ret}}\dot{a}^\alpha+(2e_a^2/3)\times(\dot{a}_m\ddot{a}_\alpha-\ddot{a}_m\dot{a}_\alpha)\dot{a}^\alpha+e_aF_{m\alpha\ \text{inc}}\dot{a}^\alpha, \quad (44)$$

conceals from view the existence of the advanced part of the fields of Schwarzschild and Fokker. We shall find it convenient to use for the field decomposition of Table I and the dynamical Eq. (44) the term "retarded field formulation of electrodynamics."

The field which enters the third term in the equation of motion (44) vanished in the case of a completely absorbing system. Its appearance

in the present case has led us to give it the name of "incident field," which Dirac applied to a quantity having an identical role in the equation of motion. However, on the origin of this field we go beyond Dirac's treatment in giving a prescription for its unique determination in terms of the movements of all the particles of the system. This prescription reveals that the field in question contains the advanced effects of the theory of action at a distance.

Some properties of the incident field may be noted before use is made of this concept in the analysis of special problems. The quantity F_{inc} has a singularity at the site of none of the charged particles. Consequently it satisfies Maxwell's equations for free space. Although completely determined by the motion of the charges, it thus has the character of a disturbance produced by sources at infinity. Now we already have in the retarded field, $\sum_{(\text{all particles})} F^{(k)}_{\text{ret}}$, a quantity whose behavior at all distances is likewise uniquely fixed by the motions of all the charges. Consequently we can expect to be able to deduce the incident field everywhere from a knowledge of the retarded field at large distances from the system of particles. Thus, in the determination of the incident fields we can, if we wish, avoid explicit reference to the movements of the charges, and base our considerations on the asymptotic behavior of their retarded fields alone. This point will be clearer after a consideration of a few examples, and can then be formulated in a general mathematical form as a by-product of an investigation primarily aimed at examining the problem of complete and incomplete absorption.

The simplest example will be the idealized case of a single-charged particle, alone in otherwise charge-free space, which is accelerated either by the gravitational attraction of a passing mass or by some other non-electromagnetic force. For the three electromagnetic forces of the equation of motion (44) we then have the following accounting: (1) There are no other particles, so the retarded field of the first term vanishes. (2) The second term is different from zero and represents the conventional force of radiative reaction. (3) The incident field of the third term is in the present case equal to half the advanced field minus half the advanced field owing to the particle itself. If we imagine the acceleration of the charge to be limited to a short stretch of time, then the incident field represents a disturbance which, long before the moment in question, was converging upon the particle from great distances. It focuses upon the particle at the period of acceleration and subsequently appears as a wave diverging from the charge. This disturbance, apparently produced by sources at infinity, exerts on the particle a force which is just sufficient in magnitude and in sign to cancel the normal force of radiative reaction. The description just given is the rather involved translation into the language of the retarded field theory of the conclusion immediately apparent from the theory of action at a distance with its half-advanced, half-retarded fields; an isolated charge neither experiences a force of radiative reaction nor radiates away electromagnetic energy.

The incident field of the preceding problem could have been determined equally well without knowledge of the motion of the particle itself, by reference to the retarded field, F_{ret}, of the charge at large distances. The latter quantity represents an electromagnetic disturbance which was negligible before the moment of acceleration, and which considerably later than that instant had the character of a diverging spherical wave. We can find a solution, S, of Maxwell's equations for free space, the diverging wave in the asymptotic expansion for which has exactly the same behavior as the field $-\frac{1}{2}F_{\text{ret}}$. By this condition the solution in question is furthermore uniquely determined. On this account it must be identical with another field which also satisfies Maxwell's equations for free space and has for its diverging wave at large distances the same form as $-\frac{1}{2}F_{\text{ret}}$; namely, the incident field, $F_{\text{inc}} \equiv \frac{1}{2}F_{\text{adv}} - \frac{1}{2}F_{\text{ret}}$. Consequently we may write $F_{\text{inc}} = S$, where S is the solution of Maxwell's equations defined as above. From this means of arriving at the value of the incident field we conclude that the incident field is that solution of the wave equation for free space which, when added to the known retarded field, F_{ret}, will reduce by one-half the strength of the diverging wave in the asymptotic representation of F_{ret}.

As next idealized example of incomplete absorption we consider a source at the center of a blackbody with two opposed openings out into

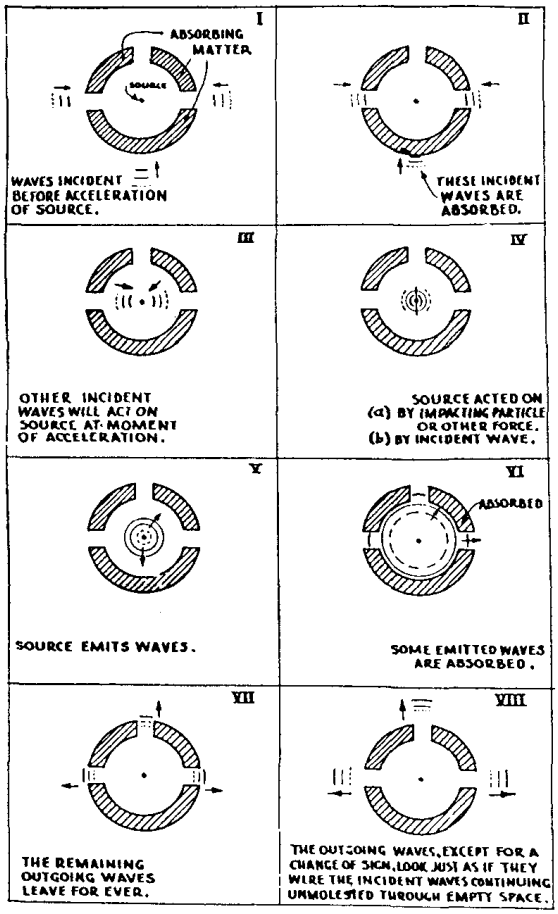


FIG. 1. Advanced effects in two examples of an incompletely absorbing system.

charge-free space (Fig. 1). In those directions not shielded by the absorber an incident field—to use the language of the retarded field theory—will enter, converge on the radiating particle, diverge and go out to infinity, as in the preceding example. This time, however, the wave incident from one side covers but the amount Ω of the whole solid angle of 4π. Consequently at the instant when it is focused upon the source, it reduces the force of radiative reaction only fractionally below the conventional value of this force. The fraction in question is equal in the case of a pair of small opposed openings and a slowly moving source to the product of the following factors:

$\Omega/4\pi$,
the fraction of the whole solid angle spanned by one hole.

2,
factor allowing for the existence of the two openings.

$\frac{3}{2}\sin^2(\mathfrak{A}, r)$,
factor allowing for the orientation of the holes relative to the direction of acceleration. Here $(\mathfrak{A}, r)$ is the angle between the vectorial acceleration of the source and the vector from this particle to one opening. The given expression for the polarization factor assumes that the quantities $d\mathfrak{A}/dt$ and $\mathfrak{A}$ are parallel, and has to be replaced by a more complicated term when this parallelism does not exist.

When the source is moving at the time of acceleration with a speed comparable to that of light, then the angular distribution of the radiation is represented by an expression more complicated than $\sin^2(\mathfrak{A}, r)$ but the general principle is the same.

In the present case of an absorber complete except for two inversion-symmetric openings to charge-free space we conclude that there is a continuous transition, as the size of the apertures is increased, from the full conventional force of radiative reaction on a central source, to the case of no radiative reaction at all. Furthermore, we note that a test charge placed in one of the two openings will receive a disturbance some time before as well as some time after the moment when the source itself is given its acceleration. Generally we may say that the explicit appearance of advanced effects is unavoidable in the case of a system which is an incomplete absorber. However, in neither of the examples so far examined do advanced fields of the source produce explicit advanced effects on any other particle than a test charge: in the first example, because there is no other charged particle; and in the second case, because the incident field is restricted to a region of space where there are no particles to be disturbed, except a possible test particle.

ADVANCED EFFECTS ASSOCIATED WITH INCOMPLETE ABSORPTION

Recognizable advanced effects appear for the first time in a third example, a source at the center of a cavity completely absorbing except for a single passage to charge-free space. (See Fig. 1.) For simplicity we consider the source to be a slowly moving particle. Also we shall denote as "antipassage" that portion of the absorber which is marked out by the inversion, with respect to the source, of the passage itself. Apart from the fields which come from and go to the passage and antipassage, we have the usual solution of the problem of a completely absorbing system. In the language of the retarded formu-

lation of field theory, we can say that there is in all other directions no incident field, and that in those directions the absorber experiences only a normal full strength retarded field as a result of the acceleration of the source. In the language of the theory of action at a distance the result is the same. The source gives out a half-advanced, half-retarded field, the advanced part of which in the direction of the normal portions of the absorber is cancelled by a portion of the advanced fields generated in the absorber itself. The remaining portion of this advanced field combines with the half-strength retarded field of the source to give the full retarded disturbance demanded by experience. The advanced field of the absorber at the location of the source itself produces a force of radiative reaction which is below the conventional amount in proportion only to the small solid angle which we have so far left out of consideration. So far the results are quite as expected.

We have now to consider the effect on the antipassage of what, in the theory of action at a distance, is the half-advanced field of the source. If the passage itself were filled with absorbent, this material would have generated a field, the advanced part of which would have compensated the effect we are now considering. As it is, we visualize two possible solutions of the problem of motion. In the first, the uncompensated half-advanced field of the source sets into motion ahead of time the particles on the inner face of the antipassage. In the second solution, this advanced field is compensated by a mechanism yet to be explained, and the particles in question are not disturbed before the moment of acceleration of the source.

In the first solution, not illustrated in Fig. 1, the particles on the inner face of the antipassage spontaneously accelerate at a moment sufficiently early so that their retarded fields reach the source at just the moment when it is radiating. The retarded field of the charges of the antipassage, evaluated in the cavity, has in the present solution the following properties: (1) It vanishes except in the directions of the passage and antipassage. (2) In those directions it has a value completely determined by the movement of the source. (3) It combines with the Schwarzschild-Fokker field of the source to cancel its outgoing component travelling in the direction of the passage and to build up its advanced component on the side of the antipassage to a full strength advanced wave. The combined advanced wave has a magnitude exactly sufficient to account for the disturbance of particles of the antipassage ahead of time. (4) The field of these particles at the location of the source acts in the opposite sense from the conventional force of radiative reaction. The magnitude of that force is reduced by a fraction which, apart from a polarization factor, is equal to twice the solid angle subtended by the passage, divided by 4π.

The particles of the antipassage, in addition to the anticipatory movements already discussed, undergo, after the moment of acceleration of the source, a disturbance similar to that experienced by the charges which neighbor them on the inner face of the cavity. In this way they are caused to generate fields, the advanced part of which (1) combines with the half-retarded field of the source in the given direction to produce a full retarded disturbance that accounts for the motions in question and (2) cancels the advanced field of the source in the direction of the passage. Thus neither advanced nor retarded disturbance emerges from this passage to be detected by an external test charge.

The self-consistent solution which we have just described in terms of the symmetrical theory of action at a distance is easily summarized in the language of the retarded field theory. The sources of radiation are the central particle and the charges on the inner face of the antipassage. Each is considered to experience the conventional force of radiative reaction and to produce only retarded fields. The fields from the disturbed charges on the inner face of the cavity focus on the central charge at the moment of its acceleration, thus (1) partially compensating the conventional force of radiative reaction and (2) cancelling that part of its retarded field which is travelling in the direction of the passage. No retarded field gets outside the system. The incident field, determinable as we have seen before from the asymptotic behavior of the retarded field, consequently vanishes. Thus in the given illustration the equation of motion of each particle contains only the retarded fields of all the other particles, plus the full conventional force of radiative damping, a conclusion consistent with the solution which we have just

given. To complete the picture, we have to express in terms of these equations of motion, the explanation of the early motion of the particles on the inner face of the antipassage. This movement we attribute to the influence of the retarded fields coming from other portions of the wall of the cavity, and reinforcing at just this particular region of the surface. The same type of reasoning can be followed back step by step in the past, along a course which is very like the reversal in time of the mechanism by which a burst of radiative energy dissipates itself. Granted that the existence of incomplete absorption requires, in our example, the explicit occurrence of advanced effects, we have in the given solution one reasonable picture how these advanced effects may build up until the time of disturbance of the source and may then be followed by a succession of retarded effects of magnitudes diminishing as absorption and reflection at the inner walls of the cavity have their effects.

In this third example an equally consistent solution of the problem of advanced effects may be briefly outlined (Fig. 1). The fields are in this case such that a test charge placed within the cavity experiences only the full retarded field of experience. The particles on the wall of the cavity are set into motion only after the time when the source was struck and caused to radiate. These particles, by the now familiar mechanism of absorber response, generate fields, the advanced parts of which in the cavity cancel the advanced field of the source and bring its retarded field up to full strength. In particular, the advanced field of the source in the direction of the passage is compensated, so that an external test charge on that side of the absorber will experience no advanced disturbance. It will, however, on our picture undergo a full strength retarded disturbance. How the half-retarded disturbance of the source in this direction is built up to full strength, and how the half-advanced field in the direction of the antipassage is cancelled, is a question still to be cleared up. For explanation we cannot (1) call upon the advanced fields of the absorber in the direction of the passage, for there is no matter in this direction. Nor can we (2) call upon retarded fields of particles on the inner face of the antipassage for our purpose—although they lie in just the right direction thus to produce the field in question—because they have been assumed not to have been set in motion prior to the time of acceleration of the source. To account for the directional field of so far unknown origin in the cavity, we are by (1) restricted to an explanation in terms of a retarded wave from some source yet to be found, lying in the direction of the antipassage and by (2) this source cannot consist of the particles on the inner face of that portion of the absorber. Consequently we must interpret the field in question as owing to particles set in motion ahead of time on the outer face of the antipassage. This conclusion, like many of the considerations to which we are led in the study of incompletely absorbing systems, appears paradoxical. Nevertheless it leads to a self-consistent solution of our problem. In the first place, the half-retarded field of the surface particles compensates within the antipassage as well as within the cavity the half-advanced field of the source. There is, therefore, no question of the propagation of any disturbance through the thickness of the absorber. Secondly, the half-advanced field of the surface particles appears from the point of view of an external test particle to be a wave front of limited cross section which comes from outer space and which would converge upon the source if the antipassage did not block its path. This half-advanced field in the region exterior to the absorber adds to the half-advanced field from the central source to give a full strength disturbance convergent upon the surface particles in question. Thus an account is given for the force and for the acceleration which they experience ahead of the time of acceleration of the source within the cavity. The energy which is absorbed at the outer surface is later paid out by the source. That accelerated particle experiences the full conventional force of radiative reaction.

The solution just given is readily translated into the language of the retarded formulation of field theory, where the force on each particle is attributed to three sources: the conventional force of damping, the retarded fields of all other particles and an incident field. The incident field has the appearance of a disturbance convergent from outer space upon the external face of the antipassage. By it the particles are accelerated and caused to generate a full strength retarded field which at greater depths within the medium

cancels the incident field, a phenomenon which is the normal mechanism of absorption. Thus there is within the cavity no disturbance convergent upon the source, and consequently only the usual retarded effects are observed, part of which also is propagated out through the passage into charge-free space.

In the retarded formulation of field theory, there is no apparent reason for a correlation between the surface absorption and the radiation process within the cavity. The requirement for a connection between the two comes into evidence only in the condition which must be satisfied by the incident field, and which has been discussed above. That there must be such a field follows from the existence of a full strength retarded field, R, diverging outward from the source through the passage in the absorber. Therefore, denote the strength of the incident field in this region of space and at this instant of time by nR, where n is a factor now to be found. Being also divergent from the source, but free of singularities there, the incident field must in this neighborhood and in the given cone of directions be a multiple of the radiation field of the source. On being followed backwards in time to moments previous to the acceleration of the particle, it must, therefore, have in the direction of the antipassage the magnitude $-nA$, where A denotes the advanced solution of Maxwell's equation for the accelerated source. Thus the field incident from great distances upon the particles on the outer surface of the antipassage must have the magnitude $-nA$. These particles generate a retarded field which within the absorber compensates the incident field and therefore has the magnitude $+nA$. This field, followed onward in the direction of the original source, where it naturally has no singularity, at first converges and then diverges to give the appearance of a retarded wave from the source itself. In this neighborhood the retarded field of the surface particles behaves much as does the radiation field of the source. Consequently the strength of the field in question, evaluated in the direction of the passage, is $-nR$. Thus the sum of the retarded fields of all the particles of the system, evaluated outside the passage way, is R (from the source) $-nR$ (from the surface particles). To determine the strength of the incident field, we now apply the condition that the divergent term, nR, in its asymptotic representation must have a strength equal to $-\frac{1}{2}$ times that of the divergent wave owing to the retarded fields of all the particles of the medium:

$$nR=-\tfrac{1}{2}(R-nR). \tag{45}$$

The solution of this implicit equation gives for the magnitude of the incident wave in the direction of the passage $nR=-R$, and consequently for the strength of the incident wave converging upon the other side of the system $+A$. Thus we check the properties of this second solution of our problem as obtained previously by using the language of half-advanced, half-retarded fields, with no reference to the concept of incident field.

Between the two self-consistent solutions of this third example of an incomplete absorber we make no attempt to choose. We have to accept the fact that the dynamical system in question possesses a number of degrees of freedom which is in direct proportion to the number of particles present. Once it is granted that advanced effects of some kind must be connected with the acceleration of the source, it does not follow uniquely upon which particles these advanced effects must act. The selection is a matter of initial conditions, not of equations of motion. The two solutions so far described are only two relatively simple samples from an infinite number of possible solutions, distinguished from one another by the requirements put upon the initial state of the particles of the absorber. It is only in the case of a completely absorbing system that there is the possibility to find a set of initial conditions which is relatively well determined by statistical considerations.

Self-consistency being the only requirement which has to be met by a solution of the problem of an incomplete absorber, and this requirement in the retarded field formulation being largely contained in a condition to be satisfied by the incident field, it may be of interest to have our so far informal statement of this relation put into mathematical terms. The condition in question furnishes a connection between the incident field, here abbreviated as I, and the sum, R, of the retarded fields of all the particles.

To derive the desired relation, we note that the present formulation of electromagnetic theory expresses the incident field as half the difference between the advanced and retarded fields owing to all the particles. Thus the advanced field of the system is given by the expression $R+2I$. From (1) a knowledge of this advanced field everywhere in space and (2) a knowledge of the values taken on the points, ξ, of a surface surrounding the system by an arbitrary solution, S, of Maxwell's equations for the same charge distribution, we can derive the values of this arbitrary solution at all other points in space, x, from the relation

$$S(x)=(R(x)+2I(x))+(2\pi)^{-1}\Big\{\int\int\int d\xi^2 d\xi^3 d\xi^1(\delta\partial S/\partial\xi^1-S\partial\delta/\partial\xi^1)\Big|_{\xi_{-1}}^{\xi_{+1}}+\cdots+\cdots$$
$$-\int\int\int d\xi^1 d\xi^2 d\xi^3(\delta\partial S/\partial\xi^4-S\partial\delta/\partial\xi^4)\Big|_{\xi_{-4}}^{\xi_{+4}}\Big\}.^{13} \quad (46)$$

In Eq. (46) the symbol δ stands for the delta function, $\delta(x^\mu-\xi^\mu\cdot x_\mu-\xi_\mu)$. The integral is to be taken only over the immediate neighborhood of those points on the surface from which an advanced wave can reach the point x^1, x^2, x^3 at the time x^4. In the first integrand we have the difference between the values of a certain quantity calculated for the largest and smallest values of ξ^1 consistent with given values of ξ^2, ξ^3, ξ^4; and similarly for the other three integrals. We shall write Eq. (46) symbolically in the form

$$S=(R+2I)+\text{Adv.}\,[S]. \quad (47)$$

We now apply this general relation to the special half-advanced, half-retarded solution of Maxwell's equations for the system of charges, $S=R+I$. In this way we arrive at an implicit equation by means of which to derive the incident field from the retarded field:

$$0=I+\text{Adv.}\,[R+I]. \quad (48)$$

This relation is the generalization of Eq. (45) from which we determined the strength of the incident field in the third example above. With this generalization we end our study of the behavior of idealized systems with incomplete absorption and come to the wider question what we should say about the absorbing properties of the system with which we have to deal in nature.

There would be no problem in interpreting the universe as a completely absorbing system if it were an indefinitely extended Euclidean space. The existence of the electron-positron field gives an mechanism by which, even in a vacuum, radiation of some frequencies can undergo absorption processes, and light of all wave-lengths can be scattered. These processes are sufficient ultimately to degrade all the radiation given out by an accelerated charge.

The universe is however now generally regarded as a closed space, in harmony with the illuminating theory put forward by Einstein. In this space present observations suggest that the absorption of radiation is far from complete even at the greatest depths so far plumbed, of the order of one-tenth the calculated radius of the universe. If this conclusion is correct, then a complete electrodynamic description of the mechanism of radiation would require us to take into account not only the curvature of space but also the phenomena summarized under the term "expanding universe." At the present time we know too little about these matters to carry out such a complete description. Moreover, there is yet no compelling reason to attempt this description. We know of course that electrodynamics remains, in other respects as well, to be tied to gravitational phenomena. But we recognize that in this sense our present theory of electrodynamics, like the theories in all other parts of science, is an idealization.

So long as we limit ourselves to the idealization based on the concept of a Euclidean space, we have to consider the question of complete and incomplete absorption on a purely empirical basis. In this connection we will obtain a satis-

[13] The analogue of Eq. (46), for determination of the arbitrary solution from a knowledge of the retarded solution, has been given in a rather different form by W. R. Morgans, Phil. Mag. 9, 148 (1930). The present form is most easily derived by use of the relation
$(\partial/\partial\xi^\mu)(\partial/\partial\xi_\mu)\delta(x^\nu-\xi^\nu\cdot x_\nu-\xi_\nu)$
$=-4\pi\delta(x^1-\xi^1)\delta(x^2-\xi^2)\delta(x^3-\xi^3)\delta(x^4-\xi^4)$,
and application of Green's theorem in four dimensions.

factory account of experience, as we have seen, on the assumption that the universe behaves as a completely absorbing system.

PRE-ACCELERATION

Is there in the case of a completely absorbing universe any consequence of the act of radiation which is so apparently paradoxical as the obviously advanced effects encountered in the instance of an incompletely absorbing system? If so, what words can we reasonably use to assimilate such a phenomenon into our experience?

Any advanced effects to appear in the case of a completely absorbing system must be deducible from the conventional force of radiative reaction, for the only other electrodynamical effects appearing in this case in the equation of motion (42) of the typical particle are the retarded fields of all other particles. That the damping term does lead to an advanced effect follows from an interesting example already considered by Dirac.[9] A source sends a sharp pulse of radiation towards a particle of charge e and mass m. At the instant of arrival the speed of the particle would be expected abruptly to increase if the force of damping were proportional to the first derivative of the displacement. Actually the radiative resistance is proportional to the third derivative of the displacement, and the nature of the solution of the equation of motion is changed. The particle commences to move before the time of arrival of the pulse; and e^2/mc^3 seconds ahead of time it attains a velocity comparable with its final speed.

As a suitable way to speak of this most interesting phenomenon of pre-acceleration brought to light by him, Dirac suggests saying that "it is possible for a signal to be transmitted faster than light through the interior of an electron. The finite size of the electron now reappears in a new sense, the interior of the electron being a region of failure, not of the field equations of electromagnetic theory, but of some of the elementary properties of space-time." This choice of language is perhaps suitable in certain respects to describe the pre-acceleration of the single charge in the example considered by Dirac. It may also be of value in other special instances. However, the given mode of speaking suggests in the case of a medium of closely packed charges the possibility of transmission of signals with a speed greater than that of light over microscopic distances, a conclusion which appears to be denied by a direct investigation of the point. Also the idea that the properties of space time fail in a region of the order of e^2/mc^2 around a charge appear to have possibilities of suggesting misleading conclusions sufficiently great to call for a later search for a more suitable means of expression.

We shall now attempt to test the idea suggested by the term "speed greater than that of light" that the phenomenon of pre-acceleration might be cumulative when charges are spaced at a distance from one another comparable to the quantity e^2/mc^2. The method of analysis will be very nearly that followed by Sommerfeld[14] and Brillouin[15] in their classic resolution of the question how it can be that the speed of propagation of a disturbance in a dispersive medium never exceeds the velocity of light even when the phase velocity for certain frequencies is far above this upper limit. The only significant mathematical difference between the two cases is the change of the damping force from proportionality to the first power of the frequency to proportionality to the third power.

The first step in the procedure of Sommerfeld and Brillouin is to determine the refractive index of the medium, n, as a function of frequency. The charges of the material are assumed normally to be at rest. Consequently the magnetic permeability is unity. According to the standard result of electromagnetic theory the square of the refractive index is in this case equal to the dielectric constant:

$$n^2 = \frac{\text{(electric field in a thin slot cut normal to the field)}}{\text{(electric field in a thin cavity cut parallel to the field)}} = (E+4\pi P)/E = 1+(4\pi P/E). \quad (49)$$

Here P represents the polarization of the medium:

$$P = \text{(number of charges per unit volume)} \cdot \text{(charge of each)} \cdot \text{(displacement from equilibrium)}. \quad (50)$$

[14] A. Sommerfeld, Ann. d. Physik **44**, 177 (1914).
[15] L. Brillouin, Ann. d. Physik **44**, 203 (1914).

For the force which determines the displacement of the charges in a homogeneous isotropic medium it is reasonable, according to Lorentz and Lorenz, to take the result valid for a cavity of spherical form,

$$(\text{force}) = (\text{charge per particle})(E+(4\pi P/3)). \quad (51)$$

The displacement itself is related to the force by the equation of motion,

$$\begin{aligned}(\text{force}) = m(&\text{displacement}) \\ &+(\text{a constant})\cdot(\text{displacement}) \\ &-(2e^2/3c^3)(\text{displacement}). \end{aligned} \quad (52)$$

Here we have visualized in the second term of the right side the possibility that the particles are bound to equilibrium positions by elastic forces. Without such forces we should be led by Earnshaw's theorem to expect that the medium would form a dynamically unstable system. We now follow Lorentz and Lorenz in selecting a function of refractive index which is easy to evaluate as a function of frequency:

$$\begin{aligned}3(n^2-1)/(n^2+2) &= 4\pi P/[E+(4\pi P/3)] \\ &= \frac{4\pi(\text{particle density})(\text{charge})^2(\text{displacement})}{m(\text{disp})+(\text{const})(\text{disp})-(2e^2/3c^3)(\dddot{\text{disp}})}. \end{aligned} \quad (53)$$

We consider a monochromatic disturbance of circular frequency $(3mc^3/2e^2)\omega$, and express the elapsed time as $(2e^2/3mc^3)\tau$, where both τ and ω are dimensionless quantities. We assume that the displacement of the typical particle varies with time as $\exp(-i\omega\tau)$. Also we express the number of particles per unit volume in the form $(N/3\pi)(3mc^2/2e^2)^3$, where N is also a magnitude without dimensions. Then from (53) we obtain the refractive index as a function of frequency in the form:

$$n(\omega) = [1+2N/(\omega_0^2-\omega^2-i\omega^3)]^{\frac{1}{2}}. \quad (54)$$

Here we have introduced the abbreviation

$$\omega_0^2 = (2e^2/3mc^3)^2(\text{constant}/m)-(2N/3), \quad (55)$$

where ω_0 is a measure of the extent to which the assumed quasi-elastic force over-compensates the otherwise inherent electrical instability of the system.

The propagation through a vacuum of an electrical disturbance of circular frequency $(3mc^3/2e^2)\omega$ is conveniently described by an electrical field of the form

$$\exp(i\omega\xi-i\omega\tau), \quad (56)$$

when we use the quantity $(2e^2/3mc^2)\xi$ as a measure of distance in the direction of propagation. We suppose this disturbance to be incident on a medium occupying the infinite half-space from $\xi=0$ to $\xi=+\infty$. Then the transverse electric field of the monochromatic wave will be represented in the medium by the expression

$$2(n+1)^{-1}\exp(i\omega n\xi-i\omega\tau). \quad (57)$$

As a measure of the disturbance in the medium we shall take the displacement of the typical particle or, what is up to a constant the same thing, the polarization:

$$\begin{aligned}P &= (n^2-1)E/4\pi, \\ &= (2\pi)^{-1}(n-1)\exp(i\omega n\xi-i\omega\tau). \end{aligned} \quad (58)$$

We are interested in following the progress through the medium, not of a monochromatic wave, but of an initially well defined pulse. We shall idealize the incoming electric field as a delta function, $\delta(\xi-\tau)$, with the property $\delta(u)=0$ when

$$u\neq 0, \quad \int_{-\infty}^{+\infty}\delta(u)du = 1.$$

We recall the representation of the delta function as a superposition of monochromatic waves:

$$\delta(\xi-\tau) = (2\pi)^{-1}\int_{-\infty}^{+\infty}\exp(i\omega\xi-i\omega\tau)d\omega. \quad (59)$$

This expression will represent the electric field in the vacuum. In the medium the electric polarization will accordingly be given as a function of position and time by the integral

$$\begin{aligned}P(\xi,\tau) = (2\pi)^{-2}\int_{-\infty}^{+\infty}&(n(\omega)-1) \\ &\times\exp(i\omega n\xi-i\omega\tau)d\omega, \end{aligned} \quad (60)$$

where the refractive index is obtained as a function of frequency from (54).

Of the mathematical details of evaluating the polarization of the medium from Eq. (60) it is enough to say that it is convenient to displace the path of integration in the complex plane, and to apply the familiar saddle point method of approximation. This procedure is sufficiently accurate for our purpose when we accept the following reasonable conditions:

(1) We consider a medium of macroscopic dimensions. With the quantity $(2e^2/3mc^2)$ equal to 1.88×10^{-13} cm, it follows that the values of the quantities ξ and τ are of the order of magnitude of 10^{+13}.

(2) The number, N, of particles per volume element $3\pi(2e^2/3mc^2)^3$ is of the order of, or greater than, unity.

(3) At the depth, $\xi(2e^2/3mc^2)$, in the medium, the disturbance, if propagated with the speed of light in vacuum, would arrive at that time, $\tau(2e^2/3mc^2)$, for which $\tau=\xi$; or more briefly, we shall say that the value of the "light-instant," τ, at the depth ξ is given by the equation $\tau=\xi$. We limit our interest to the disturbance at times ahead of the light-instant by an amount which, expressed in the dimensionless measure, $\xi-\tau$, is small in comparison with $\xi\sim10^{13}$, although this difference may otherwise range all the way from a value very small in comparison with unity to a value as great as several orders of magnitude of 10.

(4) The dimensionless measure of natural frequency of oscillation of the system, ω_r, in order of magnitude is not large in comparison with unity.

Under these conditions we obtain an approximate representation of the polarization of the medium in the form

$$4\pi^2P\doteqdot(\pi N/\xi)^{\frac{1}{2}}[(1/2N\xi)-(\xi-\tau)+\cdots], \quad (61)$$

for values of $(\xi-\tau)$ in a range of order $1/(N\xi)^{\frac{1}{2}}$ on either side of the light-instant; and

$$4\pi^2P\doteqdot(4\pi N/3\xi)^{\frac{1}{2}}[(\xi-\tau)/2N\xi]^{\frac{1}{2}} \exp[-(3/4)(2N\xi/\xi-\tau)^{\frac{1}{2}}(\xi-\tau)] \cos[(\pi/3)+(3^{\frac{1}{2}}/4)(2N\xi/\xi-\tau)^{\frac{1}{2}}(\xi-\tau)], \quad (62)$$

for values of $\xi-\tau$ between the rough limits $1/(N\xi)^{\frac{1}{2}}$ on the little side and some small fraction of the quantity $N\xi$ on the big side.

We obtain from expressions (61) and (62) the following picture of the displacement of the charges of the medium at the depth $\xi(2e^2/3mc^2)$ before and at the light-instant:

(1) The typical particle receives a displacement before the light-instant, thus justifying the use of the descriptive term "pre-acceleration" even in the case of a medium containing many particles.

(2) The displacement of the typical particle, instead of increasing with time according to the simple exponential law, $\exp(\tau-\xi)$, derived by Dirac for an isolated particle, is here before the light-instant an oscillatory function of time of much more rapidly increasing amplitude.

(3) The last full oscillation before the light-instant is in the negative sense, that is, opposite to the direction of the field in the original pulse. This oscillation is completed only slightly before the light-instant, so at that time the displacement of the typical particle is positive but small in comparison with its magnitude in the last few preceding vibrations. However, the velocity of the particle at a time about equal to the light-instant has reached the maximum value so far experienced. The condition of approximately zero displacement and high velocity has a certain correspondence with the result which would be expected at the time of arrival of the disturbance in the absence of the phenomenon of pre-acceleration.

(4) The characteristic time of pre-acceleration may reasonably be taken to be measured by the interval between the last two nodes of the oscillation, a quantity which has the order of magnitude $(N\xi)^{-\frac{1}{2}}(2e^2/3mc^2)$, a very small fraction of the so-called classical radius of the charged particle. Another estimate for the time of pri-acceleration of the same order of magnitude is obtained by studying the exponentially increasing envelope of the oscillatory motion described by Eq. (60).

From the tentative conception that the classical radius of a charged particle defines a region within which disturbances are propagated with a speed faster than the velocity of light, it would have appeared reasonable to expect in a very dense medium a macroscopic velocity of propagation significantly greater than the normal limiting value. If this were the case, the interval of pre-acceleration, $\xi-\tau$, would have increased in proportion to the depth, ξ, and would have been appreciable in comparison with ξ. In contrast, we have now found that the characteristic time of pre-acceleration not only decreases slowly with depth in a dense medium, but also is an exceedingly small fraction of the value obtained by Dirac for the case of a single particle. We conclude that it is misleading to attribute the phenomenon of pre-acceleration to an abnormal velocity of light or to a failure of the usual conceptions of space-time in the immediate neighborhood of a charged particle. We are therefore obliged to look to other terms for a suitable way to describe the phenomenon.

PRE-ACCELERATION AS WITNESS TO THE INTERACTION OF PAST AND FUTURE

Pre-acceleration and the force of radiative reaction which calls it forth are both departures from that view of nature for which one once hoped, in which the movement of a particle at a given instant would be completely determined by the motions of all other particles at earlier moments. All thought was excluded of a dependence of the force experienced by the particle upon the future behavior of either that charge itself or any other charges. The past was considered to be completely independent of the future. This idealization is no longer valid when

we have a particle commencing to move in anticipation of the retarded fields which have yet to reach it from surrounding charges. Still less is it a good approximation to the truth in the case of an incompletely absorbing system, where we have in addition to the normal damping force an incident field seen above to depend explicitly upon the advanced fields of the individual particles, and where we encounter advanced effects even more striking than preacceleration.

The mechanism by which the future affects the past is illuminated by considering a system of three or more charges in the light of the half-advanced, half-retarded fields of the theory of action at a distance. The retarded field produced by the acceleration of a affects b; the advanced field of b sets c in motion; and c generates a field, the advanced part of which affects a before the moment of its acceleration. By an extension of this line of reasoning it is apparent that the past and future of all particles are tied together by a maze of interconnections. The happenings in neither division of time can be considered to be independent of those in the other. Nevertheless, in a system containing particles sufficient in number to provide effective absorption, an interference takes place between these forces. All the advanced effects are cancelled out except those which are comprised in the conventional force of radiative reaction; and these are limited in their influence to a time of the order of magnitude of the quantity (e^2/mc^3). Therefore, to the extent that the force of radiative reaction can be neglected, we have in the case of a completely absorbing system the possibility to make a cut between past and future; but the cleanness of this cut is limited to times of the order of e^2/mc^3 or greater. Those phenomena which take place in times shorter than this figure require us to recognize the complete interdependence of past and future in nature, an interdependence due to an elementary law of interaction between particles which is perfectly symmetrical between advanced and retarded fields.

SUMMARY

Use of action at a distance with field theory as equivalent and complementary tools for the description of nature has so far been prevented by inability of the first point of view fully to account for the mechanism of radiation. Elucidation of this process in both theories comes from a 23-year old suggestion of Tetrode, that the absorber may be an essential element of the act of emission. A quantitative formulation of this idea is given here on the basis of the following postulates: (1) An accelerated charge in otherwise charge-free space does not radiate energy. (2) The fields which act on a given particle arise only from other particles. (3) These fields are represented by one-half the retarded plus one-half the advanced Lienard-Wiechert solutions of Maxwell's equations.

In a system containing particles sufficient in number ultimately to absorb all radiation, the absorber reacts upon an accelerated charge with a field, the advanced part of which, evaluated in the neighborhood of the source on the basis of these postulates, is found to have the following properties: (1) It is independent of the properties of the absorbing medium. (2) It is completely determined by the motion of the source. (3) It exerts on the source a force which is finite, is simultaneous with the moment of acceleration, and is just sufficient in magnitude and direction to take away from the source the energy which the act of radiation imparts to the surrounding particles. (4) It is equal in magnitude to one-half the retarded field minus one-half the advanced field of the accelerated charge itself, just the field postulated by Dirac as the source of the force of radiative reaction. (5) This field compensates the half-advanced field of the source and combines with its half-retarded field to produce the full retarded disturbance which is required by experience. Radiation is concluded to be a phenomenon as much of statistical mechanics as of pure electrodynamics. A complete correspondence is established between action at a distance and the usual formulation of field theory in the case of a completely absorbing system. In such a system the phenomenon of pre-acceleration appears as the sole evidence of the advanced effects of the theory of action at a distance. Other advanced effects appear in the case of an incompletely absorbing system and are also discussed.

REVIEWS OF MODERN PHYSICS VOLUME 21, NUMBER 3 JULY, 1949

Classical Electrodynamics in Terms of Direct Interparticle Action[1]

JOHN ARCHIBALD WHEELER AND RICHARD PHILLIPS FEYNMAN[2]
Princeton University, Princeton, New Jersey

". . . the energy tensor can be regarded only as a provisional means of representing matter. In reality, matter consists of electrically charged particles. . . ."[3]

INTRODUCTION AND SUMMARY

MANY of our present hopes to understand the behavior of matter and energy rely upon the notion of field. Consequently it may be appropriate to re-examine critically the origin and use of this century-old concept. This idea developed in the study of classical electromagnetism at a time when it was considered appropriate to treat electric charge as a continuous substance. It is not obvious that general acceptance in the early 1800's of the principle of the atomicity of electric charge would have led to the field concept in its present form. Is it after all essential in classical field theory to require that a particle act upon itself? Of quantum theories of fields and their possibilities we hardly know enough to demand on quantum grounds that such a *direct* self-interaction should exist. Quantum theory defines those possibilities of measurement which are consistent with the principle of complementarity, but the measuring devices themselves after all necessarily make use of classical concepts to specify the quantity measured.[4] For this reason it is appropriate to begin a re-analysis of the field concept by returning to classical electrodynamics. We therefore propose here to go back to the great basic problem of classical physics—the motion of a system of charged particles under the influence of electromagnetic forces—and to inquire what description of the interactions and motions is possible which is at the same time (1) well defined (2) economical in postulates and (3) in agreement with experience.

We conclude that these requirements are satisfied by the theory of action at a distance of Schwarzschild,[5] Tetrode,[6] and Fokker.[7] In this description of nature no direct use is made of the notion of field. Each particle moves in compliance with the principle of stationary action,[8]

$$J=-\sum_a m_a c\int(-da_\mu da^\mu)^{\frac{1}{2}}+\sum_{a<b}(e_a e_b/c)\times\int\int\delta(ab_\mu ab^\mu)(da_\nu db^\nu)=\text{extremum}. \quad (1)$$

All of mechanics and electrodynamics is contained in this single variational principle.

However unfamiliar this direct interparticle treatment compared to the electrodynamics of Maxwell and Lorentz, it deals with the same problems, talks about the same charges, considers the interaction of the same current elements, obtains the same capacities, predicts the same inductances and yields the same physical conclusions. Consequently action at a distance must have a close connection with field theory. But never does it consider the action of a charge on itself. The theory of direct interparticle action is equivalent, not

[1] Part II of a critique of classical field theory of which another part here referred to as III appeared in Rev. Mod. Phys. **17**, 157 (1945). For related discussion see also R. P. Feynman, Phys. Rev. **74**, 1430 (1948).

[2] Now at Cornell University, Ithaca, N. Y.

[3] A. Einstein, *The Meaning of Relativity* (Princeton University Press, Princeton, New Jersey, 1945), second edition, p. 82.

[4] See in this connection Niels Bohr, *Atomic Theory and the Description of Nature* (Cambridge University Press, 1934) and chapter by Bohr in *Einstein*, of the *Living Philosophers Series* (Northwestern University, scheduled for 1949).

[5] K. Schwarzschild, Göttinger Nachrichten, 128, 132 (1903).

[6] H. Tetrode, Zeits. f. Physik **10**, 317 (1922).

[7] A. D. Fokker, Zeits. f. Physik **58**, 386 (1929); Physica **9**, 33 (1929) and **12**, 145 (1932).

[8] Here the letters a, $b\cdots$ denote the respective particles. Particle a has in c.g.s. units a mass of m_a grams, a charge of e_a franklins (e.s.u.), and has at a given instant the coordinates

$a^1=a_1$, $a^2=a_2$, $a^3=a_3$ } the three space coordinates, measured in cm.

$a^4=-a_4$, a quantity which has also the dimensions of a length, and which represents the product of the time coordinate by the velocity of light, c (ct="cotime").

(Note: In comparing formulas here with those in the literature, note that not all authors use the same convention about signs of covariant and contravariant components.)

The expression ab^m is an abbreviation for the vector, a^m-b^m. Greek indices indicate places where a summation is understood to be carried out over the four values of a given label. The argument $ab_\mu ab^\mu$ of the delta-function thus vanishes when and only when the locations of the two particles in space-time can be connected by a light ray. Here the delta-function $\delta(x)$ is the usual symbolic operator defined by the conditions $\delta(x)=0$ when $x\neq 0$ and $\int_{-\infty}^{+\infty}\delta(x)dx=1$. In the evaluation of the action, J, from (1), the world lines of the several particles are considered to be known for all time; i.e., the coordinates a^m are taken to be given functions of a single parameter, a, which increases monotonically along the world line of the first particle; likewise for b, c, etc. An arbitrary assumed motion of the particles is not in general in accord with the variation principle: a small change of the first order, $\delta a^m(a)$, $\delta b^m(b)$, $\cdots$ in the world lines of the particles (this change here being limited for simplicity to any finite interval of time, and the length of this time interval later being increased without limit) produces in general a non-zero variation of the first order, δJ, in J itself. Only if all such first order variations away from the originally assumed motion produce no first order change in J is that originally assumed motion considered to satisfy the variational principle. It is such motions which are in this article concluded to be in agreement with experience.

to the usual field theory, but to a modified or *adjunct field theory*, in which

(1) the motion of a given particle is determined by the sum of the fields produced by—or *adjunct to*—every particle *other* than the given particle.

(2) the field adjunct to a given particle is uniquely determined by the motion of that particle, and is given by half the retarded plus half the advanced solution of the field equations of Maxwell for the point charge in question.

This description of nature differs from that given by the usual field theory in three respects:

(1) There is no such concept as "the" field, an independent entity with degrees of freedom of its own.

(2) There is no action of an elementary charge upon itself and consequently no problem of an infinity in the energy of the electromagnetic field.

(3) The symmetry between past and future in the prescription for the fields is not a mere logical possibility, as in the usual theory, but a postulational requirement.

There is no circumstance of classical electrodynamics which compels us to accept the three excluded features of the usual field theory. Indeed, as regards the question of the action of a particle upon itself, there never was a consistent theory, but only the hope of a theory. It is therefore appropriate now and hereafter to formulate classical electrodynamics in terms of the adjunct field theory or the theory of direct interparticle action. The agreement of these two descriptions of nature with each other and with experience assures us that we arrive in this way at *the natural and self-consistent generalization of Newtonian mechanics to the four-dimensional space of Lorentz and Einstein.*

It is easy to see why no unified presentation of classical electrodynamics along these lines has yet been given, though the elements for such a description are all present in isolated form in the literature. The development of electromagnetic theory came before the era of relativity. Most minds were not prepared for the requirement that interactions should be propagated with a certain characteristic speed, still less for the possibility of both advanced and retarded interactions. Newtonian instantaneous action at a distance with its century and a half of successes seemed the natural framework about which to construct a description of electromagnetism. Attempt after attempt failed.[9] And unfortunately uncompleted was the work of Gauss, who wrote to Weber on the 19th of March, 1845: "I would doubtless have published my researches long since were it not that at the time I gave them up I had failed to find what I regarded as the keystone, *Nil actum reputans si quid superesset agendum:* namely, the derivation of the additional forces—to be added to the interaction of electrical charges at rest, when they are both in motion—from an action which is propagated not instantaneously but in time as is the case with light."[10] These failures and the final success via the apparently quite different concept of field were taken by physicists generally as convincing arguments against electromagnetic action at a distance.

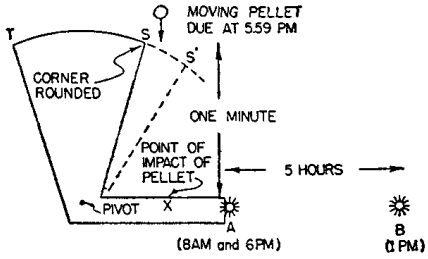


FIG. 1. The paradox of advanced effects. Does the pellet strike X at 6 p.m.? If so, the advanced field from A sets B in motion at 1 p.m., and B moves A at 8 a.m. Thereby the shutter TS is set in motion and the path of the pellet is blocked, so it cannot strike X at 6 p.m. If it does not strike X at 6 p.m., then its path is not blocked at 5.59 p.m. via this chain of actions, and therefore the pellet ought to strike X.

Field theory taught gradually and over seven decades difficult lessons about constancy of light velocity, about relativity of space and time, about advanced and retarded forces, and in the end made possible by this circuitous route the theory of direct interparticle interaction which Gauss had hoped to achieve in one leap. On this route and historically important was Liénard[11] and Wiechert's[12] derivation from the equations of Maxwell of an expression for the elementary field generated by a point charge in an arbitrary state of motion. With this expression as starting point Schwarzschild arrived at a law of force between two point charges which made no reference to field quantities. Developed without benefit of the concept of relativity, and expressed in the inconvenient notation of the prerelativistic period, his equations of motion made no appeal to the physicists of the time. After the advent of relativity Schwarzschild's results were rederived independently by Tetrode and Fokker. These results are most conveniently summarized in Fokker's principle of stationary action of Eq. (1).

To investigate the consistency of the Schwarzschild-Tetrode-Fokker theory of direct interparticle interaction and its relation to field theory, we have first to

[9] For a stimulating and instructive if not always objective account of early researches on field theory and action at a distance see A. O'Rahilly, *Electromagnetics* (Longmans, Green and Company, New York (1938)). See also J. J. Thomson, Report of the British Assn. for the Adv. of Science for 1885, p. 97; J. C. Maxwell, *Electricity and Magnetism* (Oxford University Press, London, 1892), third edition, Chapter 23); R. Reif and A. Sommerfeld, Encyclopädie der Math. Wiss. 5, Part 2, Section 12 (1902). A recent very brief account has been given by H. J. Groenewold, report on Puntladingen en stralingsveld, Ned. Nat. Ver., Amsterdam (May 1947). M. Schönberg regards field and direct action not as two equivalent representations of the same force, but as two different parts of the total force: Phys. Rev. 74, 738 (1948); Sum. Bras. Math. 1, Nos. 5 and 6 (1946); J. L. Lopes and M. Schönberg, Phys. Rev. 67, 122 (1945).

[10] C. F. Gauss, *Werke* 5, 629 (1867).

[11] A. Liénard, L'Éclairage Électrique 16, pp. 5, 53, 106 (1898).

[12] E. Wiechert, Archives Néerland (2) 5, 549 (1900); Ann. d Physik 4, 676 (1901). Compare these derivations in prerelativistic notation with that given for example by W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, New York, 1944), second edition, p. 19, or A. Sommerfeld, Ann. d. Physik 33, 668 (1910).

examine in the next section the paradox of advanced interactions. In the following section is recalled the derivation of the equations of motion from the variation principle. Next these equations of motion are shown to satisfy the principle of action and reaction as generalized to the non-instantaneous forces of a relativistic theory of action at a distance. In a subsequent section the corresponding formulation of the laws of conservation of energy and momentum is given. Finally the connection is established between these conservation laws and the field-theoretic description of a stress-energy tensor defined throughout space and time.

THE PARADOX OF ADVANCED ACTIONS

The greatest conceptual difficulty presented by the theory of direct interparticle interaction is the circumstance that it associates with the retarded action of a on b, for example, an advanced action of b on a. A description employing retarded forces alone would violate the law of action and reaction or, in mathematical terms, could not be derived from a single principle of stationary action.

Advanced actions appear to conflict both with experience and with elementary notions of causality. Experience refers not to the simple case of two charges, however, but to a universe containing a very large number of particles. In the limiting case of a universe in which all electromagnetic disturbances are ultimately absorbed it may be shown[1] that the advanced fields combine in such a way as to make it appear—except for the phenomenon of radiative reaction—that each particle generates only the usual and well-verified retarded field. It is only necessary to make the natural postulate that we live in such a completely absorbing universe to escape the apparent contradiction between advanced potentials and observation.

In a universe consisting of a limited number of charged particles advanced effects occur explicitly. It is no objection if the character of physics under such idealized conditions conflicts with our experience. It is only required that the description should be logically self-consistent. In particular in analyzing the behavior of an idealized universe containing only a few particles we cannot introduce the human element as we know it into the systems under study. To do so would be to assume tacitly the possibility of a clean cut separation between the effects of past and future. This possibility is denied in a description of nature in which both advanced and retarded effects occur explicitly.

The apparent conflict with causality begins with the thought: If the present motion of a is affected by the future motion of b, then the observation of a attributes a certain inevitability to the motion of b. Is not this conclusion in direct conflict with our recognized ability to influence the future motion of b?

All essential elements of the general paradox appear in the following idealized example: Charged particles a and b are located in otherwise charge-free space at a distance of 5 light-hours. A clockwork mechanism is set to accelerate a at 6 p.m. Thereby b will be affected, not only at 11 p.m. via retarded effects, but also at 1 p.m. via advanced forces. This afternoon motion will cause a to suffer a premonitory movement at 8 a.m. Seeing this motion in the morning, we conclude the clockwork will go off in the evening. We return to the scene a few seconds before 6 p.m. and block the clockwork from acting on a. But then why did a move in the morning?

To formulate the paradox acceptably, we have to eliminate human intervention. We therefore introduce

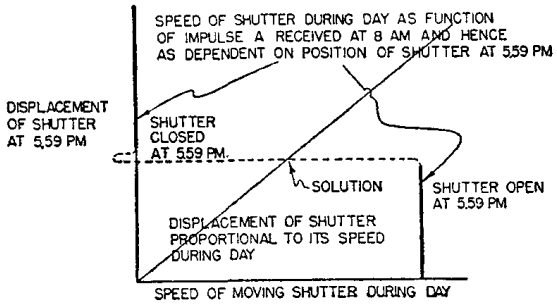


FIG. 2. Analysis and resolution of the paradox of advanced effects. The action of the shutter on the pellet—the interaction of past and future—is continuous (dashed line in diagram) and the curves of action and reaction cross. See text for physical description of solution.

a mechanism which saves charge a from a blow at 6 p.m. only if this particle performs the expected movement at 8 a.m. (Fig. 1). Our dilemma now is this: Is a hit in the evening or is it not? If it is, then it suffered a premonitory displacement at 8 a.m. which cut off the blow, so a is *not* struck at 6 p.m.! If it is not bumped at 6 p.m. there is no morning movement to cut off the blow and so in the evening a *is* jolted!

To resolve, we divide the problem into two parts: effect of past of a upon its future, and of future upon past. The two corresponding curves in Fig. 2 do not cross. We have no solution, because the action of the shutter on the pellet, of the future on the past, has been assumed discontinuous in character.

The paradox, and the case it presents against advanced potentials, evidently depends on the postulate that discontinuous forces can exist in nature. From a physical point of view we are led to make just the contrary assumption, that the influence of the future upon the past depends in a continuous manner upon the future configuration.

Our general assumption about continuity is explicitly verified in the present case. The action of shutter on pellet is not discontinuous. The pellet will strike the point S a glancing blow if the shutter lies only part way across its path (dashed curve in Fig. 2).

Of the problem of influence of future upon past, and past upon future, we now have in Fig. 2 a self-consistent solution: Charge a by late afternoon has moved a

very slight distance athwart the path of the pellet. Thus one second before 6 p.m. it receives a glancing blow in the counter-clockwise sense and at 6 p.m. a stronger acceleration in the clockwise direction. The accelerations received by a at these two moments are by electromagnetic interaction transmitted in reduced measure to b at 1 p.m. and back from b in yet greater attenuation to a. Thus this particle receives one second before 8 a.m. a certain counter-clockwise impulse and at 8 a.m. an opposite impulse. The net rotational momentum imparted to the lever is clockwise. It carries the point S in the course of 10 hours the necessary distance across the path of the pellet. The chain of action and reaction is completed. The paradox is resolved.

Generalizing, we conclude advanced and retarded interactions give a description of nature logically as acceptable and physically as completely deterministic as the Newtonian scheme of mechanics. In both forms of dynamics the distinction between cause and effect is pointless. With deterministic equations to describe the event, one can say: the stone hits the ground because it was dropped from a height; equally well: the stone fell from a height because it was going to hit the ground.

The distinction between Newtonian and relativistic mechanics is one of detail—instantaneous interactions *versus* forces unconfined to a single plane in space time. The interrelations between the world lines are more complicated than those of Newtonian mechanics, but just as definite. There a well-defined division of past and present was possible; here these divisions of time are inextricably mixed.

EQUATIONS OF MOTION

Advanced and retarded forces being accepted on equal footing in the description of nature, we now reproduce the derivation from Fokker's action principle of equations of motion which contain them both. Let the world line of a typical particle a be altered from $a^m(a)$ to $a^m(a)+\delta a^m(a)$. Let the abbreviation be introduced,

$$A_m^{(b)}(x)=e_b\int\delta(xb_\mu xb^\mu)db_m(b) \tag{2}$$

(vector potential of particle b at point x). Also denote by $a^{m\prime}(a)$ the derivative $da^m(a)/da$. Then the change in action produced by the alteration of the world line of a is

$$\delta J=m_ac\int\{a_\mu'(\delta a^\mu)'/(-a_\nu'a^{\nu\prime})^{\frac{1}{2}}\}da + \delta\sum_{b\neq a}(e_a/c)\int A_\nu^{(b)}(a)a^{\nu\prime}(a)da,$$

or, by partial integration, and dropping terms at the limits where the variations δa^m vanish,

$$\delta J=\int da\sum_m\delta a^m(a)\{-m_ac(d/da)[a_m'/(-a_\nu'a^{\nu\prime})^{\frac{1}{2}}] + (e_a/c)\sum_{b\neq a}[(\partial A_\rho^{(b)}/\partial a^m)-(\partial A_m^{(b)}/\partial a^\rho)]a^{\rho\prime}\}. \tag{3}$$

The condition that δJ be zero to the first order for arbitrary δa^m is the vanishing of the curly bracket in (3) for all four values of m, whence result the four components of the equation of motion for particle a. Instead of expressing the motion in terms of the arbitrary parameter a, introduce a new parameter, $\alpha=\alpha(a)$, the "proper cotime," defined in terms of a up to an unimportant additive constant by the equation $d\alpha/da=(-a_\nu'a^{\nu\prime})^{\frac{1}{2}}$ and denote by dots derivatives with respect to the proper cotime. Introduce also the abbreviation

$$F_{mn}^{(b)}(x)=\partial A_n^{(b)}(x)/\partial x^m-\partial A_m^{(b)}(x)/\partial x^n \tag{4}$$

(field at point x due to b).[13] Then the four-vector *equation of motion* takes a form,

$$m_ac^2\ddot{a}_m=e_a\sum_{b\neq a}F_{m\rho}^{(b)}(a)\dot{a}^\rho, \tag{5}$$

identical with that of Lorentz, with the following exceptions: self-actions are explicitly excluded; no fields act except those adjunct to the other particles; each such adjunct field is uniquely determined by the prescription of Eqs. (2) and (4).

Now we come to the well known proof that each adjunct field satisfies Maxwell's equations when for charge and current are introduced the appropriate expressions for the given particles. We employ Dirac's identity[14]

$$(\partial^2/\partial x_\mu\partial x^\mu)\delta(xb_\nu xb^\nu)=-4\pi\delta(x_1-b_1)\delta(x_2-b_2)\times\delta(x_3-b_3)\delta(x_4-b_4), \tag{6}$$

multiply both sides by $db_m(\beta)=\dot{b}_m(\beta)d\beta$, integrate with respect to β from $-\infty$ to $+\infty$, and conclude that $A_m^{(b)}(x)$ satisfies the equation

$$(\partial^2/\partial x_\mu\partial x^\mu)A_m^{(b)}(x)=-4\pi j_m^{(b)}(x). \tag{7}$$

Here

$$j_m^{(b)}(x)=e_b\int_{-\infty}^{+\infty}\delta(x_1-b_1)\delta(x_2-b_2)\times\delta(x_3-b_3)\delta(x_4-b_4)\dot{b}_m(\beta)d\beta \tag{8}$$

is an abbreviation for the density-current four-vector at point x due to particle b, an obviously singular quantity, obeying certain evident conservation rela-

[13] The electric field E_x is $F_{14}=-F_{41}$ and the magnetic field H_x is $F_{23}=-F_{32}$: the vector potential A_x is $A^1=A_1$ and the scalar potential is $A^4=-A_4$. Likewise in Eq. (8) $j^4=-j_4$ represents the charge density in franklins (e.s.u.)/cm³ and $j^1=j_1$ gives $(1/c)$ times (x-component of the charge flux in franklins/cm² sec.).

[14] P. A. M. Dirac, Proc. Roy. Soc. London. A167, 148 (1938).

tions. The vector potential (2), in addition to satisfying the inhomogeneous wave equation (6), has a four-dimensional divergence which vanishes:

$$(\partial/\partial x_\mu)A_\mu^{(b)}(x)=e_b\int_{-\infty}^{+\infty}\delta'(xb_\nu xb^\nu)2xb^\mu\dot{b}_\mu d\beta$$
$$=2e_b\delta(xb_\nu xb^\nu)\Big|_{-\infty}^{+\infty}=0. \quad (9)$$

We differentiate this zero divergence with respect to x^m and subtract from it (7), obtaining the field equations

$$\partial F_{m\mu}^{(b)}(x)/\partial x_\mu=4\pi j_m, \quad (10)$$

equivalent to the usual relations $\mathrm{div}E=4\pi\rho$ and $\mathrm{curl}H=\dot{E}/c+4\pi J_x/C$. The other pair of Maxwell's equations follow identically from the definition (4) of the F's in terms of the A's.

The fields (2) are distinguished from all other solutions of Maxwell's equations by being half the sum of the advanced and retarded Liénard-Wiechert potentials of particle b:

$$A_m^{(b)}(x)=e_b\int\frac{db_m/d\beta}{d(xb_\mu xb^\mu)/d\beta}\delta(xb_\nu xb^\nu)d(xb_\rho xb^\rho)$$
$$=(1/2)R_m^{(b)}(x)+(1/2)S_m^{(b)}(x). \quad (11)$$

Here, for example, R represents the retarded potential

$$R^{m(b)}(x)=e_b\dot{b}^m/\dot{b}^\mu bx_\mu, \quad (12)$$

evaluated at that point on the world line of b which intersects the light cone drawn from the point of observation into the past:

$$xb_\mu xb^\mu=0;\quad x^4>b^4, \quad (13)$$

and S similarly represents the advanced potential.

By way of illustration of these results in familiar cases consider first the case of a point charge, b, at rest at the origin. Then retarded and advanced fields are identical, all components of the four-potential vanish except the last, $\dot{b}^4=d(\text{cotime})/d(\text{proper cotime})=1$, $bx_4=b_4-x_4=x^4-b^4=$elapsed cotime$=$distance to point of observation$=r$, and the scalar potential has the familiar value e_b/r. Next, in the case of a slowly moving point charge, it similarly follows that $A^m=e_b(\dot{b}^m/2r)_{\text{ret}}+e_b(\dot{b}^m/2r)_{\text{adv}}$. If this point charge is at the same time being accelerated, then the derived electric field has at large distances the value $E=-e_b(\ddot{b}_\perp/2r)_{\text{ret}}-e_b(\ddot{b}_\perp/2r)_{\text{adv}}$, where $b_\perp$ is the component of the three-vector b perpendicular to the line r. This result refers only to the field of the particle in question. In the idealized case of a universe containing charged particles sufficient in number to absorb all electromagnetic disturbances, the advanced fields of the particles of the absorber will combine with the given field to produce the full retarded field of experience, $-e_b(\ddot{b}_\perp/r)_{\text{ret}}$, as shown in III. As a final example consider a fixed linear conductor past any point of which flow per second i/e particles of charge e. The interval of cotime between the kth and the $(k+1)$st particle is ce/i. The coordinates of the kth particle are

$$k^m(\gamma)=s^m(\gamma)\quad(m=1,2,3)$$
$$k^4(\gamma)=s^4(\gamma)+kce/i\quad(k=-\infty,\cdots,-1,0,1,\cdots) \quad (14)$$

where $s^m(\gamma)$ is the parametric representation of the curve of the wire. The four-potential at a point of observation an appreciable distance from the wire is obtained by summing over all the particles or equivalently, because of the close spacing of the charges, by integrating over k:

$$A^m(x)=e\int\int\delta[r^2(\gamma)-(ct-s^4(\gamma)-kce/i)^2]$$
$$\times dk(ds^m(\gamma)/d\gamma)d\gamma$$
$$=\begin{cases}i\int ds^m(\gamma)/cr(\gamma) & \text{for}\quad m=1,2,3\\ e\int dk/r & \text{for}\quad m=4.\end{cases} \quad (15)$$

Here $r(\gamma)$ is the magnitude of the vector $x(\gamma)$, $y(\gamma)$, $z(\gamma)$ which runs from the point γ of the curve to the point of observation. The scalar potential of Eq. (15) will normally be compensated wholly or in part by contributions from opposite charges at rest and need not be considered here. From the vector potential follows an expression for the magnetic field

$$H=\mathrm{curl}A=i\int(\mathbf{ds}\times\mathbf{r})/cr^3, \quad (16)$$

identical with that due to Ampere.

To go further in deriving well known results would be pointless. Adequate textbooks exist. They treat well defined problems of electromagnetism, where there is no compelling reason to consider a particle to act on itself. Thus all their analyses are immediately translatable into terms of the present modified or *adjunct field theory*. However, this point of view is mathematically identical with that of action at a distance. Consequently the theory of direct interparticle action, far from attempting to replace field theory, joins with field theory to provide the science of electromagnetism with additional techniques of mathematical analysis and to facilitate deeper physical insight. The rest of this article may illustrate how the two points of view join hands to elucidate in four-dimensional mechanics the principle of action and reaction and the laws of conservation of momentum and energy.

ACTION AND REACTION

Laws of conservation of angular momentum, energy and linear momentum are well known to exist in any

theory for which the equations of motion are derivable from an action principle which is invariant with respect to rotation, translation, or displacement of the time coordinate.[15] Thus Fokker[16] has derived an energy-momentum conservation principle for an idealized situation in which there are only two particles, of which a acts on b via purely retarded forces. The present treatment is the natural generalization of Fokker's analysis to the case of a theory which is symmetric between every pair of particles and which is based on the action principle (1). It will be sufficient to prove the conservation law for a single pair of particles in order to see the corresponding result for a system of particles.

For the typical particle a let the four-vector of energy and comomentum be denoted by

$$mcv(1-v^2/c^2)^{-\frac{1}{2}}=\left\{\begin{matrix}G^1=G_1\\G^2=G_2\\G^3=G_3\end{matrix}\right\}$$ three space components of the kinetic comomentum (velocity of light times kinetic momentum: expressible in energy units).

$mc^2(1-v^2/c^2)^{-\frac{1}{2}}=G^4=-G_4$, kinetic energy plus rest-mass energy.

Then the change in kinetic comomentum and energy in the interval of proper cotime, $d\alpha$, on account of the action of particle b follows directly from the equations of motion (5) and the expressions (4) and (2) for the force coefficients:

$$dG_m^{(a)}(\alpha)=m_ac^2\ddot{a}_m d\alpha=e_ae_b da^\mu \times\left\{(\partial/\partial a^m)\int\dot{b}_\mu-(\partial/\partial a^\mu)\int\dot{b}_m\right\}\delta(ab_\nu ab^\nu)d\beta. \quad (17)$$

We carry out the differentiations with respect to the coordinates a and add to the result the following zero quantity

$$e_ae_b da_m\int_{-\infty}^{+\infty}(d/d\beta)\delta(ab_\nu ab^\nu)d\beta, \quad (18)$$

thus finding for the impulse

$$dG_m^{(a)}(\alpha)=2e_ae_b\int_{\beta=-\infty}^{\beta=+\infty}\delta'(ab_\nu ab^\nu) \times(ab_m da^\mu db_\mu-db_m da^\mu ab_\mu-da_m db^\mu ab_\mu). \quad (19)$$

In this expression the integrand is changed in sign but unaltered in value by an interchange of the roles of particles a and b.

To the result just obtained we give the following obvious interpretation:

(1) The right hand side of (19), after removal of the integral sign, represents in terms of the symbolic delta-function the transfer of impulse or energy to a during the stretch of cotime $d\alpha$ from effects which originate at b in the cotime interval $d\beta$.

(2) There is no energy or impulse transfer except when the stretch $d\beta$ of the world line of b is intersected by either the forward or backward light cone drawn from a: i.e., b acts on a through both retarded and advanced forces.

(3) The impulse communicated to a over the portion $d\alpha$ of its world line via retarded forces, for example, from the stretch $d\beta$ of the world line of b is equal in magnitude and opposite in sign to the impulse transfer from a to b via advanced forces over the same world line intervals (*equality of action and reaction*).

The relativistic generalization of the Newtonian principle of action and reaction as just stated is obviously not identical with the non-relativistic formulation. In no Lorentz frame of reference are action and reaction simultaneous. For the instant at which a experiences a force from b there is not one corresponding time at which b gets a back reaction, but two instants.[17] Thus for a given point on the world line of a we can make two statements about the transfer of energy (or impulse) from b. Each statement refers to a single one of the two parts of the total transfer. It is evidently reasonable that the law of action and reaction should have this Jacob's ladder character in 4-dimensional space-time.

ENERGY AND MOMENTUM OF INTERACTION

Considering two isolated particles a and b, we immediately conclude from the law of action and reaction as just stated the constancy in time of the *total energy and comomentum four-vector*

$$G_m(\alpha,\beta)=m_ac^2\dot{a}_m(\alpha)+m_bc^2\dot{b}_m(\beta) +2e_ae_b\left\{-\int_{-\infty}^{\alpha}\int_{\beta}^{\infty}+\int_{\alpha}^{\infty}\int_{-\infty}^{\beta}\right\}\delta'(ab_\nu ab^\nu)$$

$$(ab_m da^\mu db_\mu-db_m da^\mu ab_\mu-da_m db^\mu ab_\mu)=(\text{constant})_m. \quad (20)$$

In the case of more particles we have a corresponding expression with a kinetic term for each individual particle and an interaction term for each pair of charges. Thus G_m becomes a function of as many parameters $\alpha, \beta, \gamma, \cdots$ as there are particles. To prove constancy with respect to a given parameter, such as α, we have only to differentiate (20) and insert for $m_ac^2\ddot{a}_m(\alpha)$ the quotient $dG_m^{(a)}(\alpha)/d\alpha$ obtained from (19).

Evidently we have in (20) what may be called a many-time formulation of the conservation laws, derived of course from the equations of motion, but from which conversely the equations of motion are derivable with equal ease.

The interpretation of the double integral in (20) as an interaction energy is obvious in the case of two stationary charges separated by a distance R. Thus by integration we find for G^4 the familiar result $m_ac^2+m_bc^2+e_ae_b/R$.

[15] E. Noether, Göttinger Nachrichten, Math. Phys. Klasse. 235 (1918): E. Bessel-Hagen, Math. Ann. 84, 258 (1921).

[16] A. D. Fokker, Zeits. f. Physik 58, 386 (1929).

[17] L. Page, Am. J. Phys. 13, 141 (1945), has reviewed the complications which come from comparing action and reaction at the *same* time.

In the case of individual moving charges it is sometimes convenient to add to the idea of kinetic comomentum and energy $G_m^{(a)}$ the notion of potential comomentum and energy

$$U_m^{(a)} = e_a \sum_{b \neq a} A_m^{(b)}(a(\alpha)), \tag{21}$$

and total comomentum and energy,

$$P_m^{(a)} = G_m^{(a)} + U_m^{(a)}. \tag{22}$$

In terms of these expressions, the four-vector of energy and comomentum of the whole system takes the form

$$G_m(\alpha, \beta, \cdots) = \sum_a P_m^{(a)}(\alpha) + \sum_{a<b} 2e_a e_b \times \left\{ \int_\alpha^\infty \int_{-\infty}^\beta - \int_{-\infty}^\alpha \int_\beta^\infty \right\} \delta'(ab_\nu ab^\nu) ab_m da^\mu db_\mu. \tag{23}$$

The summation of the potential energies so to speak counts twice the interaction between each pair of particles. The double integrals in (23) correct for this overcount.

From either Eq. (20) or Eq. (23) for the energy of the system it is clear (see Fig. 3) that the electromagnetic energy of a finite number of particles is definable from a knowledge of only a finite stretch of their world lines. It is also evident that particles which come together in otherwise charge free space, interact, and then separate in a regular way, will in the end experience no net loss of energy to outer space. Both features of the four-vector G_m are reasonable in the mathematical description of a physically closed system.

RELATION OF INTERACTION ENERGY TO FIELD ENERGY

In field theory it is customary to attempt to define throughout space a symmetrical stress energy tensor[18] $T_{mn}(x)$ with the following properties:

(1) The divergence $\partial T_{m\nu}/\partial x_\nu$ vanishes at every place where there is no particle.

(2) At the location of a typical charge a this divergence becomes singular in such a way that its integral over a small volume element containing the charge gives the value of the electromagnetic force acting on that charge:

$$-\dot{a}_4 \underset{\substack{\text{neighborhood} \\ \text{of } a}}{\int\int\int} (\partial T_{m\nu}/\partial x_\nu) dx^1 dx^2 dx^3 = m_a c^2 \ddot{a}_m \tag{24}$$

when the integration extends over a region of constant time which contains a. When the integration proceeds over an arbitrary space-like region or "surface," σ, such that no pair of points in the surface can be connected by a light ray, then the corresponding statement is

$$m_a c^2 \ddot{a}_m + \dot{a}_\mu \underset{\substack{\text{neighborhood} \\ \text{of } a}}{\int\int\int} (\partial T_{m\nu}/\partial x_\nu) d\sigma^\mu = 0. \tag{25}$$

Here, if the surface is defined by a parametric representation in terms of three quantities u, v, w, then

$$d\sigma^4 = [\partial(x^1, x^2, x^3)/\partial(u, v, w)] du dv dw$$

with corresponding expressions for the other three components of $d\sigma^m$.

(3) For every space-like surface σ there is defined a *four-vector of energy and comomentum*

$$G_m(\sigma) = \Sigma_a m_a c^2 \dot{a}_m(\alpha) + \int\int\int T_{m\alpha} d\sigma^\alpha, \tag{26}$$

which is conserved in the sense that its value is completely independent of the choice of σ. Thus consider a change $\delta\sigma$ in the surface σ—i.e., an alteration from $x^m(u, v, w)$ to $x^m + \delta x^m(u, v, w)$—and the associated alterations $d\alpha, d\beta, \cdots$ in the points where the respective world lines intersect this surface. Then the change in G_m is expressible via the theorem of Gauss in terms of an integral over the volume, ω, comprised between the two surfaces:

$$\delta G_m = \Sigma_a m_a c^2 \ddot{a}_m d\alpha + \int\int\int\int (\partial T_{m\alpha}/\partial x_\alpha) d\omega. \tag{27}$$

But the integrand vanishes everywhere except in the immediate neighborhood of the typical particle, a, and there—writing $d\omega = da_\mu d\sigma^\mu$, and using (25)—we conclude that the contribution from the integral just cancels out the first term in δG_m.

Is there any choice of the tensor T_{mn} in the adjunct field theory which will yield for the energy-comomentum vector $G_m(\sigma)$ of (26) a value identical with the corresponding vector $G_m(\alpha, \beta \cdots)$ of the theory of direct interparticle action? The appropriate tensor may be constructed when one recalls that the field of a given particle is to produce changes only in the motions of the other particles, and that the principle of action and reaction connects the retarded effects exerted for example by a on b via the retarded field $(1/2)R_{mn}^{(a)}$ with the advanced effects exerted by b on a via the advanced field $(1/2)S_{mn}^{(b)}$:

$$T_{mn}(x) = \sum_{a \neq b} (R^{(a)}(x) \;\&\; S^{(b)}(x))_{mn}. \tag{28}$$

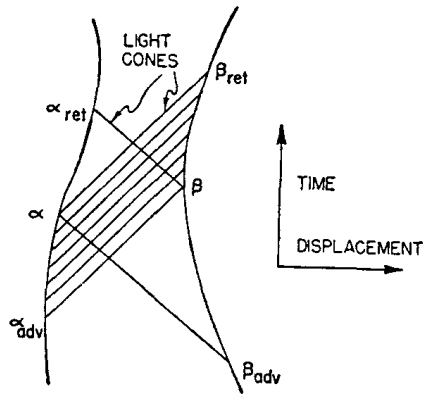


FIG. 3. Interactions considered in formulating the law of conservation of momentum and energy. Note that the stretches of world line from α_{adv} to α_{ret} and from β_{adv} to β_{ret} completely determine the value of the energy-momentum four vector $G_m(\alpha, \beta)$. It is also natural to specify these two world-line segments as initial conditions in dealing with the two-particle problem.

[18] Typical components are

T_{11}, force in positive x-direction across unit area in yx plane exerted upon medium on negative side of plane by medium on positive side (equal in the Maxwell theory to $(8\pi)^{-1} \times (H_x^2 - H_y^2 - H_z^2 + E_x^2 - E_y^2 - E_z^2)$).

T_{14}, velocity of light times energy flux in x direction per cm^2 of yz plane and per sec. (Maxwell value $(4\pi)^{-1}(E_y H_z - E_z H_y)$).

T_{44}, negative of the energy density (usual expression $-(8\pi)^{-1} \times (E^2 + H^2)$).

Here R and S denote the retarded and advanced Liénard-Wiechert fields, so that $F_{mn}{}^{(a)}=(1/2)R_{mn}{}^{(a)}+(1/2)S_{mn}{}^{(a)}$. For a convenient abbreviation we have adopted the notation

$$(R\ \&\ S)_{mn}=(R_m{}^{\mu}S_{\mu n}+S_m{}^{\mu}R_{\mu n}+\tfrac{1}{2}g_{mn}R^{\mu\nu}S_{\mu\nu})/8\pi \quad (29)$$

with $g_{mn}=0$ for $m\neq n$ and $g_{11}=g_{22}=g_{33}=1=-g_{44}$.

That the tensor T_{mn} of (28) does lead to the energy-comomentum four-vector (20) of the theory of action at a distance is proven in the appendix. Here we shall only establish that the stress energy tensor satisfies the conditions (1) and (2) (and hence (3)). Thus, we

TABLE I. Correspondence of principal alternative expressions for interaction energy in adjoint field theory and in theory of direct interparticle interaction.

	Canonical form	Frenkel form
Basic type of field coupling envisaged	Those partial fields which are reciprocally responsible for equality of action and reaction	Total (time-symmetric) field adjunct to each of the coupled particles
Typical term in stress-energy tensor	$R^{(a)}$ & $S^{(b)}$	$F^{(a)}$ & $F^{(b)}$
Expression for interaction energy:	Eq. (20)	Eq. (20) plus expression (40)
Depends upon:	Finite stretch of the two world lines	Shape of the two world lines from $t=-\infty$ to $+\infty$

evaluate the divergence of the typical term in the tensor of Eq. (28), finding

$$\begin{aligned}\partial T_{m\mu}/\partial x_{\mu}=\sum_{a\neq b}\{&(S^{(b)\rho\eta}/16\pi)(\partial R_{m\rho}{}^{(a)}/\partial x^{\eta}\\&+\partial R_{\rho\eta}{}^{(a)}/\partial x^{m}+\partial R_{\eta m}{}^{(a)}/\partial x^{\rho})\\&+(S_m{}^{(b)\rho}/8\pi)(\partial R_{\rho\sigma}{}^{(a)}/\partial x_{\sigma})\\&+\text{similar term with } S^{(b)} \text{ and}\\&R^{(a)} \text{ interchanged}\}. \quad (30)\end{aligned}$$

Here the first three cyclically related terms cancel, as seen for example from the antisymmetrical representation of the fields via potentials; and the divergence of R gives the same charge and current distribution (8) which appeared in the time-symmetric case. Using this circumstance, and combining terms, we have

$$\begin{aligned}\partial T_{m\alpha}/\partial x_{\alpha}&=\sum_{b\neq a}F_{m\mu}{}^{(b)}(x)j^{(a)\mu}(x)\\&=\sum_{b\neq a}F_{m\mu}{}^{(b)}(x)e_a\int\delta(x^1-a^1)\delta(x^2-a^2)\\&\qquad\times\delta(x^3-a^3)\delta(x^4-a^4)\dot{a}^{\mu}(\alpha)d\alpha\\&=\sum_a\int\delta(x^1-a^1)\delta(x^2-a^2)\delta(x^3-a^3)\\&\qquad\times\delta(x^4-a^4)m_ac^2\ddot{a}_m d\alpha, \quad (31)\end{aligned}$$

in complete satisfaction of requirements (1) and (2).

As alternative choice for the stress energy tensor which also has the properties (1), (2) and (3) is that proposed by Frenkel,[19] who was among the first to stress the notion of fields as always adjunct to specific particles:

$$T_{mn}{}^{*}(x)=\sum_{a\neq b}(F^{(a)}(x)\ \&\ F^{(b)}(x))_{mn}. \quad (32)$$

Thus the difference between Frenkel's tensor and the canonical tensor (28) is a quantity

$$T_{mn}{}^{*}-T_{mn}=\sum_{a\neq b}(\tfrac{1}{2}R^{(a)}-\tfrac{1}{2}S^{(a)})\ \&\ (\tfrac{1}{2}R^{(b)}-\tfrac{1}{2}S^{(b)}) \quad (33)$$

which has everywhere a zero divergence.

The possibility of more than one expression for the stress-energy tensor with the same divergence is well known in the usual single-field formulation of electrodynamics,[20] and is not surprising here. However, the expressions for field energy also turn out to differ (Table I).

The energy-comomentum four-vector G_m defined by (26) and (28), and the alternative four-vector $G_m{}^{*}$ defined by (26) and (32), are both ordinarily finite for a system of point charges. In illustration, note that near a typical particle a the corresponding field varies as $1/r^2$, the field of any other particle b is finite, the volume element is proportional to $4\pi r^2dr$ and the integral of (26) converges, yielding for example in the interaction energy e_ae_b/r_{ab} for two stationary point charges separated by the distance r_{ab}. The density of field energy, while finite, is not positive definite, even for two particles of the same charge. Also the flow of energy and momentum may have finite values at a point in space where the total field, $F^{(a)}+F^{(b)}+\cdots$, actually vanishes. This result, unexpected from the point of view of the usual field theory, nevertheless presents no logical difficulties.

ENERGY OF RADIATION

The canonical and the Frenkel tensors, which give the same interaction energy in the case of two charges which are at rest, give different results for the case of a

TABLE II. Energy flux at distance r from accelerated charge for adjunct field theory in completely absorbing universe.

Time of observation relative to moment of acceleration	Form of stress-energy tensor: Canonical	Frenkel	Maxwell
r/c seconds earlier	no flux	$-E^2/8\pi$ *towards* the source	no flux
r/c seconds later	$E^2/4\pi$ *outward*	$E^2/8\pi$ outward	$E^2/4\pi$ outward
at other times	no flux	no flux	no flux

[19] J. Frenkel, Zeits. f. Physik **32**, 518 (1925). See also J. L. Synge, Trans. Roy. Soc. Canada **34**, 1 (1940) and Proc. Roy. Soc. London A177, 118 (1940) as well as the discussion of Synge's treatment in III.

[20] See in particular M. H. L. Pryce, Proc. Roy. Soc. London. A168, 398 (1938).

single accelerated particle in a completely absorbing universe. There we have in the neighborhood of the radiating source $F^{(a)}=(1/2)R^{(a)}+(1/2)S^{(a)}$ and $F^{(b)}+F^{(c)}+\cdots=$(sum of advanced fields of absorber particles)[21] $=(1/2)R^{(a)}-(1/2)S^{(a)}$. For the parts of these fields which are proportional to the acceleration of the charge, and which vary at large distance as $1/r$, we have for $R^{(a)}$ and $S^{(a)}$ respectively a zero value except for an instant r/c seconds after or before the moment of acceleration. The corresponding energy flux (Table II) satisfies in both the Frenkel and the canonical formulations the law of conservation of energy, but agrees only in the canonical case with customary ideas of energy localization. From the standpoint of pure electrodynamics it is not possible to choose between the two tensors. The difference is of course significant for the general theory of relativity, where energy has associated with it a gravitational mass. So far we have not attempted to discriminate between the two possibilities by way of this higher standard.

CONCLUSION

We conclude that the theory of direct interparticle action, and the equivalent adjunct field theory, provide a physically reasonable and experimentally satisfactory account of the classical mechanical behavior of a system of point charges in electromagnetic interaction with one another, free of the ambiguities associated with the idea of a particle acting upon itself.

APPENDIX

To compute the integral of the field energy which appears in (26), we express each field as a superposition of elementary fields from each infinitesimal range of path $d\alpha$, and the tensor T_{mn} or $T_{mn}{}^{*}$ as the superposition of parts due to stretches $d\alpha$ of the world line of a and $d\beta$ of b. We use the notation $T^{\dagger}d\alpha d\beta$, $F^{\dagger}d\alpha$ to indicate each such elementary contribution to T, F, etc. Thus the four-potential $\mathfrak{R}^{\dagger}d\alpha$ arises from a charge which appears for an instant at $a(\alpha)$ and disappears at $a(\alpha+d\alpha)$.

The lack of conservation of the charge which generates the elementary potential causes the four-divergence of $\mathfrak{R}^{\dagger}$ to equal a non-zero scalar, r,

$$\partial\mathfrak{R}_{\mu}{}^{(a)\dagger}/\partial x_{\mu}=(\partial/\partial x_{\mu})2e_a\dot{a}_{\mu}\begin{cases}\delta(ax_{\nu}ax^{\nu}) & \text{for } x^4>a^4\\ 0 & \text{for } x^4<a^4\end{cases}$$

$$=r^{(a)}(x,\alpha)=4e_a\dot{a}_{\mu}xa^{\mu}\begin{cases}\delta'(ax_{\nu}ax^{\nu}) & \text{for } x^4>a^4\\ 0 & \text{for } x^4<a^4\end{cases}, \quad (34)$$

whose integral however satisfies the conservation condition $\int_{-\infty}^{+\infty}r^{(a)}(x,\alpha)d\alpha=0$. This circumstance permits some latitude in the definition of the elementary field in terms of the potential. It will prove useful to adopt the definition

$$R_{mn}{}^{(a)\dagger}=\partial\mathfrak{R}_{n}{}^{(a)}/\partial x^{m}-\partial\mathfrak{R}_{m}{}^{(a)}/\partial x^{n}-(r^{(a)}/2)g_{mn}. \quad (35)$$

The elementary field is not antisymmetrical in the indices m and n, but the normal field $R_{mn}{}^{(a)}=\int R_{mn}{}^{(a)\dagger}d\alpha$ changes sign of course on this interchange of labels.

The elementary component of the stress-energy tensor is not symmetric in its two indices, but its divergence is found by direct algebra to have the simple value

$$(\partial/\partial x_{\rho})(R^{(a)\dagger}\ \&\ S^{(b)\dagger})_{m\rho}=(1/8\pi)(-\partial^2R_{\rho}{}^{(a)\dagger}/\partial x^{\sigma}\partial x_{\sigma})\times(S_{m}{}^{(b)\dagger\rho}-(S^{(b)}/2)\delta_{m}{}^{\rho})+(1/8\pi)(-\partial^2S_{\rho}{}^{(b)\dagger}/\partial x_{\sigma}\partial x^{\sigma})\times(R_{m}{}^{(a)\dagger\rho}-(r^{(a)}/2)\delta_{m}{}^{\rho}), \quad (36)$$

where the typical field d'Alembertian has the value

$$-\partial^2R_{m}{}^{(a)\dagger}/\partial x^{\sigma}\partial x_{\sigma}=4\pi e_a\delta(x^1-a^1)\delta(x^2-a^2)\delta(x^3-a^3)\delta(x^4-a^4)\dot{a}_{m}. \quad (37)$$

We integrate (36) over a four-dimensional region of the form shown in Fig. 4. Of the terms on the right the second vanishes throughout this region, and the first gives

$$(e_a/2)(S_{m\rho}{}^{(b)\dagger}(a)-(g_{m\rho}/2)s^{(b)}(a))\dot{a}^{\rho}=(e_a/2)(\partial\mathfrak{S}_{\rho}{}^{(b)\dagger}/\partial a^{m}-\partial\mathfrak{S}_{m}{}^{(b)\dagger}/\partial a^{\rho}-g_{m\rho}\partial\mathfrak{S}_{\nu}{}^{(b)\dagger}/\partial a_{\nu})\dot{a}^{\rho}=2e_ae_b(ab_m\dot{a}^{\mu}\dot{b}_{\mu}-\dot{b}_m{}^{*}ab_{\mu}\dot{a}^{\mu}-\dot{a}_mab_{\mu}\dot{b}^{\mu})\delta'(ab_{\nu}ab^{\nu}) \quad (38)$$

when $b^4>a^4$, and zero otherwise. The four-integral on the left hand side may be expressed via the theorem of the Gauss in the form

$$-\int\int\int(R^{(a)\dagger}\ \&\ S^{(b)\dagger})_{m\rho}d\sigma^{\rho}. \quad (39)$$

Here the integral, which goes over the whole of the three-dimensional region or "surface" in the figure, contributes only over the upper region because of the vanishing elsewhere of at least one of the fields in question. The elementary contributions just computed we now sum over the world line of a from $-\infty$ to α and over the world line of b from β to ∞, where α and β determine the points where the world lines of a and b intersect the space-like surface σ. We have then only to erase the daggers in (39). The converse expression, with $R^{(b)}$ & $S^{(a)}$, we obtain by interchanging the roles of b and a in (38) and in the limits of integration. In this way follows at once the identity of expression (20) for the energy in the theory of direct interparticle interaction and the canonical expression (26–28) of the adjunct field theory.

When instead the Frenkel expression (32) is used for the stress-energy tensor, then there results an increment in the energy-comomentum four-vector given by the expression

$$G_{m}{}^{*}(\alpha,\beta)-G_{m}(\alpha,\beta)=e_ae_b\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(ab_m\dot{a}^{\mu}\dot{b}_{\mu}-\dot{b}_mab_{\mu}\dot{a}^{\mu}-\dot{a}_mab_{\mu}\dot{b}^{\mu})\begin{cases}\delta'(ab_{\nu}ab^{\nu}) & \text{for } a^4<b^4\\ -\delta'(ab_{\nu}ab^{\nu}) & \text{for } a^4>b^4\end{cases}d\alpha d\beta$$

$$=e_ae_b\int\int\begin{cases}+\delta' & \text{for } a^4<b^4\\ -\delta' & \text{for } b^4>a^4\end{cases}ab_{m}da_{\mu}db^{\mu}, \quad (40)$$

a covariant which is independent of α and β and which has an interesting relation to the two world lines in question.

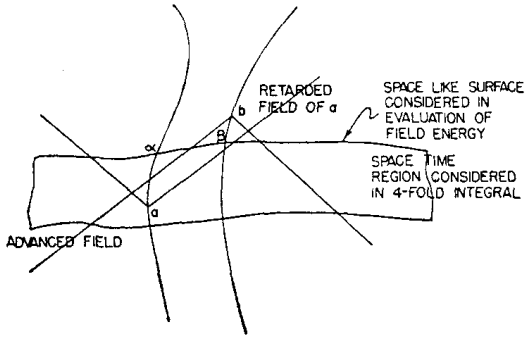


FIG. 4. Contribution to canonical expression for field energy which arises from coupling of retarded field of a and advanced field of b.

[21] See part III for fuller discussion.

PHYSICAL REVIEW VOLUME 74, NUMBER 8 OCTOBER 15, 1948

A Relativistic Cut-Off for Classical Electrodynamics

R. P. FEYNMAN
Cornell University, Ithaca, New York

(Received June 8, 1948)

Ordinarily it is assumed that interaction between charges occurs along light cones, that is, only where the four-dimensional interval $s^2=t^2-r^2$ is exactly zero. We discuss the modifications produced if, as in the theory of F. Bopp, substantial interaction is assumed to occur over a narrow range of s^2 around zero. This has no practical effect on the interaction of charges which are distant from one another by several electron radii. The action of a charge on itself is finite and behaves as electromagnetic mass for accelerations which are not excessive. There also results a classical representation of the phenomena of pair production in sufficiently strong fields.

QUANTUM electrodynamics is built from a classical counterpart that already contains many difficulties which remain upon quantization. It has been hoped that if a classical electrodynamics could be devised which would not contain the difficulty of infinite self-energy, and this theory could be quantized, then the problem of a self-consistent quantum electrodynamics would be solved. For this reason many successful attempts have been made to produce such a classical theory. The field equations can be made non-linear,[1] the fields produced by or acting on an electron can be redefined,[2,3] or one may resort to some averaging of the fields over space or time.[4] These theories have, however, met with considerable difficulties when an attempt has been made to quantize them. In this paper a consistent classical theory is described which the author believes can be quantized. Some preliminary results of the quantization of this theory will be discussed in a future paper. Some of the physical ideas of the classical form of the theory are sufficiently interesting in themselves to warrant their discussion first in a separate paper.

The potential at a point in space at a given time depends on the charge at a distance r from the point at a time previous by $t=r$ (taking the speed of light as unity). Speaking relativistically, interaction occurs between events whose four-dimensional interval, s, defined by $s^2=t^2-r^2$, vanishes. There results, however, an infinite action of a point electron on itself. The present theory modifies this idea by assuming that substantial interaction exists as long as the interval s is time-like and less than some small length, a, of order of the electron radius. When t is large since $\Delta(s^2)=2t\cdot\Delta t$ this means a spread in the time of arrival of a signal of amount of order $a^2/2t$. For charges separated by many electron radii there is, therefore, essentially no effect of the modification. For the action of an electron on itself, however, there is a considerable modification. The result is to reduce the infinite self-energy to a finite value. For accelerations which are not extreme, the action of an electron on itself appears simply as an electromagnetic mass. If desired in the classical theory, all the mass of an electron may be represented as electromagnetic. (In the quantum theory this cannot be done in a reasonable way as the electromagnetic mass comes out quite small under reasonable assumptions for a.) We have, therefore, a consistent classical theory which does not disagree with classical experience.

In the remainder of the paper we formulate this idea mathematically, and draw one or two simple consequences. We then discuss a curious

[1] M. Born and L. Infeld, Proc. Roy. Soc. London **A144**, 425 (1935).
[2] P. A. M. Dirac, Proc. Roy. Soc. London **A167**, 148 (1938). An excellent discussion of these matters is given by C. J. Eliezer, Rev. Mod. Phys. **19**, 147 (1947).
[3] J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **17**, 157 (1945).
[4] There are many theories of this nature. The author's theory is essentially that of F. Bopp, Ann d. Physik **42**, 573 (1942). R. Peierls and H. McManus have developed a theory in which the electron is pictured as a rigid distribution of charge in both space and time. The theory can be shown to be exactly equivalent to the present one, at least for a class of f functions. Their physical ideas may offer advantages over the present one in which the function f is not so directly interpretable. I thank Dr. McManus for a copy of his thesis. For a summary of another theory of this type see B. Podolsky and P. Schwed, Rev. Mod. Phys. **20**, 40 (1948). A somewhat different type is that of N. Rosen, Phys. Rev. **72**, 298 (1947).

feature of this theory. It can give a classical representation of the phenomena of pair production in sufficiently strong fields. This is of interest because the physical ideas may possibly be carried over to give a clearer understanding of the hole theory of positrons.

The main result which is to be carried over to quantum problems is this: In any process in which there is no permanent emission of quanta one must assume the field quanta to have a "density" $g(k_4,\mathbf{K})$ in frequency, and wave number space. This replaces the usual assumption that the frequency k_4 equals the magnitude of the wave number, K, and that the density in wave numbers $\mathbf{K}$, is uniform (corresponding to $g(k_4,\mathbf{K})=\delta(k_4^2-K^2)$). The properties $g(k_4,\mathbf{K})$ ought to have are discussed more fully below.

MATHEMATICAL FORMULATION

It is most convenient (but not necessary) to formulate these ideas in the language of action at a distance.[3] Hence a brief summary of that point of view is given here. We start with Fokker's action principle that the action

$$S=\sum_a m_a\int(da_\mu da_\mu)^{\frac{1}{2}}+\sum_{a,b}{}'e_ae_b\int\int\delta(s_{ab}^2)da_\mu db_\mu \quad (1)$$

is an extremum. Here a_μ represents, for $\mu=1$ to 4, the three space coordinates and the time coordinate of a particle a of mass m_a, charge e_a. We shall later consider them as functions of a parameter α, say. The b_μ are corresponding quantities for a particle b, etc. The symbol $x_\mu y_\mu$ means $x_4y_4-x_1y_1-x_2y_2-x_3y_3$ and $s_{ab}^2=(a_\mu-b_\mu)(a_\mu-b_\mu)$. The δ is Dirac's delta function.

The integrals are taken over the trajectories of the particles. The $\sum'$ means the sum over all pairs a, b with $a\neq b$. We consider varying the path $a_\mu(\alpha)$ of particle a. Defining

$$A_\mu^{(b)}(x)=e_b\int\delta(s_{xb}^2)db_\mu, \quad (2)$$

where x stands for x_μ, a point in space time, we can write (1) as

$$S=\sum_a m_a\int(da_\mu da_\mu)^{\frac{1}{2}}+\sum_a\sum_{b\neq a}e_a\int A_\mu^{(b)}(a)da_\mu.$$

The result of seeking an extremum of this is to lead in the well-known way to the equations of motion,

$$m_a\frac{d}{d\tau_a}\left(\frac{da_\nu}{d\tau_a}\right)=e_a\frac{da_\mu}{d\tau_a}\sum_{b\neq a}F_{\mu\nu}^{(b)}(a), \quad (3)$$

where we can call $F_{\mu\nu}^{(b)}(x)$ the field at x caused by particle b. It is given by

$$F_{\mu\nu}^{(b)}(x)=\partial A_\mu^{(b)}(x)/\partial x_\nu-\partial A_\nu^{(b)}(x)/\partial x_\mu.$$

We have written $d\tau_a=(da_\mu da_\mu)^{\frac{1}{2}}$ for the proper time along the path of a.

Since $\Box^2\delta(s_{xb}^2)=4\pi\delta(x_1-b_1)\delta(x_2-b_2)\delta(x_3-b_3)\times\delta(x_4-b_4)$, where $\Box^2=(\partial/\partial x_\mu)(\partial/\partial x_\mu)$, Eq. (2) gives

$$\Box^2A_\mu^{(b)}(x)=4\pi e_b\int\delta(x_1-b_1)\delta(x_2-b_2)\times\delta(x_3-b_3)\delta(x_4-b_4)db_\mu, \quad (4)$$

which is 4π times the current four-vector of a point charge e_b. Thus $F_{\mu\nu}^{(b)}(x)$ satisfies Maxwell's equations. But the special solution (2) is not the usual retarded solution but is rather half the retarded plus half the advanced solution of Lienard and Wiechert[5] (since $\delta(t^2-r^2)=(1/2r)\times(\delta(t+r)+\delta(t-r).)$ Thus we may write (dots representing derivatives with respect to τ_a, and the fields being calculated at the point $x_\mu=a_\mu$),

$$m_a\ddot{a}_\nu=e_a\dot{a}_\mu\sum_{b\neq a}(\tfrac{1}{2}F^{(b)}{}_{\mu\nu\ \mathrm{ret}}+\tfrac{1}{2}F^{(b)}{}_{\mu\nu\ \mathrm{adv}}). \quad (5)$$

This can be compared to the usual theory which just uses retarded effects by writing it in the form

$$m_a\ddot{a}_\nu=e_a\dot{a}_\mu\left\{\sum_{b\neq a}F^{(b)}{}_{\mu\nu\ \mathrm{ret}}+\tfrac{1}{2}\sum_{\text{all }b}[F^{(b)}{}_{\mu\nu\ \mathrm{adv}}-F^{(b)}{}_{\mu\nu\ \mathrm{ret}}]-\tfrac{1}{2}[F^{(a)}{}_{\mu\nu\ \mathrm{adv}}-F^{(a)}{}_{\mu\nu\ \mathrm{ret}}]\right\}, \quad (6)$$

as in the paper[3] by Wheeler and Feynman. As in that paper the first term is the retarded field of other charges, the second term vanishes in a world where all emitted light is eventually ab-

[5] This use of advanced and retarded potentials is really unnecessary for an understanding of the modifications of electrodynamics which is the main point of the paper. It results from the aurhor's desire to start with a principle of least action, for it is in this form that the transition to quantum theory can be made.

sorbed,[6] and the third term, depending only on the motion of a, is the force of radiative damping. Thus (1) is equivalent to (6) and thus satisfactorily describes the known laws of classical electrodynamics. There is no self-energy.

According to the above, a particle does not act upon itself, as the term with $a=b$ in the sum $\sum'_{a,b} e_a e_b \int \cdots$ in the action has been omitted. (Radiation resistance is pictured as in indirect effect of source on absorber and absorber on source.) The field of each particle must be kept separate in order to exclude, when asking for the force on a particle, the field of the particle itself.

There is no *need* to do so, but it is an interesting question to try to reinstate the idea of a universal field. This requires that a particle be allowed to act on itself and the term $a=b$ included in the action sum. This leads immediately to an infinite self force. This difficulty can be eliminated if the $\delta(s_{ab}^2)$ is replaced, as Bopp[4] has suggested, by some other function $f(s_{ab}^2)$ of the invariant s_{ab}^2, which behaves like $\delta(s_{ab}^2)$ for large dimensions but differs for small. (We shall discuss the properties of this function later, but as an example to keep in mind, consider $f(s^2) = (1/2a^2)\exp(-|s|/a)$ for $s^2>0$, and $f(s^2)=0$ for $s^2<0$ with a of order of the electron radius e^2/mc^2.)

We study the consequences of replacing (1) by the law that S is extremum if

$$S=\sum_a m_a \int (da_\mu da_\mu)^{\frac{1}{2}} + \tfrac{1}{2}\sum_a \sum_b e_a e_b \int\int f(s_{ab}^2) da_\mu db_\mu. \quad (7)$$

The term with $a=b$ may be written

$$\tfrac{1}{2} e_a^2 \int\int f(s^2_{aa'}) da_\mu da_\mu', \quad (8)$$

where a and a' are two points on the world-line of a. The variation problem clearly leads to

$$m_a \ddot{a}_\nu = e_a \dot{a}_\mu \left[\sum_{b\neq a} \bar{F}_{\mu\nu}{}^{(b)}(a) + \bar{F}_{\mu\nu}{}^{(a)}(a) \right], \quad (9)$$

where

$$\bar{F}_{\mu\nu}{}^{(b)}(x) = \partial \bar{A}_\mu{}^{(b)}(x)/\partial x_\nu - \partial \bar{A}_\nu{}^{(b)}(x)/\partial x_\mu \quad (10)$$

and the bar over the field quantities indicate that they are calculated from the f function rather than the δ function. That is,

$$\bar{A}_\mu{}^{(b)}(x) = e_b \int f(s_{ab}^2) db_\mu. \quad (11)$$

This theory differs from the usual in two respects: A. There is an extra force $h_\nu = e_a \dot{a}_\mu \bar{F}_{\mu\nu}{}^{(a)}(a)$ on particle a depending only on the motion of a. This we shall study in a moment and show that it represents inertia. B. The fields of other particles are given by the curl of a potential but the potential (11) no longer solves the Maxwell equations (4). However, since $f(s^2)$ is close to $\delta(s^2)$ this means that except for particles very close together nothing is changed very much. Thus $f(t^2-r^2)$ is large only when $t=r$ is nearly satisfied, but for large t near $+r$, say, $f(t^2-r^2) \cong f(2t(t-r))$ so that the function which has width a^2 in its argument s^2 has width $a^2/2t$ in $t-r$. Thus for increasing distances from the source the potentials satisfy Maxwell's equations ever more accurately.

The analog of Eq. (6) becomes

$$m_a \ddot{a}_\nu = e_a \dot{a}_\mu \Big\{ \sum_{b\neq a} \langle F\rangle^{(b)}{}_{\mu\nu\ \mathrm{ret}} + \tfrac{1}{2} \sum_{\text{all } b} [F^{(b)}{}_{\mu\nu\ \mathrm{adv}} - F^{(b)}{}_{\mu\nu\ \mathrm{ret}}] - \tfrac{1}{2}[F^{(a)}{}_{\mu\nu\ \mathrm{adv}} - F^{(a)}{}_{\mu\nu\ \mathrm{ret}}] + \bar{F}^{(a)}{}_{\mu\nu} \Big\}, \quad (12)$$

where we define $\langle F\rangle_{\mathrm{ret}} = \bar{F} + \tfrac{1}{2}F_{\mathrm{ret}} - \tfrac{1}{2}F_{\mathrm{adv}}$. Thus only the δ part, so to speak, of the fields becomes retarded. It would not do to replace F by $\bar{F}$ throughout in (6) for then we could not deduce that the second term is zero at the source because it was zero at infinity for it would not then be a solution of Maxwell's equation in empty space. The damping term is unaltered. It plus the self-force can be written $\langle F\rangle^{(a)}{}_{\mu\nu\ \mathrm{ret}}$ (see

[6] That the second term vanishes in these circumstances may be seen as follows. If a source radiates for a time, at a very long time afterwards the total retarded field vanishes, for all the light is absorbed. But also the total advanced field vanishes at this time (for charges are no longer accelerating and the advanced field exists only at times previous to their motion). Hence, the difference vanishes everywhere at this time and, since it is a solution of Maxwell's homogeneous equations, at all times.

footnote 5), so in practice one can write simply

$$m_a\ddot{a}_\nu = e_a\dot{a}_\mu \sum_{\text{all } b} \langle F\rangle^{(b)}_{\mu\nu \text{ ret.}}$$

The effect of the modification in the theory using retarded fields is therefore to change, slightly, the field of one particle on another when they are very close, and to add a self-force h_μ.

We now turn to a study of the self-force h_μ. This can be calculated directly from the formulae (10), (11) but a simpler way is from the action term (8). This term in the action can be re-expressed approximately if we assume that the accelerations are not too great. Only values of a' near a are important. Let us define a parameter along the path and say a corresponds to the value α of this parameter, a' to the value $\alpha'=\alpha+\epsilon$. Assuming a' not to vary too rapidly with ϵ we can approximate $s_{aa'}{}^2=(a_\mu-a_\mu')\times(a_\mu-a_\mu')$ by $\epsilon^2(da_\mu/d\alpha)(da_\mu/d\alpha)=\epsilon^2(d\tau_a/d\alpha)^2$. Likewise $da_\mu da_\mu'$ is to sufficient accuracy $(da_\mu/d\alpha)(da_\mu/d\alpha)d\alpha d\epsilon$. Thus the self-action term is approximately $\frac{1}{2}e_a{}^2\int\int f(\epsilon^2(d\tau_a/d\alpha)^2)\cdot(d\tau_a/d\alpha)^2 \times d\epsilon d\alpha$. Then calling $\eta=\epsilon(d\tau_a/d\alpha)$ we can write this as

$$\mu_a\int(d\tau_a/d\alpha)d\alpha=\mu_a\int(da_\mu da_\mu)^{\frac{1}{2}}, \qquad (13)$$

where we have set

$$\mu_a=\tfrac{1}{2}e_a{}^2\int_{-\infty}^{\infty}f(\eta^2)d\eta. \qquad (14)$$

That is, the self-action term to this approximation represents pure electrodynamic mass. The term readily combines with $m_a\int d\tau_a$ for the mass is correctly invariant. We can go further and assume that originally m_a is zero and all mass of electrons is electrodynamic, but for protons this would then not be so.

The function $f(s^2)$ is to be normalized such that

$$\int_{-\infty}^{\infty}f(s^2)d(s^2)=1. \qquad (15)$$

The condition (14) says the range in η of $f(\eta^2)$ is of order e^2/μ, or if μ is the electron mass, of order of the electron radius. The function $f(s^2)$ is chosen so that it is symmetrical near past and future light cones since any asymmetry drops out in the form (7) of the action. Other than these conditions, there are strictly no further conditions on $f(s^2)$. It is convenient to assume $f(s^2)$ to be zero if s^2 is negative (space-like). It is also very desirable to have $f(s^2)$ fall rapidly away from the light cone, rather than oscillate indefinitely, and to have $f(s^2)$ finite everywhere.

By taking the Fourier transform of (11), one can represent the field as a superposition of the effects of harmonic oscillators in the usual way. However, the oscillators corresponding to waves of wave number k_1, k_2, k_3 need not have a frequency k_4 equal to the magnitude of the wave number. Instead we can take the density of the oscillators to be k_4 times, $g(k_\mu k_\mu)dk_1dk_2dk_3dk_4$ where g is defined for positive k_4 only, and is

$$g(k_\mu k_\mu)=(1/4\pi^2)\int f(s^2{}_{xy})\cos(k_\mu(x_\mu-y_\mu)) \times dx_1dx_2dx_3dx_4.$$

It is a function of the invariant $k_\mu k_\mu$ only. The ordinary case, $f(s^2)=\delta(s^2)$ corresponds to $g(k_\mu k_\mu)=\delta(k_\mu k_\mu)$. The condition that $f(s^2)$ be finite on the light cone implies that $g(k_\mu k_\mu)$ can be written in the form

$$g(k_\mu k_\mu)=\int_0^\infty[\delta(k_\mu k_\mu)-\delta(k_\mu k_\mu-\lambda^2)]G(\lambda)d\lambda. \qquad (16)$$

Here $G(\lambda)$ is normalized such that $\int_0^\infty G(\lambda)d\lambda=1$, in view of (15). It is otherwise arbitrary, as $f(s^2)$ is. The λ values for which g must exist must be large, going up to order μ/e^2.

If G is chosen as $\delta(\lambda-\lambda_0)$ the resulting $f(s^2)$ is (for $s^2\geqslant 0$) the Bessel function, $\lambda_0J_1(\lambda_0 s)/s$. For large t, if $s=(t^2-r^2)^{\frac{1}{2}}$, this does not die off fast with $t-r$, but oscillates with phase varying as $\lambda_0(t^2-r^2)^{\frac{1}{2}}$. That is, it oscillates with frequency $\lambda_0(1-r^2/t^2)^{-\frac{1}{2}}$ at a time corresponding to arrival of signals with velocity r/t and thus in quantum mechanics would represent arrival of radiated "particles" of mass $\hbar\lambda_0$. The free emission of such "particles" is removed in classical theory by interference among the various values of λ if a smooth distribution, $G(\lambda)$, of λ is used. This is required if f is to represent say a function decaying rather than oscillating (see appendix).

It appears that the quantum mechanical result is simply this: For processes without permanent radiation the oscillator density g is to replace $\delta(k_\mu k_\mu)$. The negative sign in (16) proves embarrassing (see appendix) if quanta of mass λ_0 can be freely radiated so a wide distribution

in λ corresponding to a monatonic $f(s^2)$ is preferable. As an example, for $f(s^2)=(1/2a^2)\times\exp(-|s|/a)$ find $G(\lambda)=(3a^2\lambda)(1+a^2\lambda^2)^{-5/2}$.

The electrostatic potential at a distance r from a stationary charge, is according to (11),

$$\bar{A}_4(r)=e\int_{-\infty}^{\infty}f(t^2-r^2)dt. \tag{17}$$

For large r, in view of (15) this is readily seen to be e/r. At the origin $r=0$ however, it is finite being $e\int_{-\infty}^{\infty}f(t^2)dt$ or $2\mu/e$. This has a simple interpretation if all mass is electromagnetic. The energy released in bringing a positron and electron charge together and so canceling out all external fields is just 2μ, the rest mass these particles have in virtue of their fields. Or put otherwise, the rest mass particles have is simply the work done in separating them against their mutual attraction after they are created. No energy is needed to create a pair of particles at the same place. (These ideas do not have direct quantum counterparts since in quantum theory all mass does not appear to be electromagnetic self energy, at least in the same simple way.) There may be a maximum field of attraction between two like charges at some separation since, for some functions f the force arising from (17) vanishes at the origin, and of course again at infinity.

There remains to discuss a curious point about the solutions of the least action principle (7) with the mechanical mass term m_a absent. First let us study the simple problem of an ordinary single particle of mass μ in a potential A_4 (caused by other charges) which depends only on one coordinate x. Call the time t, and use this for the parameter α. The action is

$$S=\mu\int(1-\dot{x}^2)^{\frac{1}{2}}dt+e\int A_4dt. \tag{18}$$

Now suppose the potential A_4 is zero outside a small band in x say $|x|<b/2$ (potential barrier) and that it is large positive, and constant within the region. Consider, in Fig. 1, the paths from a point 1 to a point 2 which make S a local maximum. A typical solution is the solid line which is kinked out of the straight line so as to increase the time integral of A_4. This represents a particle moving from 1 rapidly toward the barrier,

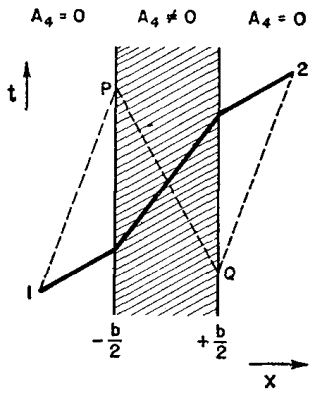


FIG. 1. If two points 1, 2 are separated by a high potential barrier, there are two paths which make action an extremum. One (solid line) represents passage of a fast electron. The other (dotted line) has a section reversed in time and is interpreted as the effective penetration of the barrier by a slow electron by means of a pair production at Q and annihilation at P, section PQ representing the motion of the positron.

entering the region of high potential, losing energy and thus going slower in this region. The high velocity is regained on passing out of the region to 2. Slow particles cannot penetrate the barrier.

But there may be another local maximum. Consider the path $1PQ2$. In the interval PQ the proper time integral must be taken positively as can be verified from a study of the derivation of (13). Now moving the point P upward by Δt might be expected to increase the action by over $2\mu\Delta t$ because the length of $P1$ and PQ are both increased. On the other hand, the integral $\int A_4dt$ is now negative and if A_4 exceeds 2μ such a curve may be a local maximum. Thus for A_4 greater than 2μ there is a new way that slow particles can penetrate the barrier. This is a classical analog of the Klein Paradox.

How would such a path appear to someone whose future gradually becomes past through a moving present? He would first see a single particle at 1, then at Q two new particles would suddenly appear, one moving into the potential to the left, the other out to the right. Later at P the one moving to the left combines with the original particle at 1 and they both disappear, leaving the right moving member of the original pair to arrive at 2. We therefore have a classical description of pair production and annihilation. The particle whose trajectory has its proper time opposed in sign to the true time t (section

PQ) would behave as a particle of opposite sign, for changing the sign of db_μ in (7) is equivalent to changing the sign of e_b. This idea that positrons might be electrons with the proper time reversed was suggested to me by Professor J. A. Wheeler in 1941.

The field at $x=\pm b/2$ is infinite. If it is finite the action (18) does not show such a local maximum, the sharp corner at P becoming a cusp which can go indefinitely into the future. On the other hand, if the correct self-force from (7) is used instead of the approximation (13), a path reversal again becomes a possibility. It is only necessary that the field exceed a critical value, namely, that maximum value of attraction of two unlike particles mentioned above. This field represents a potential of 2μ in a distance of the order of an electron radius and must be as great as this to get the pair of newly created particles apart over the potential barrier of their mutual attraction. (The actual field required to produce pairs in quantum mechanics is 137 times weaker. One might ascribe this to a quantum mechanical penetration of the potential barrier over a Compton wave-length.)

There are many interesting problems presented by these ideas. For example, will pairs be produced *ad infinitum* by the field, or only to that extent that we can guarantee that the positrons will be annihilated by electrons in the future? Again, in a weak field can a large number of pairs be created which separate slightly in the field (which is insufficient to tear the two apart) and thus produce a large polarization of that field? It is hoped that an application of these ideas to a study of positron hole theory will appear in a future paper.

I should like to thank Professor J. A. Wheeler for inoculating me with many ideas without which this work would not have been done.

APPENDIX

The difficulties in a theory such as the one presented here have been discussed by many authors. A very brief discussion of them will be given in this appendix.

The first point is that the action S defined in (1) or (7) is infinite and meaningless because of the infinite extent of the integrals along the path of the particles. The principle of extreme action which we mean to apply can be more rigorously defined as follows. Consider any given variation in paths δa_μ such that $\delta a_\mu \to 0$ as $\alpha \to \pm\infty$. Then define δS as the limit of the quantity δS calculated from (7) with this path variation, the limit being taken as the range of integration passes to infinity. Then the law of motion shall be $\delta S=0$ for all variations which satisfy $\delta a_\mu \to 0$ as $\alpha \to \pm\infty$. The equations of motion (9) are then consequences of the action principle, of course, but not all solutions of these equations satisfy the principle of least action as defined here. There are certain runaway solutions of the equations of motion, such as those discussed by Dirac[2] in the case of Eq. (6), in which the kinetic energy and momentum of a particle increase exponentially with time. These are excluded for they do not satisfy the principle of least action.

Bopp[4] has studied in great detail the consequences of equations of motion. However, he assumes that the function f acts only at retarded times. He finds that the radiation resistance of an oscillator decreases below the normal value with increasing frequency of the oscillator. However, if an oscillator is enclosed in a large, light tight box the fields at the walls of the box are effectively unchanged by the use of f rather than δ. (We are assuming that f decays and does not oscillate. Below, we discuss the situation if f oscillates.) Hence the energy absorbed by the walls will not, apparently, decrease with the increase in oscillator frequency and the radiation resistance will not keep up with it. In the modification described in this paper, in such a box, only the δ-part of f is to be retarded. The radiation resistance has its normal value at all frequencies, and the energy lost will all be found eventually in the walls of the box.

The conservation of quantities like energy (and momentum) can be demonstrated directly if a theory starts from a principle of least action, the form of which action is invariant under a change of origin of the time (and space). For the action (7) consider the quantity

$$g_\nu = \sum_a \left\{ m_a \dot{a}_\nu + e_a \sum_b \bar{A}_\nu^{(b)}(a) \right\}_{\text{at } a_\mu(\alpha_0)} - \sum_a \sum_b e_a e_b \int_{-\infty}^{\alpha_0} \int_{\beta_0}^{\infty} 2(a_\nu - b_\nu) f'(s_{ab}^2) da_\mu db_\mu \quad (19)$$

where $f'(x)=df(x)/dx$. The points $\alpha_0, \beta_0, \ldots$ are an arbitrary set of points, one chosen on the world-line of each particle. In virtue of the equations of motion (9) the derivative of g_ν with respect to any of the α_0's is zero. This is a generalization of the usual conservation of energy. Ordinarily, we would choose all the α_0 such that all $a_4(\alpha_0)$ are equal, and find g_ν is independent of this common value of a_4. The energy is seen to consist of a self-energy, an energy caused by the presence of the potential on each particle, and (since this last would count each interaction twice) a further correction to interaction energy. This is described as a line integral over the paths of the particles, but since one point is in the future and the other in the past, the actual range of integration does not extend beyond the time during which b could interact with a at α_0 and that a could interact with b at β_0. This is the way in which energy which is usually spoken of as being in the field is represented in a theory of action at a distance between particles. Since it is an integral only over a limited range, the energy of motion of the particles is conserved in the long run. (It is easy to generalize (19) to the case that paths may reverse themselves in time.)

We now consider the situation in which the function f oscillates rather than decays. If, as was discussed by Bopp[4] and others (e.g., B. Podolsky and P. Schwed[4]), f is replaced by a Bessel function $\lambda_0 J_1(\lambda_0 s)/s$, the theory corresponds as we have seen to that of interaction through ordinary "quanta" *minus* those corresponding to a mass $\hbar\lambda_0$. The f function does not appear as a pure δ-function at large distances, but another component appears if the frequency of the source exceeds λ_0. Thus, a source at high frequency ω emits waves of two wave-lengths, light of wave number $K=\omega$ and "λ_0-quanta" of wave number $K=(\omega^2-\lambda_0^2)^{\frac{1}{2}}$. Again Bopp's equations (using retarded potentials only) show that the radiative resistance force on a dipole oscillator of amplitude x, frequency ω, is constant at $2e^2\omega^3\dot{x}^2/3c^3$ for $\omega<\lambda_0$ and falls off as ω exceeds λ_0, as $(2e^2x^2/3c^3)[\omega^3-(\omega^2+\frac{1}{2}\lambda_0^2)(\omega^2-\lambda_0^2)^{\frac{1}{2}}]$, remaining positive, however, for all frequencies. The decrease at higher frequencies must correspond to a negative contribution to radiation resistance accompanying the emission of "λ_0-quanta." That is, the λ_0-quanta behave as though they had negative energy. That this is so and that it results in fundamental difficulties may be seen from a few examples. If through interference the rate of emission of "λ_0-quanta" can be enhanced relative to the rate for the ordinary light quanta, a net negative radiation resistance will result. For example, a set of two vertical dipoles oscillating in phase (at frequency $\omega=2\lambda_0/3^{\frac{1}{2}}$), separated horizontally by one-half the wave-length of light, and one-fourth the wave-length of the λ_0-quanta of the same frequency, shows a negative net radiation resistance. It would oscillate with ever-increasing amplitude, the large emission of negative energy λ_0-quanta supplying the increase in energy of the oscillators and the energy of the light quanta emitted. Again a beam of λ_0-quanta passing through a medium containing damped (energy-absorbing) oscillators increases in amplitude. The wave of λ_0-quanta scattered by the oscillators in the forward direction which ordinarily interferes destructively with the incident wave, in this case has a reversed sign and enhances the incident wave. (The light scattered forward has the incorrect wave-length to make an appreciable effect by interference.) A beam of λ_0-quanta can be separated from light of the same frequency by having the radiation from a source of a given frequency impinge on a diffraction grating of scattering centers. The λ_0-quanta and light quanta will then be scattered in different directions as they have different wave-lengths.

What results if instead of using only retarded waves for λ_0-quanta, we start from the least action principle and analyze the situation of a source enclosed in a box? Then the derivation of Eq. (12) is still incomplete as $\langle F\rangle_{\text{ret}}$ still contains both advanced and retarded components (of λ_0-quanta) at large distances. We could now split off the advanced parts for λ_0-quanta just as we did for light. The resulting equation is just that used by Bopp, namely all retarded interactions but negative contribution of λ_0-quanta to radiation resistance, and therefore leads this time to conservation of energy but to diverging solutions. Such diverging solutions are, as indicated above, excluded by the least action principle so this form of the equation is not convenient. Non-divergent motions do exist. To see this it is better to split off the retarded part of

the λ_0-quanta. What results is that light goes by retarded waves, λ_0-quanta by advanced waves,[7] and the radiation resistance of both contribute positively. Thus an accelerating charge will emit light, but it is predestined that negative energy λ_0-quanta were coming toward it to be absorbed, still further increasing the radiation resistance. This avoids the divergent solutions only to predict observable advanced effects.

For these reasons it is better to restrict oneself to the case of a decaying f-function (distribution of λ) for which a consistent theory can be made. Then the modifications of classical electrodynamics will only appear at very small distances from a charge. On the other hand, these distances are well within the Compton wave-length so that modifications caused by quantum mechanics would in any case appear before the ones here discussed. There is, therefore, little reason to believe that the ideas used here to solve the divergences of classical electrodynamics will prove fruitful for quantum electrodynamics. Nevertheless, the corresponding modifications were attempted with quantum electrodynamics and appear to solve some of the divergence difficulties of that theory. This will be discussed in a future paper.

[7] This may be understood in that, as indicated above, the energy-absorbing walls of the box absorb retarded light waves, but cannot be presumed to absorb retarded λ_0-quanta. Instead, in fact, they spontaneously emit such waves (warming up in the process) and non-divergent solutions result only if they emit just exactly the λ_0-quanta which can later be absorbed by the accelerating charge at the center.

II.C *Quantum Electrodynamics*

Because Feynman's unconventional approach to QED did not use quantum field theory, he found it difficult to convince his peers of its validity. When he outlined his views to them at a small elite gathering, the Pocono Conference of March 30–April 1, 1948, as he later recalled, his lecture was "a hopeless presentation." He was treated as an alien: "I had too much stuff. My machines came from too far away."[1] He thus thought that it would be best first to describe his important physical results, and to show how they could be obtained by more standard methods. In this way he hoped to motivate his colleagues to make the effort to follow his more powerful new methods, which he would present later.

In this section we group the papers containing physical results in QED, some already published by others, but some obtained correctly for the first time. The papers that detail Feynman's new mathematical tools and derivations will be given in Part III below.

Feynman began this acclimatization process by publishing paper [9], which includes a quantum-mechanical version of the relativistic high-frequency cutoff that he had introduced for classical electrodynamics in paper [8]. This method, an example of what would later be called "regularization," made the divergent integrals of QED finite — except for the so-called vacuum polarization integrals, which would require a stronger form of regularization.[2] Using the "old-fashioned" perturbation theory of Dirac, Feynman showed that his cutoff procedure led to the same results for the electromagnetic shift of energy levels of hydrogen (Lamb shift) and the radiative corrections to potential scattering as had been published earlier by others, including Bethe, Schwinger, and Weisskopf. However, Feynman's results avoided the subtraction of infinite integrals, using only finite ones that he showed were very insensitive to the value chosen for the cutoff.

Paper [12] treats the "motion of electrons and positrons in given external potentials." It introduces the overall space–time view, emphasizing the solutions of the Dirac relativistic electron equation in the integral form, the Green's function or "propagator," which is the probability amplitude for the electron to pass from one point in space–time to another. The allowed paths can propagate forward or backward in time, the time-backward paths being interpreted as positrons.[3] Thus here are introduced the famous trademark "Feynman diagrams." In an appendix, Feynman derives his formulation from the (misnamed) "second quantization" theory of the Dirac electron–positron field.

Paper [13], submitted a month later than paper [12], appeared in the same issue of the *Physical Review*. A continuation of the first paper, it uses the same notation and diagrams, but now it attacks the general problem of QED, involving, besides the external potentials and photons, also the fields of the charged particles themselves, involving both real and virtual photons. This paper constitutes a textbook of Feynman's powerful methods for solving real problems in QED, and describes his propagators and their formulation in momentum space. He discusses the electron self-energy problem, the convergence of processes with virtual quanta, the radiationless scattering problem, and the Lamb shift. His footnote (numbered 13 — appropriately, as he remarked) on the history of the calculation of the relativistic Lamb

[1]Feynman interview with S.S. Schweber, 1984. See Schweber's *QED*, p. 436.

[2]W. Pauli and F. Villars, "On the invariant regularization in relativistic quantum theory," *Rev. Mod. Phys.* **21**, 434 (1949).

[3]This provocative notion is a major technical advance, but it has engendered as much (mostly meaningless) speculation about its "meaning" among the uninformed as has Einstein's "relativity."

shift apologizes for delays caused by the disagreement of his slightly wrong result of early 1948 with other (correct) calculations. He also calculates the vacuum polarization part of the Lamb shift, using a regularization method that he credits to Pauli and Bethe (the Pauli–Villars paper had not yet appeared). In footnote 18 he writes: "It would be very interesting to calculate the Lamb shift accurately enough to be sure the 20 megacycles expected from vacuum polarization are actually present." This shows that he had not yet fully accepted the need for a quantum field theory, as opposed to delayed action-at-a-distance!

Paper [18] is a calculation of the radiation corrections to the Klein–Nishina formula for Compton scattering. Even when making full use of the Feynman calculational tricks, it was a long and tedious calculation. However, the scattering of a photon by an electron can be regarded as the fundamental interaction described by QED, and unlike the other processes mentioned (Lamb shift, electron anomalous magnetic moment), it is not a single quantity but a complete differential cross-section, and it does not make use of an approximate nonrelativistic potential, like the radiationless scattering and the Lamb shift. Thus it is an especially appropriate testing ground for QED, as well as being an important observable effect at higher electron energies, which would be employed to analyze experiments in the decades that followed.

As stated in the introduction, paper [45] is a review of QED as of 1961, included here for its originality and stimulating language. Feynman concludes it by referring to his "long-held strong prejudice that [QED] *must* fail significantly (other than being simply incomplete) at around 1 GeV virtual energy."[4] Furthermore, he wrote, "I still hold this belief, and do not subscribe to the philosophy of renormalization."

Selected Papers

[9] Relativistic cut-off for quantum electrodynamics. *Phys. Rev.* **74** (1948): 1430–1438.
[12] The theory of positrons. *Phys. Rev.* **76** (1949): 749–759.
[13] Space–time approach to quantum electrodynamics. *Phys. Rev.* **76**: 769–789.
[18] With Laurie M. Brown. Radiative corrections to Compton scattering. *Phys. Rev.* **85** (1952): 231–244.
[45] The present status of quantum electrodynamics. *Extrait des rapports et discussions, Solvay, Institut International de Physique* (1961).

[4]Reference [45], p. 89.

PHYSICAL REVIEW VOLUME 74, NUMBER 10 NOVEMBER 15, 1948

Relativistic Cut-Off for Quantum Electrodynamics

RICHARD P. FEYNMAN
Cornell University, Ithaca, New York

(Received July 12, 1948)

A relativistic cut-off of high frequency quanta, similar to that suggested by Bopp, is shown to produce a finite invariant self-energy for a free electron. The electromagnetic line shift for a bound electron comes out as given by Bethe and Weisskopf's wave packet prescription. The scattering of an electron in a potential, without radiation, is discussed. The cross section remains finite. The problem of polarization of the vacuum is not solved. Otherwise, the results will in general agree essentially with those calculated by the prescription of Schwinger. An alternative cut-off procedure analogous to one proposed by Wataghin, which eliminates high frequency intermediate states, is shown to do the same things but to offer to solve vacuum polarization problems as well.

THE main problems of quantum electrodynamics have been essentially solved by the observations of Bethe[1] and of Weisskopf[2] that the divergent terms in the line shift problem can be thought to be contained in a renormalization of the mass of a free electron. That this principle applies as well to other problems was demonstrated by Lewis[3] in analyzing the radiationless scattering of an electron in a potential. Ambiguities which remained in the subtraction procedures are removed by Schwinger.[2,4] He formulated, in a general way, which terms are to be identified in a future correct theory with rest mass, and hence should be omitted from a calculation which does not renormalize the mass. These results are remarkable because they solve the problem without the addition of any new fundamental lengths or dimensions.

The solution given by Schwinger does, however, assume that in some future theory the divergent self-energy terms will be finite. Therefore, it is of interest to point out that there is a model, a modification of ordinary electrodynamics, for which all quantities automatically do come out finite. With this model the ideas of Bethe, Oppenheimer, and Lewis and Schwinger can be directly confirmed.

The model results from the quantization of a classical theory described in a previous paper.[5] In this paper we describe only the results for processes in which only virtual quanta are emitted and absorbed. The problems of permanent emission and the position of positron theory must be more completely studied. It is hoped that a complete physical theory may be published in the near future. Lacking such a complete picture, the present paper may be looked upon merely as presenting an arbitrary rule to cut off at high frequencies in a relativistically invariant manner, the otherwise divergent integrals appearing in quantum field theories. For electrodynamics the rule is to consider the (positive) frequency ω and wave number $\mathbf{k}$ of the field oscillators as independent and to integrate them over the density function $g(\omega^2-k^2)d\omega d\mathbf{k}$ where

$$g(\omega^2-k^2)=\int_0^\infty [\delta(\omega^2-k^2) - \delta(\omega^2-k^2-\lambda^2)]G(\lambda)d\lambda. \quad (1)$$

Here $\delta(x)$ is Dirac's delta function and $G(\lambda)$ is some smooth function such that $\int_0^\infty G(\lambda)d\lambda=1$ and for which the mean values of λ which are important are of order of the frequency 137 mc^2/h, or higher. Ordinary quantum electrodynamics replaces the function $g(\omega^2-k^2)$ by $\delta(\omega^2-k^2)$. According to (1), the density g is not everywhere positive.[5] Therefore, the model is essentially that due to Bopp.[6]

The model therefore contains an arbitrary function and the numerical results depend on the

[1] H. A. Bethe, Phys. Rev. **72**, 339 (1947); **73**, 1271A (1948).
[2] J. Schwinger and V. Weisskopf, Phys. Rev. **73**, 1272A (1948).
[3] H. W. Lewis, Phys. Rev. **73**, 173 (1948).
[4] J. Schwinger, Phys. Rev. **73**, 415A (1948).
[5] R. P. Feynman, Phys. Rev. **74**, 939 (1948).
[6] F. Bopp, Ann. d. Physik **42**, 573 (1942).

form of $G(\lambda)$. However, the only term that depends seriously (logarithmically) on the cut-off frequency is the self-energy, which can be used to renormalize the electron mass. After this is done, the remaining terms are nearly independent of the function $G(\lambda)$.

We shall illustrate these points by studying the particular examples of self-energy and radiationless scattering. We shall then discuss an alternative cut-off procedure in which the density of electron states is cut off rather than that of the quanta. This promises to solve problems of vacuum polarization which are not touched by the former procedure.

SELF-ENERGY

The transverse self-energy of a free electron, of mechanical mass μ, in state of momentum $\mathbf{P}_0$ energy $E_0=(\mu^2+P_0^2)^{\frac{1}{2}}$ is given to the first order in e^2 by the second-order perturbation theory, using the one-electron theory of Dirac, by

$$\Delta E=-\frac{e^2}{4\pi^2}\sum_i\int\frac{d\mathbf{k}}{k}\left\{\sum_+\frac{(0|\alpha_i|f)(f|\alpha_i|0)}{E_f-E_0+k}+\sum_-\frac{(0|\alpha_i|f)(f|\alpha_i|0)}{-E_f-E_0+k}\right\}. \quad (2)$$

Here the intermediate state f arises from the initial state through emission of a quantum of momentum $\mathbf{k}$ and of energy $k=|\mathbf{k}|$ (the velocity of light is taken as unity, as is Planck's constant). Thus in the intermediate state the electron has momentum $\mathbf{P}_f=\mathbf{P}_0-\mathbf{k}$ and an energy of magnitude $E_f=+(\mu^2+P_f^2)^{\frac{1}{2}}$ but which may be either plus or minus in sign. The sums indicate the sum over all such intermediate states (actually just two) for each sign of the energy. The terms for positive and negative energy have been separated and the sums are written $\sum_+$ and $\sum_-$ for these two cases. The $(f|\alpha_i|0)$ are the matrix elements of Dirac's α-matrices, the sum on i being over the two directions of polarization of the quanta. We shall henceforth write the integral $d\mathbf{k}/k$ over $\mathbf{k}$ space by its equivalent $2\int d\omega d\mathbf{k}\delta(\omega^2-k^2)$, the integral being over all *positive* ω, and all wave numbers $\mathbf{k}$. We shall also write ω for k in the energy denominators as we shall later wish to distinguish the energy of a quantum and the magnitude of the momentum change that its recoil represents. We may further simplify the expression by the use of the well-known projection operators:

$$\Lambda_f^{\pm}=(E_f\pm H_f)/2E_f=(E_f\pm\boldsymbol{\alpha}\cdot\mathbf{P}_f\pm\beta\mu)/2E_f.$$

According to the theory of holes, the last term, the transition to negative energy states, is to be left out; such transitions are prevented because the negative levels are already occupied. On the other hand, in the vacuum, electrons in state of energy $-E_f$ could make virtual transitions to positive energy state E_0. This is now prevented by the presence of an electron in the state E_0, so that, relative to the vacuum, the transverse self-energy is

$$\Delta E=-\frac{e^2}{2\pi^2}\sum_i\int d\omega d\mathbf{k}\delta(\omega^2-k^2)\times\left\{\frac{(0|\alpha_i\Lambda_f^+\alpha_i|0)}{E_f-E_0+\omega}-\frac{(0|\alpha_i\Lambda_f^-\alpha_i|0)}{E_f+E_0+\omega}\right\}. \quad (3)$$

The treatment of the longitudinal self-energy is usually different, for the longitudinal oscillators are first eliminated from the Hamiltonian, their effect being the term e^2/r_{00} where r_{00} is the meaningless distance of the electron from itself. These terms must be expressed as integrals over oscillators and combined with (3) before the change suggested by (1) is to be performed. An additional point of confusion is that the longitudinal elimination assumes the intermediate states to form a complete set as they do in (2), but the situation in (3) is not so clear. Fortunately, all these points may be most easily circumvented by simply not eliminating the longitudinal oscillators from the field Hamiltonian at all. One need simply to specify that the sum on i in (3) now be interpreted to mean the sum over each of three perpendicular space directions minus a term for the time direction. We may write $\sum_i\alpha_i\Lambda\alpha_i=\boldsymbol{\alpha}\cdot\Lambda\boldsymbol{\alpha}-\Lambda$, which is a relativistic combination since $\alpha_4=1$. One does not need to be concerned about the gauge condition in a problem in which all quanta are virtual, for the quanta are created by a charge which is conserved. This solution automatically insures the gauge condition just as the Lienard Wiechert classical solution of the Maxwell equations will automatically satisfy the gauge

condition if the charge which produces the potential is conserved.

With this convention for $\sum_i$, Eq. (3) represents the total self-energy. It is easily calculated. The numerator of first term may be written as $1/2E_f$ times $\sum_i(0|\alpha_i(H_f+E_f)\alpha_i|0)$ where H_f is $\boldsymbol{\alpha}\cdot\mathbf{P}_f+\beta\mu$. Now since $\sum_i \alpha_i\alpha_i=+2$, $\sum_i \alpha_i\beta\alpha_i=-4\beta$, and $\sum_i \alpha_i\boldsymbol{\alpha}\alpha_i=-2\boldsymbol{\alpha}$, this becomes

$$-2(0|-E_f+2\beta\mu+\boldsymbol{\alpha}\cdot\mathbf{P}_f|0).$$

The diagonal elements of β and $\boldsymbol{\alpha}$ for the state 0 are μ/E_0 and $\mathbf{P}_0/E_0$, respectively.

The change in energy ΔE_0 can, since the momentum is given, be represented as a change $\Delta\mu$ in rest mass of the electron. In virtue of the general relation $E^2=\mu^2+P^2$, the relation between these quantities is $\mu\Delta\mu=E_0\Delta E_0$. Thus we find, treating the sum of negative energies in a similar manner,

$$\Delta\mu_0=\frac{e^2}{2\pi^2\mu}\int d\omega d\mathbf{k}\delta(\omega^2-k^2)\left\{\frac{2\mu^2-E_0E_f+\mathbf{P}_0\cdot\mathbf{P}_f}{E_f(E_f-E_0+\omega)}+\frac{2\mu^2+E_0E_f+\mathbf{P}_0\cdot\mathbf{P}_f}{E_f(E_f+E_0+\omega)}\right\}. \quad (4)$$

The integral diverges logarithmically and $\Delta\mu_0$ defined here is meaningless. If the $\delta(\omega^2-k^2)$ is replaced by $g(\omega^2-k^2)$ defined in (1), the result is finite and invariant (i.e., does not depend on the momentum $\mathbf{P}_0$ of the electron).

How this comes about may be seen by calculating the integral in (4) for

$$g(\omega^2-k^2)=\delta(\omega^2-k^2)-\delta(\omega^2-k^2-\lambda^2)$$

and reserving an integration on λ until later. The integral (4) will converge with this $g(\omega^2-k^2)$, but it is convenient to divide it for purposes of calculation into the difference of two diverging ones.

This is legitimate providing the divergent integrals are first both computed over the same finite region of $\mathbf{k}$ space, the difference taken, and then the region allowed to pass to infinity. Therefore, we shall define $\Delta\mu_0$ by (4), in which we choose the region arbitrarily to be first over all (positive) ω and then over a sphere in $\mathbf{k}$ space of very large radius K. Likewise $\Delta\mu_\lambda$ is defined as expression (4) with $\delta(\omega^2-k^2-\lambda^2)$ replacing $\delta(\omega^2-k^2)$, and the integration taken over the same region.

The true self-mass is therefore

$$\Delta\mu=\int_0^\infty[\operatorname{Lim}_{K\to\infty}(\Delta\mu_0-\Delta\mu_\lambda)]G(\lambda)d\lambda. \quad (5)$$

We may now calculate these integrals, starting with $\Delta\mu_\lambda$. Since $\mathbf{P}_0\cdot\mathbf{P}_f=\mathbf{P}_0\cdot(\mathbf{P}_0-\mathbf{k})=E_0^2-\mu^2-\mathbf{P}_0\cdot\mathbf{k}$ and $E_f^2=E_0^2+k^2-2\mathbf{P}_0\cdot\mathbf{P}_f$, the $\mathbf{P}_0\cdot\mathbf{P}_f$ term in the numerator of the first term may be eliminated, the numerator becoming

$$\tfrac{1}{2}(E_f^2+E_0^2-k^2)-E_0E_f+\mu^2=\mu^2+\tfrac{1}{2}(\omega^2-k^2)+\tfrac{1}{2}(E_f-E_0-\omega)(E_f-E_0+\omega).$$

Thus the first term in $\Delta\mu_\lambda$ becomes

$$\int\frac{\mu^2+\frac{1}{2}(\omega^2-k^2)}{E_f(E_f-E_0+\omega)}\delta(\omega^2-k^2-\lambda^2)d\omega d\mathbf{k}+\tfrac{1}{2}\int\delta(\omega^2-k^2-\lambda^2)d\omega d\mathbf{k}(E_f-E_0-\omega)/E_f. \quad (6)$$

Adding the corresponding second term which differs from the first only in the sign of E_0, and performing the integral on ω (which requires simply division by 2ω), we find

$$(2\pi^2/e^2)\mu\Delta\mu_\lambda=(\mu^2+\tfrac{1}{2}\lambda^2)\int\frac{1}{(E_f+\omega)^2-E_0^2}\cdot\frac{E_f+\omega}{E_f}\frac{d\mathbf{k}}{\omega}+\frac{1}{2}\int\frac{d\mathbf{k}}{\omega}-\frac{1}{2}\int\frac{d\mathbf{k}}{E_f}, \quad (7)$$

where $\omega=(k^2+\lambda^2)^{\frac{1}{2}}$ and the integration is to be taken over a sphere of radius K in $\mathbf{k}$ space. The first and, obviously, the second integrals turn out to be invariant; the third is not, but its contribution will cancel out on taking $\Delta\mu_0-\Delta\mu_\lambda$ as it does not depend on λ.[7] The result of the integrations[8] is, dropping terms of order $1/K$ and

[7] Pais has suggested that one subtract from $\Delta\mu_0$ the $-\Delta\mu_\lambda$ that one gets not from electrodynamics but from the scalar f field (for which $\beta\cdots\beta$ replaces $\Sigma_i\,\alpha_i\cdots\alpha_i$). Proceeding in this way the integrals $\int d\mathbf{k}/E_f$ do not appear with the same coefficient. Therefore, although this procedure leads to a finite rest mass it is not invariant in the sense here, that the limits of $\mathbf{k}$ space integration can be taken to be independent of the momentum of the electron. A. Pais, Kon. Ned. Akad. v. Wet. Verh. D1, 19, 1 (1947).

[8] The first integral may be performed in the following manner: First integrate over the directions in $\mathbf{k}$ space at constant magnitude k. Only E_f depends on the direction of k and one may therefore replace the solid angle integral by one on E_f. The limits of E_f are $E_+=(\mu^2+(P_0+k)^2)^{\frac{1}{2}}$ and $E_-=(\mu^2+(P_0-k)^2)^{\frac{1}{2}}$ but both terms may be considered together as one if the integral on k be extended from $-K$ to K instead of 0 to K. To integrate this on k, substitute the variable $x=E_++\omega-E_0$ and (the algebra is long) integrate by parts to reduce it to elementary integrals.

smaller:

$$(\pi/e^2)\mu\Delta\mu_\lambda=(\mu^2+\tfrac{1}{2}\lambda^2)[N_\lambda+\mu^2X_\lambda(\mu,\mu)]+\tfrac{1}{2}[K^2-\lambda^2(\ln(2K/\lambda)-\tfrac{1}{2})]-\tfrac{1}{2}[K^2-\tfrac{1}{3}P_0^2-\mu^2(\ln(2K/\mu)-\tfrac{1}{2})],$$

where

$$N_\lambda=N_0-[\lambda^2/(\lambda^2-\mu^2)]\ln(\lambda/\mu), \tag{7a}$$

with $N_0=\ln(2K/\mu)-\frac{1}{2}$, and the quantity $X_\lambda(\mu,\mu)$ is finite as $K\to\infty$. It is given by setting $\mu_0=\mu$ in the complicated expression

$$2\mu_0^4X_\lambda(\mu,\mu_0)=((\lambda^2-\mu^2-\mu_0^2)^2-4\mu^2\mu_0^2)^{\frac{1}{2}}\times\ln\frac{\lambda^2+\mu^2-\mu_0^2+((\lambda^2-\mu^2-\mu_0^2)^2-4\mu^2\mu_0^2)^{\frac{1}{2}}}{2\lambda\mu}+\left(\lambda^2-\mu^2+\mu_0^2-\frac{2\lambda^2\mu_0^2}{\lambda^2-\mu^2}\right)\ln\frac{\mu}{\lambda}+\mu_0^2. \tag{7b}$$

Thus $X_0(\mu,\mu)=1/2\mu^2$ and for λ large compared to μ, $X_\lambda(\mu,\mu)=1/4\lambda^2$. Hence

$$(\pi/e^2)(\Delta\mu_0-\Delta\mu_\lambda)=\frac{3\mu}{2}\cdot\frac{\lambda^2}{\lambda^2-\mu^2}\cdot\ln\frac{\lambda}{\mu}+\frac{\mu}{2}-(\mu^2+\tfrac{1}{2}\lambda^2)\mu X_\lambda(\mu,\mu), \tag{8}$$

which is independent of K (in the limit $K\to\infty$). If the important values of λ are much greater than μ, we find approximately (to terms of order $(\mu/\lambda)^2$)

$$\Delta\mu=\mu(e^2/\pi)[\tfrac{3}{2}\ln(\lambda_0/\mu)+\tfrac{3}{8}], \tag{9}$$

where

$$\ln\lambda_0=\int_0^\infty\ln\lambda G(\lambda)d\lambda.$$

Judging from the classical case we would have expected to take λ_0 of order 137μ, for then all mass would be electromagnetic. But $\Delta\mu$ here is too small for this to represent a real possibility. The experimental electron mass m is of course $\mu+\Delta\mu$.

The value of λ would have to be of phenomenal size ($\sim e^{137}\mu$) before $\Delta\mu$ can represent a sizeable fraction of the experimental mass. However, to go to the limit of the conventional electrodynamics, λ_0 should be taken as infinite. Then the self-energy diverges logarithmically in the manner found by Weisskopf.[9]

[9] V. Weisskopf, Phys. Rev. 56, 72 (1939).

The emission and subsequent absorption of a quantum acts similarly to the effect of a change in mass not only on the diagonal matrix element which we have just calculated, but on non-diagonal elements as well. Consider that the state appearing on the left of all the matrices in (3) were arbitrary, say x. Then the numerator of the first term can be expressed, as we have seen, by $(-1/E_f)(x|-E_f+2\beta\mu+\boldsymbol{\alpha}\cdot\mathbf{P}_f|0)$. The second term can be expressed similarly. The two terms can be combined so that the whole expression in brackets in (3) can be written

$$2\left\{\frac{(x|-E_fE_0+(E_f+\omega)(2\beta\mu+\boldsymbol{\alpha}\cdot\mathbf{P}_0-\boldsymbol{\alpha}\cdot\mathbf{k})|0)}{E_f((E_f+\omega)^2-E_0^2)}\right\}. \tag{10}$$

This expression may be multiplied by

$$\delta(\omega^2-k^2-\lambda^2)$$

and integrated with respect to ω and over a sphere of radius K in $\mathbf{k}$ space. We make use of the following integrals which can be directly verified:

$$\int\frac{1}{(E_f+\omega)^2-E_0^2}\cdot\frac{E_f+\omega}{E_f}\delta(\omega^2-k^2-\lambda^2)d\omega d\mathbf{k}/\pi=N_\lambda+\mu_0^2X_\lambda(\mu,\mu_0),$$

$$\int\frac{1}{(E_f+\omega)^2-E_0^2}\delta(\omega^2-k^2-\lambda^2)d\omega d\mathbf{k}/\pi=\tfrac{1}{2}N_\lambda+\tfrac{1}{2}(\mu^2+\mu_0^2-\lambda^2)X_\lambda(\mu,\mu)_0, \tag{11}$$

$$\int\frac{\mathbf{k}}{(E_f+\omega)^2-E_0^2}\cdot\frac{E_f+\omega}{E_f}\delta(\omega^2-k^2-\lambda^2)d\omega d\mathbf{k}/\pi=\tfrac{1}{2}\mathbf{P}_0[\tfrac{1}{3}+N_\lambda+(\lambda^2+\mu_0^2-\mu^2)X_\lambda(u,\mu_0)].$$

The integrals have been calculated under the assumption that $E_0^2=\mu_0^2+P_0^2$. In our application we should take $\mu_0=\mu$. The quantities N_λ and $X_\lambda(\mu,\mu_0)$ are given by (7a), (7b). (The extra parameter μ_0 is helpful in obtaining other integrals, useful in the radiationless scattering problem, by differentiations with respect to the various parameters under the integral sign.)

The result of integration (10) with the density $\delta(\omega^2-k^2)-\delta(\omega^2-k^2-\lambda^2)$ is therefore

$$(e^2/\pi)(x|-E_0(\tfrac{1}{2}(N_0-N_\lambda)+\tfrac{1}{2}-(\mu^2-\tfrac{1}{2}\lambda^2)X_\lambda)+(2\beta\mu+\boldsymbol{\alpha}\cdot\mathbf{P}_0)(N_0-N_\lambda+\tfrac{1}{2}-\mu^2X_\lambda)-\tfrac{1}{2}\boldsymbol{\alpha}\cdot\mathbf{P}_0(N_0-N_\lambda-\lambda^2X_\lambda)|0).$$

Now the energy of state 0 is E_0 so that $\boldsymbol{\alpha}\cdot\mathbf{P}_0=H_0-\beta\mu$ is equivalent to $E_0-\beta\mu$, since it operates on state 0 (no implication about state x is involved). Making this replacement, all the terms in E_0 are seen to cancel and the result is simply

$$(x|\beta|0)\cdot(\Delta\mu_0-\Delta\mu_\lambda), \tag{12}$$

where $\Delta\mu_0-\Delta\mu_\lambda$ is given in (8). On integrating over $G(\lambda)d\lambda$ then we find $(x|\beta\Delta\mu|0)$. But this is just the perturbation element which would result from a change of mass by $\Delta\mu$ in the Dirac equation.

We may use this result to show that the level shift for an electron in a bound state given in the present theory will be essentially that given by Weisskopf and Bethe according to their so-called wave-packet method. The change in energy of our electron in a bound state may be calculated in a straightforward manner according to the present formulation. One would simply start with Eq. (2) but with the wave functions and energies for states 0 and f being appropriate for the potential by which the electron is bound. Then one would integrate over $g(\omega^2-k^2)$ rather than $\delta(\omega^2-k^2)$ and obtain a definite finite result. The result would show a fairly large change in E_0 depending logarithmically on λ.

A good part of this change could be accounted for as simply due to the change in E_0 that would occur if the mass of the electron were altered from μ to $m=\mu+\Delta\mu$. We can define the true term shift, then, as the complete change in E_0, less $\Delta\mu(\partial E_0/\partial\mu)$, the change due to using μ instead of m in computing the energy with radiation absent. But $\partial E_0/\partial\mu$ is by perturbation theory the expected value $(\psi_0^*|\beta|\psi_0)$ of β for the state ψ_0 in question. From the result (12), however, this is also equivalent to computing the self-energy of a wave packet ψ_0, assuming the electron as free. But Bethe[1] and Weisskopf[2] compute their term shift by just this prescription: the total effect less the self-energy of the free packet. The only difference here is that we would compute the term shift integral on $g(\omega^2-k^2)$ rather than $\delta(\omega^2-k^2)$. But since the integral converges either way, the difference between the two results is very small, being of order of (μ^2/λ_0^2) times smaller than the result.

RADIATIONLESS SCATTERING

We can study the radiationless scattering problem in a similar manner. This problem is the correction to the scattering by a first-order potential due to the possibility of emission and absorption of a virtual quantum. For example, this emission and absorption can occur at any time previous to the scattering. (It would, in this case, be nearly equivalent to a change in mass in the wave function of the electron arriving at the scatterer.) There will be a large change in cross section, which would be expected as the result of a change in mass of the electron plus a smaller change caused essentially by emissions previous to and absorptions subsequent to the scattering. As in the case of the self-energy in a field and, in fact, in all such problems, we will really be interested in those effects of radiation over and above that resulting from the change in mass. It is, therefore, simpler to compute the difference between the desired quantity calculated with no radiation and electrons of mass m, and the same quantity computed with the possibility of a virtual quantum emission and absorption with an electron of mass μ. This difference, which we shall call the radiative correction, can be looked upon as the result of perturbation due to the addition to the Hamiltonian of both the radiative interaction terms and a term $-\beta\Delta\mu$. The latter term can, as we have shown, be represented by the integral over oscillators of

$$-\sum_i\left(\frac{\alpha_i\Lambda_f^+\alpha_i}{E_f-E_0+\omega}-\frac{\alpha_i\Lambda_f^-\alpha_i}{E_f+E_0+\omega}\right) \tag{13}$$

when acting on a free electron state of positive energy E_0 and momentum $\mathbf{P}_0$. When acting on a state of negative energy $-E_0$, the term can be shown in a similar manner to be the expression (13) with the sign of E_0 changed in the denominator.

Terms like these are just the ones that Schwinger[4] thought should be omitted from the Hamiltonian if one wishes to get meaningful results, so that the present model agrees with Schwinger's prescription.

When this proces is applied to the scattering problem to obtain the radiative correction to the matrix elements, we are left with several

residual terms. First, the emissions and absorptions previous to scattering are not exactly equivalent to a change in mass. If the emission occurs too close (in time) to the scattering, the absorption must occur in a restricted time, rather than at leisure as for a free electron forming $\beta\Delta\mu$. The correction to the matrix element (in the theory of holes) for this is proportional to

$$-\tfrac{1}{2}\sum_i \frac{(2|\,V\Lambda_1{}^+\alpha_i\Lambda_f{}^+\alpha_i|\,1)}{(E_f+\omega-E_1)^2} -\tfrac{1}{2}\sum_i \frac{(2|\,V\Lambda_1{}^+\alpha_i\Lambda_f{}^-\alpha_i|\,1)}{(E_f+\omega+E_1)^2}. \quad (14)$$

We assume the potential V (vector or scalar) depending on position like $e^{i\mathbf{q}\cdot\mathbf{R}}$ and time like e^{-iQt} induces transitions from a state 1 of momentum $\mathbf{P}_1$, energy E_1, to the state 2 of momentum $\mathbf{P}_2=\mathbf{P}_1+\mathbf{q}$, energy $E_2=E_1+Q=(\mu^2+P_2^2)^{\frac{1}{2}}$. The operator V is just 1 for scalar potential, α_x for vector potential in x direction, etc. The term (14) represents only that contribution due to a quantum of momentum $\mathbf{k}$, frequency ω. We expect later to integrate over ω and $\mathbf{k}$, times $g(\omega^2-k^2)$. We put $\mathbf{P}_f=\mathbf{P}_1-\mathbf{k}$, $E_f=(\mu^2+P_f^2)^{\frac{1}{2}}$. This term can also be regarded as due to the second-order normalization correction in the ordinary perturbation theory on the incoming wave function. There is a corresponding correction for the final wave function resulting from virtual quanta emitted and absorbed after the scattering: ($\mathbf{P}_g=\mathbf{P}_2-\mathbf{k}$, $E_g=(\mu^2+P_g^2)^{\frac{1}{2}}$).

$$-\tfrac{1}{2}\sum_i \frac{(2|\,\alpha_i\Lambda_g{}^+\alpha_i\Lambda_2{}^+V|\,1)}{(E_g+\omega-E_2)^2} -\tfrac{1}{2}\sum_i \frac{(2|\,\alpha_i\Lambda_g{}^-\alpha_i\Lambda_2{}^+V|\,1)}{(E_g+\omega+E_2)^2}. \quad (15)$$

All the effects of $\beta\Delta\mu$ are now included. The remaining terms are those for which the potential scattering occurs between the emission and absorption. They may be worked out as by Dancoff[10] (except that we include the longitudinal waves by summing i from 1 to 4). They are

$$+\sum_i \frac{(2|\,\alpha_i\Lambda_g{}^+V\Lambda_f{}^+\alpha_i|\,1)}{(E_f+\omega-E_1)(E_g+\omega-E_2)} +\sum_i \frac{(2|\,\alpha_i\Lambda_g{}^-V\Lambda_f{}^-\alpha_i|\,1)}{(E_f+\omega+E_1)(E_g+\omega+E_2)} \quad (16)$$

and

$$-\sum_i \frac{(2|\,\alpha_i\Lambda_g{}^+V\Lambda_f{}^-\alpha_i|\,1)}{(E_g+\omega-E_2)(E_f+\omega+E_1)} \times\left[1+\frac{2\omega}{E_f+E_g-E_2+E_1}\right] -\sum_i \frac{(2|\,\alpha_i\Lambda_g{}^-V\Lambda_f{}^+\alpha_i|\,1)}{(E_f+\omega-E_1)(E_g+\omega+E_2)} \times\left[1+\frac{2\omega}{E_f+E_g+E_2-E_1}\right]. \quad (17)$$

Although each separate term diverges, the sum of (14), (15), (16), (17) will lead to an integral convergent for large $\mathbf{k}$ even if integrated in the conventional manner on $\delta(\omega^2-k^2)$. This is the result of Lewis. Integration on $g(\omega^2-k^2)$ will make each term converge for large $\mathbf{k}$, but will then only make correction to the sum of order $(\mu/\lambda)^2$ smaller. These we shall neglect.

The integrals do, however, diverge logarithmically at the lower limit of small momentum transfer. This infra-red catastrophe has been completely cleared up by Bloch and Nordsieck.[11] They show that for very long wave-length quanta the amplitude for emission and reabsorption of more than one quantum is not negligible. Inclusion of these higher order terms, which is necessary only in the non-relativistic region, solves the problem. To keep the results given here in a simple form, we can imagine the integrals to be performed down to some minimum momentum $k_{\min}$, small compared to μ. What is effectively the same thing but which is easier (because relativistic invariance is maintained) for practical purposes, is to imagine that the quanta have a very small rest mass $\lambda_{\min}$. Thus we integrate the density

$$\delta(\omega^2-k^2-\lambda_{\min}{}^2)d\omega d\mathbf{k}$$

[10] S. M. Dancoff, Phys. Rev. 55, 959 (1939).

[11] F. Bloch and A. Nordsieck, Phys. Rev. 52, 54 (1937).

and assume $\lambda_{\min} \ll \mu$. The two methods are equivalent if one replaces $\ln\lambda_{\min}$ by $\ln(2k_{\min})-1$.

The integrals may be expanded in powers of $\mathbf{q}$ and Q, say up to the second.[12] The constant term vanishes on integration. The integrals appearing may all be expressed in terms of various parametric derivatives of the integrals already given in (11). The result may be expressed in terms of a general potential in a very simple way. A term linear in $\mathbf{q}$, such as proportional to q_x say, is equivalent to taking the matrix element $(2|q_x \exp(i\mathbf{q}\cdot\mathbf{R})|1)$ directly between the two states 2, 1. But this is also equivalent to the matrix element of $-i(\partial/\partial x)\exp(i\mathbf{q}\cdot\mathbf{R})$. Thus if the potential varied in any other manner in space, one sees by superposition that the matrix element is the same as that of $-i\partial V/\partial x$. Thus the terms up to second order can be represented by matrix elements of first and second space and time derivatives of the potential. That is, the radiative correction to the scattering in any potential is equivalent to the first order in e^2 and in the potential, to the scattering produced by a perturbation ΔH to the Dirac Hamiltonian. The perturbation up to terms of first and second derivatives of the vector potential $\mathbf{A}$ and the scalar potential φ is calculated in this manner to be

$$\Delta H=\frac{e^2}{2\pi\hbar c}\left\{-\frac{\hbar e}{2\mu c}(\beta(\boldsymbol{\sigma}\cdot\mathbf{B})-i\beta\boldsymbol{\alpha}\cdot\mathbf{E})+\frac{2\hbar^2 e}{3\mu^2c^2}(\Box^2\varphi-\boldsymbol{\alpha}\cdot\Box^2\mathbf{A})\left(\ln\frac{\mu}{\lambda_{\min}}-\frac{3}{8}\right)\right\}. \quad (18)$$

The first term, where $\mathbf{B}=\nabla\times\mathbf{A}$ and $\mathbf{E}=-\nabla\varphi-(1/c)\partial\mathbf{A}/\partial t$, has the same effect as an alteration in the electron magnetic moment[13] by a fraction $e^2/2\pi\hbar c$. This effect was first discovered by Schwinger.[4]

LINE SHIFT

The perturbation to H given here is useful not only for scattering problems but also for the line-shift problem. The actual motion of an electron in a binding potential can be visualized as simply a continued sequence of scatterings in this potential. For each scattering we can calculate the effect of virtual quanta in the way outlined above. However, it is possible, if the potential is strong, that *two* scatterings occur between the emission and reabsorption of the quantum, in which case the above formula for ΔH is incorrect. In hydrogen the potential over most of the atom is sufficiently weak that this does not occur with effective probability. The very long wave-length quanta do have a tendency to exist in the virtual state for long periods, but they have been eliminated by the cut-off $\lambda_{\min}$ at low frequencies.

In hydrogen, then, the line shift due to quanta above minimum wave number $k_{\min}$ is the expected value, for the state in question, of

$$\Delta H=\frac{e^2}{2\pi\hbar c}\left\{-\frac{\hbar e i}{2\mu c}\beta\boldsymbol{\alpha}\cdot\nabla\varphi+\frac{2\hbar^2 e}{3\mu^2c^2}(\nabla^2\varphi)\times\left(\ln\frac{\mu c}{2\hbar k_{\min}}+\frac{5}{8}\right)\right\}, \quad (19)$$

where $\varphi=e/r$, r being the distance to the proton, and we have used the relation

$$\ln\lambda_{\min}=\ln(2k_{\min})-1.$$

The first term insures that the fine structure separation correction will be that expected from the change in the electron's magnetic moment. The second may be combined with Bethe's nonrelativistic calculation for quanta below $k_{\min}$.[14]

APPLICATION TO OTHER PROCESSES

The important problem of verifying that the self-energy will not diverge in higher-order approximations has not been carried to completion. It appears unlikely that trouble will arise here. If that is true the model probably gives sensible answers to all problems of quantum electrodynamics other than those involving Uehling polarization effects, discussed below. It has been found to give finite self-mass if we have, instead of a vector field, a scalar field or a pseudoscalar field, coupled to the electron in the simplest way possible without gradient operators. If the field

[12] The integrals have also been worked out, by other methods, for arbitrarily large $\mathbf{q}$ and Q. These will appear in a future publication.

[13] W. Pauli, *Handbuch der Physik* (1933), Vol. 24/1, p. 233.

[14] Using Eq. (18), Professor Bethe finds 1050 megacycles for the separation between $2p_{3/2}$ and $2s_{1/2}$ in hydrogen. (Solvay Report.)

quanta have mass M, $g(\omega^2-k^2)$ is replaced by $g(\omega^2-k^2-M^2)$, and the values of λ of importance are chosen to be large compared to M.

The results for electrodynamics, then, after mass renormalization, depend only slightly on the form of $G(\lambda)$ and the size of λ_0. Since λ_0 may be taken to be extremely large without spoiling the smallness of $\Delta\mu$, there would appear to be good reason to drop the dependence on λ altogether. Thus the $G(\lambda)$ appears only as a complicated scaffold which is removed after the calculation is done.

On the other hand, electrodynamics probably does break down somewhere and it is interesting to keep the terms in λ for various phenomena to see if one might be selected which is particularly sensitive to λ. This phenomena would then be a promising one to study experimentally. The Møller interaction between two electrons is modified by the present theory. There is, of course, the radiative correction, but in addition to that there is simply a change due to the change in the density function for the quanta which can be exchanged. The Møller interaction ordinarily is proportional to $1/q^2$, where q is the magnitude of the momentum transferred from one electron to the other in the center of gravity system. The modification is only that this factor is changed to $\int_0^\infty(1/q^2-1/(q^2+\lambda^2))G(\lambda)d_\lambda$. This represents a decrease in cross section for hard collisions. If λ is of order 137 μc^2, we would need electrons in the center of gravity system of roughly 30 Mev to find a strong effect. This corresponds, however, to bombardment of stationary electrons by electrons of $3\frac{1}{2}$ Bev.[15]

It is interesting to note that the Møller interaction can be viewed as simply a correction to self-energy due to the exclusion principle. The self-energy of two electrons, 1 and 2, is not the sum of the self-energy of each, for one of the virtual states that 2 could ordinarily enter by emission of a quantum is now occupied by 1. The difference between the self-energy of two electrons and the sum of the self-energy of each separately comes out to be just their interaction energy.

[15] A more promising way to obtain processes with high momentum transfer would be wide-angle scattering of electrons from nuclei. But here deviations from expectations might be associated with uncertainties in the nuclear charge distributions rather than electrodynamics. Very wide angle pair production is a phenomena which does occur for high energy incident γ-rays with large momentum transfer in a region not too close to the nucleus.

VACUUM POLARIZATION. ALTERNATIVE CUT-OFF PROCEDURES

In the above calculation, terms of the type discussed by Uehling[16] have been omitted. These terms represent processes involving a pair production followed by annihilation of the *same* pair. For example, a pair produced by the potential may annihilate again emitting a quantum. This quantum is then absorbed by the electron in state 1 transferring it to state 2. These terms are infinite and are not made convergent by the present scheme. There is some point, nevertheless, to solving problems at first without taking them into account. This is because their net effect is only to alter the effective potential in which the electron finds itself, for it may be scattered either directly or by the quantum produced by the Uehling terms. That is, if this problem of polarization of the vacuum is solved it will mean, if there is any effect, simply that the potential A, φ appearing in the Dirac equation and (to high order) in such terms as (18) should be replaced by new "polarized" potentials A', φ'.

These polarization terms can be characterized in a relativistically invariant manner. All the terms which have been calculated above contain matrix elements of operators between states in a sequence such as 1 to f, f to g, g to 2. The omitted polarization terms contain transitions like f to g, g to f, 1 to 2. For higher order processes the polarization terms are those which do not contain a continued sequence of transitions from the initial to the final state.

The polarization terms are not affected in any helpful way by the changes in the density of quanta. It is likely that this problem will have its answer in a changed physical viewpoint. However, there is a simple alternative procedure to produce finite self-energies which also makes convergent the integrals appearing in Serber's[17] treatment of the polarization problem. (Since, however, this treatment of Serber already presupposes a partial subtraction procedure of Heisenberg and Dirac, the situation is not so clear here as in the self-energy problem.)

[16] E. A. Uehling, Phys. Rev. **48**, 55 (1935).
[17] R. Serber, Phys. Rev. **48**, 49 (1935).

From the point of view of coordinate space, the reason that the electronic self-energy diverges appears to be this. A virtual light quantum emitted at one point spreads out as $\delta(t^2-r^2)$ from the origin. The wave packet of the electron spreading out after the emission of the quantum has, as a consequence of Dirac's equation, a similar discontinuous value along the light cone. It is the continued coincidence of these singularities which makes the matrix element for the subsequent absorption of the quantum infinite. The method outlined above of changing $\delta(\omega^2-k^2)$ to $g(\omega^2-k^2)$ has the effect of changing $\delta(t^2-r^2)$ to $f(t^2-r^2)$ where $f(s^2)$ is everywhere finite and goes to zero rapidly for $|s^2|>1/\lambda_0^2$. The quanta have been moved away from the electrons so that overlap on the light cone is reduced.

An obvious alternative procedure is to move the electron wave function away from the quanta. This is easily done in a very similar manner. We assume the density of electron states of energy E, momentum $\mathbf{P}$ to be $g(E^2-P^2-\mu^2)$ rather than $\delta(E^2-P^2-\mu^2)$.[18] The quanta are conventional, $\omega=k$, density $d\mathbf{k}/k$. The self-energy integrals (2) can, of course, be expressed as an integral over the intermediate state momentum $\mathbf{P}_f$ rather than $\mathbf{k}$. Replacing $d\mathbf{P}_f/E_f$ by $g(E_f^2-P_f^2-\mu^2)dE_f d\mathbf{P}_f$, we find

$$\Delta E_0' = -\frac{e^2}{2\pi^2}\int g(E_f^2-P_f^2-\mu^2)dE_f d\mathbf{P}_f \cdot\frac{E_f}{k}\sum_i\left\{\frac{(0|\alpha_i\Lambda_f^+\alpha_i|0)}{E_f+k-E_0}-\frac{(0|\alpha_i\Lambda_f^-\alpha_i|0)}{E_f+k+E_0}\right\},$$

where $k=|\mathbf{P}_f-\mathbf{P}_0|$, $E_0=(\mu^2+P_0^2)^{\frac{1}{2}}$. The projection operators are unchanged since it is only the density of states which we wish to alter. They are still $\Lambda_f^{2\pm}=(E_f\pm\boldsymbol{\alpha}\cdot\mathbf{P}_f\pm\beta\mu)/2E_f$. The result of this calculation is to verify that $\Delta E_0'$ is finite, (depending logarithmically on λ_0). The other problems can be analyzed in the same way.

In the problem of polarization of the vacuum, the wave functions of both electron and positron ordinarily spread with a singularity on the light core. The matrix element for their subsequent annihilation is therefore infinite. With the modification here described these wave functions are made less singular and their overlap integral is finite. The polarization integrals in Serber's article[17] may now be integrated to yield finite results.

Other than terms which might be removed by a small renormalization of charge (depending logarithmically on λ_0), the net effect in (17) would be to change the $-(\frac{3}{8})$ in the last term of (17) to $-(\frac{3}{8})-(\frac{1}{5})$. However, the real existence of such polarization corrections is, in the author's view, uncertain. These matters will be discussed in much more detail in future publications. Also reserved for future publication is a more complete physical theory from which the results reported here may be directly deduced. It yields much more powerful techniques for setting up problems and performing the required integrations.

The author would like to express his gratitude to Mr. P. V. C. Hough for assistance in the calculations and to Professor H. A. Bethe and Dr. F. Dyson and many others for useful discussions.

[18] This is seen to be essentially the method proposed by Wataghin. G. Wataghin, Zeits. f. Physik 88, 92 (1934).

The Theory of Positrons

R. P. FEYNMAN
Department of Physics, Cornell University, Ithaca, New York
(Received April 8, 1949)

The problem of the behavior of positrons and electrons in given external potentials, neglecting their mutual interaction, is analyzed by replacing the theory of holes by a reinterpretation of the solutions of the Dirac equation. It is possible to write down a complete solution of the problem in terms of boundary conditions on the wave function, and this solution contains automatically all the possibilities of virtual (and real) pair formation and annihilation together with the ordinary scattering processes, including the correct relative signs of the various terms.

In this solution, the "negative energy states" appear in a form which may be pictured (as by Stückelberg) in space-time as waves traveling away from the external potential backwards in time. Experimentally, such a wave corresponds to a positron approaching the potential and annihilating the electron. A particle moving forward in time (electron) in a potential may be scattered forward in time (ordinary scattering) or backward (pair annihilation). When moving backward (positron) it may be scattered backward in time (positron scattering) or forward (pair production). For such a particle the amplitude for transition from an initial to a final state is analyzed to any order in the potential by considering it to undergo a sequence of such scatterings.

The amplitude for a process involving many such particles is the product of the transition amplitudes for each particle. The exclusion principle requires that antisymmetric combinations of amplitudes be chosen for those complete processes which differ only by exchange of particles. It seems that a consistent interpretation is only possible if the exclusion principle is adopted. The exclusion principle need not be taken into account in intermediate states. Vacuum problems do not arise for charges which do not interact with one another, but these are analyzed nevertheless in anticipation of application to quantum electrodynamics.

The results are also expressed in momentum-energy variables. Equivalence to the second quantization theory of holes is proved in an appendix.

1. INTRODUCTION

THIS is the first of a set of papers dealing with the solution of problems in quantum electrodynamics. The main principle is to deal directly with the solutions to the Hamiltonian differential equations rather than with these equations themselves. Here we treat simply the motion of electrons and positrons in given external potentials. In a second paper we consider the interactions of these particles, that is, quantum electrodynamics.

The problem of charges in a fixed potential is usually treated by the method of second quantization of the electron field, using the ideas of the theory of holes. Instead we show that by a suitable choice and interpretation of the solutions of Dirac's equation the problem may be equally well treated in a manner which is fundamentally no more complicated than Schrödinger's method of dealing with one or more particles. The various creation and annihilation operators in the conventional electron field view are required because the number of particles is not conserved, i.e., pairs may be created or destroyed. On the other hand charge is conserved which suggests that if we follow the charge, not the particle, the results can be simplified.

In the approximation of classical relativistic theory the creation of an electron pair (electron A, positron B) might be represented by the start of two world lines from the point of creation, 1. The world lines of the positron will then continue until it annihilates another electron, C, at a world point 2. Between the times t_1 and t_2 there are then three world lines, before and after only one. However, the world lines of C, B, and A together form one continuous line albeit the "positron part" B of this continuous line is directed backwards in time. Following the charge rather than the particles corresponds to considering this continuous world line as a whole rather than breaking it up into its pieces. It is as though a bombardier flying low over a road suddenly sees three roads and it is only when two of them come together and disappear again that he realizes that he has simply passed over a long switchback in a single road.

This over-all space-time point of view leads to considerable simplification in many problems. One can take into account at the same time processes which ordinarily would have to be considered separately. For example, when considering the scattering of an electron by a potential one automatically takes into account the effects of virtual pair productions. The same equation, Dirac's, which describes the deflection of the world line of an electron in a field, can also describe the deflection (and in just as simple a manner) when it is large enough to reverse the time-sense of the world line, and thereby correspond to pair annihilation. Quantum mechanically the direction of the world lines is replaced by the direction of propagation of waves.

This view is quite different from that of the Hamiltonian method which considers the future as developing continuously from out of the past. Here we imagine the entire space-time history laid out, and that we just become aware of increasing portions of it successively. In a scattering problem this over-all view of the complete scattering process is similar to the S-matrix viewpoint of Heisenberg. The temporal order of events during the scattering, which is analyzed in such detail by the Hamiltonian differential equation, is irrelevant. The relation of these viewpoints will be discussed much more fully in the introduction to the second paper, in which the more complicated interactions are analyzed.

The development stemmed from the idea that in nonrelativistic quantum mechanics the amplitude for a given process can be considered as the sum of an ampli-

tude for each space-time path available.[1] In view of the fact that in classical physics positrons could be viewed as electrons proceeding along world lines toward the past (reference 7) the attempt was made to remove, in the relativistic case, the restriction that the paths must proceed always in one direction in time. It was discovered that the results could be even more easily understood from a more familiar physical viewpoint, that of scattered waves. This viewpoint is the one used in this paper. After the equations were worked out physically the proof of the equivalence to the second quantization theory was found.[2]

First we discuss the relation of the Hamiltonian differential equation to its solution, using for an example the Schrödinger equation. Next we deal in an analogous way with the Dirac equation and show how the solutions may be interpreted to apply to positrons. The interpretation seems not to be consistent unless the electrons obey the exclusion principle. (Charges obeying the Klein-Gordon equations can be described in an analogous manner, but here consistency apparently requires Bose statistics.)[3] A representation in momentum and energy variables which is useful for the calculation of matrix elements is described. A proof of the equivalence of the method to the theory of holes in second quantization is given in the Appendix.

2. GREEN'S FUNCTION TREATMENT OF SCHRÖDINGER'S EQUATION

We begin by a brief discussion of the relation of the non-relativistic wave equation to its solution. The ideas will then be extended to relativistic particles, satisfying Dirac's equation, and finally in the succeeding paper to interacting relativistic particles, that is, quantum electrodynamics.

The Schrödinger equation

$$i\partial\psi/\partial t = H\psi, \tag{1}$$

describes the change in the wave function ψ in an infinitesimal time Δt as due to the operation of an operator $\exp(-iH\Delta t)$. One can ask also, if $\psi(\mathbf{x}_1, t_1)$ is the wave function at $\mathbf{x}_1$ at time t_1, what is the wave function at time $t_2 > t_1$? It can always be written as

$$\psi(\mathbf{x}_2, t_2) = \int K(\mathbf{x}_2, t_2; \mathbf{x}_1, t_1)\psi(\mathbf{x}_1, t_1)d^3\mathbf{x}_1, \tag{2}$$

where K is a Green's function for the linear Eq. (1). (We have limited ourselves to a single particle of coordinate $\mathbf{x}$, but the equations are obviously of greater generality.) If H is a constant operator having eigenvalues E_n, eigenfunctions ϕ_n so that $\psi(\mathbf{x}, t_1)$ can be expanded as $\sum_n C_n\phi_n(\mathbf{x})$, then $\psi(\mathbf{x}, t_2) = \exp(-iE_n(t_2-t_1))\times C_n\phi_n(\mathbf{x})$. Since $C_n = \int \phi_n^*(\mathbf{x}_1)\psi(\mathbf{x}_1, t_1)d^3\mathbf{x}_1$, one finds (where we write 1 for $\mathbf{x}_1, t_1$ and 2 for $\mathbf{x}_2, t_2$) in this case

$$K(2, 1) = \sum_n \phi_n(\mathbf{x}_2)\phi_n^*(\mathbf{x}_1)\exp(-iE_n(t_2-t_1)), \tag{3}$$

for $t_2 > t_1$. We shall find it convenient for $t_2 < t_1$ to define $K(2, 1) = 0$ (Eq. (2) is then not valid for $t_2 < t_1$). It is then readily shown that in general K can be defined by that solution of

$$(i\partial/\partial t_2 - H_2)K(2, 1) = i\delta(2, 1), \tag{4}$$

which is zero for $t_2 < t_1$, where $\delta(2, 1) = \delta(t_2-t_1)\delta(x_2-x_1)\times\delta(y_2-y_1)\delta(z_2-z_1)$ and the subscript 2 on H_2 means that the operator acts on the variables of 2 of $K(2, 1)$. When H is not constant, (2) and (4) are valid but K is less easy to evaluate than (3).[4]

We can call $K(2, 1)$ the total amplitude for arrival at $\mathbf{x}_2, t_2$ starting from $\mathbf{x}_1, t_1$. (It results from adding an amplitude, $\exp iS$, for each space time path between these points, where S is the action along the path.[1]) The transition amplitude for finding a particle in state $\chi(\mathbf{x}_2, t_2)$ at time t_2, if at t_1 it was in $\psi(\mathbf{x}_1, t_1)$, is

$$\int \chi^*(2)K(2, 1)\psi(1)d^3\mathbf{x}_1 d^3\mathbf{x}_2. \tag{5}$$

A quantum mechanical system is described equally well by specifying the function K, or by specifying the Hamiltonian H from which it results. For some purposes the specification in terms of K is easier to use and visualize. We desire eventually to discuss quantum electrodynamics from this point of view.

To gain a greater familiarity with the K function and the point of view it suggests, we consider a simple perturbation problem. Imagine we have a particle in a weak potential $U(\mathbf{x}, t)$, a function of position and time. We wish to calculate $K(2, 1)$ if U differs from zero only for t between t_1 and t_2. We shall expand K in increasing powers of U:

$$K(2, 1) = K_0(2, 1) + K^{(1)}(2, 1) + K^{(2)}(2, 1) + \cdots. \tag{6}$$

To zero order in U, K is that for a free particle, $K_0(2, 1)$.[4] To study the first order correction $K^{(1)}(2, 1)$, first consider the case that U differs from zero only for the infinitesimal time interval Δt_3 between some time t_3 and $t_3 + \Delta t_3 (t_1 < t_3 < t_2)$. Then if $\psi(1)$ is the wave function at $\mathbf{x}_1, t_1$, the wave function at $\mathbf{x}_3, t_3$ is

$$\psi(3) = \int K_0(3, 1)\psi(1)d^3\mathbf{x}_1, \tag{7}$$

since from t_1 to t_3 the particle is free. For the short interval Δt_3 we solve (1) as

$$\psi(\mathbf{x}, t_3 + \Delta t_3) = \exp(-iH\Delta t_3)\psi(\mathbf{x}, t_3) = (1 - iH_0\Delta t_3 - iU\Delta t_3)\psi(\mathbf{x}, t_3),$$

[1] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).

[2] The equivalence of the entire procedure (including photon interactions) with the work of Schwinger and Tomonaga has been demonstrated by F. J. Dyson, Phys. Rev. **75**, 486 (1949).

[3] These are special examples of the general relation of spin and statistics deduced by W. Pauli, Phys. Rev. **58**, 716 (1940).

[4] For a non-relativistic free particle, where $\phi_n = \exp(i\mathbf{p}\cdot\mathbf{x})$, $E_n = \mathbf{p}^2/2m$, (3) gives, as is well known

$$K_0(2, 1) = \int \exp[-(i\mathbf{p}\cdot\mathbf{x}_1 - i\mathbf{p}\cdot\mathbf{x}_2) - i\mathbf{p}^2(t_2-t_1)/2m]d^3\mathbf{p}(2\pi)^{-3} = (2\pi i m^{-1}(t_2-t_1))^{-\frac{1}{2}}\exp(\tfrac{1}{2}im(\mathbf{x}_2-\mathbf{x}_1)^2(t_2-t_1)^{-1})$$

for $t_2 > t_1$, and $K_0 = 0$ for $t_2 < t_1$.

where we put $H=H_0+U$, H_0 being the Hamiltonian of a free particle. Thus $\psi(\mathbf{x}, t_3+\Delta t_3)$ differs from what it would be if the potential were zero (namely $(1-iH_0\Delta t_3)\psi(\mathbf{x}, t_3)$) by the extra piece

$$\Delta\psi=-iU(\mathbf{x}_3, t_3)\cdot\psi(\mathbf{x}_3, t_3)\Delta t_3, \tag{8}$$

which we shall call the amplitude scattered by the potential. The wave function at 2 is given by

$$\psi(\mathbf{x}_2, t_2)=\int K_0(\mathbf{x}_2, t_2; \mathbf{x}_3, t_3+\Delta t_3)\psi(\mathbf{x}_3, t_3+\Delta t_3)d^3\mathbf{x}_3,$$

since after $t_3+\Delta t_3$ the particle is again free. Therefore the change in the wave function at 2 brought about by the potential is (substitute (7) into (8) and (8) into the equation for $\psi(\mathbf{x}_2, t_2)$):

$$\Delta\psi(2)=-i\int K_0(2, 3)U(3)K_0(3, 1)\psi(1)d^3\mathbf{x}_1d^3\mathbf{x}_3\Delta t_3.$$

In the case that the potential exists for an extended time, it may be looked upon as a sum of effects from each interval Δt_3 so that the total effect is obtained by integrating over t_3 as well as $\mathbf{x}_3$. From the definition (2) of K then, we find

$$K^{(1)}(2, 1)=-i\int K_0(2, 3)U(3)K_0(3, 1)d\tau_3, \tag{9}$$

where the integral can now be extended over all space and time, $d\tau_3=d^3\mathbf{x}_3dt_3$. Automatically there will be no contribution if t_3 is outside the range t_1 to t_2 because of our definition, $K_0(2, 1)=0$ for $t_2<t_1$.

We can understand the result (6), (9) this way. We can imagine that a particle travels as a free particle from point to point, but is scattered by the potential U. Thus the total amplitude for arrival at 2 from 1 can be considered as the sum of the amplitudes for various alternative routes. It may go directly from 1 to 2 (amplitude $K_0(2, 1)$, giving the zero order term in (6)). Or (see Fig. 1(a)) it may go from 1 to 3 (amplitude $K_0(3, 1)$), get scattered there by the potential (scattering amplitude $-iU(3)$ per unit volume and time) and then go from 3 to 2 (amplitude $K_0(2, 3)$). This may occur for any point 3 so that summing over these alternatives gives (9).

Again, it may be scattered twice by the potential (Fig. 1(b)). It goes from 1 to 3 ($K_0(3, 1)$), gets scattered there ($-iU(3)$) then proceeds to some other point, 4, in space time (amplitude $K_0(4, 3)$) is scattered again ($-iU(4)$) and then proceeds to 2 ($K_0(2, 4)$). Summing over all possible places and times for 3, 4 find that the second order contribution to the total amplitude $K^{(2)}(2, 1)$ is

$$(-i)^2\int\int K_0(2, 4)U(4)K_0(4, 3)\times U(3)K_0(3, 1)d\tau_3d\tau_4. \tag{10}$$

This can be readily verified directly from (1) just as (9) was. One can in this way obviously write down any of the terms of the expansion (6).[5]

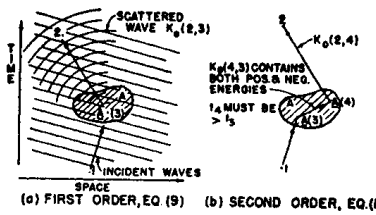


FIG. 1. The Schrödinger (and Dirac) equation can be visualized as describing the fact that plane waves are scattered successively by a potential. Figure 1 (a) illustrates the situation in first order. $K_0(2, 3)$ is the amplitude for a free particle starting at point 3 to arrive at 2. The shaded region indicates the presence of the potential A which scatters at 3 with amplitude $-iA(3)$ per cm³sec. (Eq. (9)). In (b) is illustrated the second order process (Eq. (10)), the waves scattered at 3 are scattered again at 4. However, in Dirac one-electron theory $K_0(4, 3)$ would represent electrons both of positive and of negative energies proceeding from 3 to 4. This is remedied by choosing a different scattering kernel $K_+(4, 3)$, Fig. 2.

3. TREATMENT OF THE DIRAC EQUATION

We shall now extend the method of the last section to apply to the Dirac equation. All that would seem to be necessary in the previous equations is to consider H as the Dirac Hamiltonian, ψ as a symbol with four indices (for each particle). Then K_0 can still be defined by (3) or (4) and is now a 4–4 matrix which operating on the initial wave function, gives the final wave function. In (10), $U(3)$ can be generalized to $A_4(3)-\boldsymbol{\alpha}\cdot\mathbf{A}(3)$ where A_4, $\mathbf{A}$ are the scalar and vector potential (times e, the electron charge) and $\boldsymbol{\alpha}$ are Dirac matrices.

To discuss this we shall define a convenient relativistic notation. We represent four-vectors like $\mathbf{x}$, t by a symbol x_μ, where $\mu=1, 2, 3, 4$ and $x_4=t$ is real. Thus the vector and scalar potential (times e) $\mathbf{A}$, A_4 is A_μ. The four matrices $\beta\boldsymbol{\alpha}$, β can be considered as transforming as a four vector γ_μ (our γ_μ differs from Pauli's by a factor i for $\mu=1, 2, 3$). We use the summation convention $a_\mu b_\mu=a_4b_4-a_1b_1-a_2b_2-a_3b_3=a\cdot b$. In particular if a_μ is any four vector (but not a matrix) we write $\boldsymbol{a}=a_\mu\gamma_\mu$ so that $\boldsymbol{a}$ is a matrix associated with a vector ($\boldsymbol{a}$ will often be used in place of a_μ as a symbol for the vector). The γ_μ satisfy $\gamma_\mu\gamma_\nu+\gamma_\nu\gamma_\mu=2\delta_{\mu\nu}$ where $\delta_{44}=+1$, $\delta_{11}=\delta_{22}=\delta_{33}=-1$, and the other $\delta_{\mu\nu}$ are zero. As a consequence of our summation convention $\delta_{\mu\nu}a_\nu=a_\mu$ and $\delta_{\mu\mu}=4$. Note that $\boldsymbol{ab}+\boldsymbol{ba}=2a\cdot b$ and that $\boldsymbol{a}^2=a_\mu a_\mu=a\cdot a$ is a pure number. The symbol $\partial/\partial x_\mu$ will mean $\partial/\partial t$ for $\mu=4$, and $-\partial/\partial x$, $-\partial/\partial y$, $-\partial/\partial z$ for $\mu=1$, 2, 3. Call $\nabla=\gamma_\mu\partial/\partial x_\mu=\beta\partial/\partial t+\beta\boldsymbol{\alpha}\cdot\nabla$. We shall imagine

[5] We are simply solving by successive approximations an integral equation (deducible directly from (1) with $H=H_0+U$ and (4) with $H=H_0$),

$$\psi(2)=-i\int K_0(2, 3)U(3)\psi(3)d\tau_3+\int K_0(2, 1)\psi(1)d^3\mathbf{x}_1,$$

where the first integral extends over all space and all time greater than the t_1 appearing in the second term, and $t_3>t$

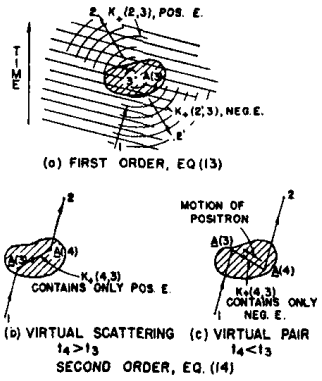


FIG. 2. The Dirac equation permits another solution $K_+(2, 1)$ if one considers that waves scattered by the potential can proceed backwards in time as in Fig. 2 (a). This is interpreted in the second order processes (b), (c), by noting that there is now the possibility (c) of virtual pair production at 4, the positron going to 3 to be annihilated. This can be pictured as similar to ordinary scattering (b) except that the electron is scattered backwards in time from 3 to 4. The waves scattered from 3 to 2′ in (a) represent the possibility of a positron arriving at 3 from 2′ and annihilating the electron from 1. This view is proved equivalent to hole theory: electrons traveling backwards in time are recognized as positrons.

hereafter, purely for relativistic convenience, that ϕ_n^* in (3) is replaced by its adjoint $\bar{\phi}_n = \phi_n^* \beta$.

Thus the Dirac equation for a particle, mass m, in an external field $A = A_\mu \gamma_\mu$ is

$$(i\nabla - m)\psi = A\psi, \tag{11}$$

and Eq. (4) determining the propagation of a free particle becomes

$$(i\nabla_2 - m)K_+(2, 1) = i\delta(2, 1), \tag{12}$$

the index 2 on ∇_2 indicating differentiation with respect to the coordinates $x_{2\mu}$ which are represented as 2 in $K_+(2, 1)$ and $\delta(2, 1)$.

The function $K_+(2, 1)$ is defined in the absence of a field. If a potential A is acting a similar function, say $K_+^{(A)}(2, 1)$ can be defined. It differs from $K_+(2, 1)$ by a first order correction given by the analogue of (9) namely

$$K_+^{(1)}(2, 1) = -i\int K_+(2, 3)A(3)K_+(3, 1)d\tau_3, \tag{13}$$

representing the amplitude to go from 1 to 3 as a free particle, get scattered there by the potential (now the matrix $A(3)$ instead of $U(3)$) and continue to 2 as free. The second order correction, analogous to (10) is

$$K_+^{(2)}(2, 1) = -\int\int K_+(2, 4)A(4) \times K_+(4, 3)A(3)K_+(3, 1)d\tau_4 d\tau_3, \tag{14}$$

and so on. In general $K_+^{(A)}$ satisfies

$$(i\nabla_2 - A(2) - m)K_+^{(A)}(2, 1) = i\delta(2, 1), \tag{15}$$

and the successive terms (13), (14) are the power series expansion of the integral equation

$$K_+^{(A)}(2, 1) = K_+(2, 1) - i\int K_+(2, 3)A(3)K_+^{(A)}(3, 1)d\tau_3, \tag{16}$$

which it also satisfies.

We would now expect to choose, for the special solution of (12), $K_+ = K_0$ where $K_0(2, 1)$ vanishes for $t_2 < t_1$ and for $t_2 > t_1$ is given by (3) where ϕ_n and E_n are the eigenfunctions and energy values of a particle satisfying Dirac's equation, and ϕ_n^* is replaced by $\bar{\phi}_n$.

The formulas arising from this choice, however, suffer from the drawback that they apply to the one electron theory of Dirac rather than to the hole theory of the positron. For example, consider as in Fig. 1(a) an electron after being scattered by a potential in a small region 3 of space time. The one electron theory says (as does (3) with $K_+ = K_0$) that the scattered amplitude at another point 2 will proceed toward positive times with both positive and negative energies, that is with both positive and negative rates of change of phase. No wave is scattered to times previous to the time of scattering. These are just the properties of $K_0(2, 3)$.

On the other hand, according to the positron theory negative energy states are not available to the electron after the scattering. Therefore the choice $K_+ = K_0$ is unsatisfactory. But there are other solutions of (12). We shall choose the solution defining $K_+(2, 1)$ so that *$K_+(2, 1)$ for $t_2 > t_1$ is the sum of (3) over positive energy states only.* Now this new solution must satisfy (12) for all times in order that the representation be complete. It must therefore differ from the old solution K_0 by a solution of the homogeneous Dirac equation. It is clear from the definition that the difference $K_0 - K_+$ is the sum of (3) over all negative energy states, as long as $t_2 > t_1$. But this difference must be a solution of the homogeneous Dirac equation for all times and must therefore be represented by the same sum over negative energy states also for $t_2 < t_1$. Since $K_0 = 0$ in this case, it follows that our new kernel, *$K_+(2, 1)$, for $t_2 < t_1$ is the negative of the sum (3) over negative energy states.* That is,

$$\begin{aligned} K_+(2, 1) &= \textstyle\sum_{POS\ E_n} \phi_n(2)\bar{\phi}_n(1) \times \exp(-iE_n(t_2 - t_1)) \quad \text{for} \quad t_2 > t_1 \\ &= -\textstyle\sum_{NEG\ E_n} \phi_n(2)\bar{\phi}_n(1) \times \exp(-iE_n(t_2 - t_1)) \quad \text{for} \quad t_2 < t_1. \end{aligned} \tag{17}$$

With this choice of K_+ our equations such as (13) and (14) will now give results equivalent to those of the positron hole theory.

That (14), for example, is the correct second order expression for finding at 2 an electron originally at 1 according to the positron theory may be seen as follows (Fig. 2). Assume as a special example that $t_2 > t_1$ and that the potential vanishes except in interval $t_2 - t_1$ so that t_4 and t_3 both lie between t_1 and t_2.

First suppose $t_4 > t_3$ (Fig. 2(b)). Then (since $t_3 > t_1$)

the electron assumed originally in a positive energy state propagates in that state (by $K_+(3, 1)$) to position 3 where it gets scattered ($A(3)$). It then proceeds to 4, which it must do as a positive energy electron. This is correctly described by (14) for $K_+(4, 3)$ contains only positive energy components in its expansion, as $t_4 > t_3$. After being scattered at 4 it then proceeds on to 2, again necessarily in a positive energy state, as $t_2 > t_4$.

In positron theory there is an additional contribution due to the possibility of virtual pair production (Fig. 2(c)). A pair could be created by the potential $A(4)$ at 4, the electron of which is that found later at 2. The positron (or rather, the hole) proceeds to 3 where it annihilates the electron which has arrived there from 1.

This alternative is already included in (14) as contributions for which $t_4 < t_3$, and its study will lead us to an interpretation of $K_+(4, 3)$ for $t_4 < t_3$. The factor $K_+(2, 4)$ describes the electron (after the pair production at 4) proceeding from 4 to 2. Likewise $K_+(3, 1)$ represents the electron proceeding from 1 to 3. $K_+(4, 3)$ must therefore represent the propagation of the positron or hole from 4 to 3. That it does so is clear. The fact that in hole theory the hole proceeds in the manner of and electron of negative energy is reflected in the fact that $K_+(4, 3)$ for $t_4 < t_3$ is (minus) the sum of only negative energy components. In hole theory the real energy of these intermediate states is, of course, positive. This is true here too, since in the phases $\exp(-iE_n(t_4 - t_3))$ defining $K_+(4, 3)$ in (17), E_n is negative but so is $t_4 - t_3$. That is, the contributions vary with t_3 as $\exp(-i|E_n|(t_3 - t_4))$ as they would if the energy of the intermediate state were $|E_n|$. The fact that the entire sum is taken as negative in computing $K_+(4, 3)$ is reflected in the fact that in hole theory the amplitude has its sign reversed in accordance with the Pauli principle and the fact that the electron arriving at 2 has been exchanged with one in the sea.[6] To this, and to higher orders, all processes involving virtual pairs are correctly described in this way.

The expressions such as (14) can still be described as a passage of the electron from 1 to 3 ($K_+(3, 1)$), scattering at 3 by $A(3)$, proceeding to 4 ($K_+(4, 3)$), scattering again, $A(4)$, arriving finally at 2. The scatterings may, however, be toward both future and past times, an electron propagating backwards in time being recognized as a positron.

This therefore suggests that negative energy components created by scattering in a potential be considered as waves propagating from the scattering point toward the past, and that such waves represent the propagation of a positron annihilating the electron in the potential.[7]

With this interpretation real pair production is also described correctly (see Fig. 3). For example in (13) if $t_1 < t_3 < t_2$ the equation gives the amplitude that if at time t_1 one electron is present at 1, then at time t_2 just one electron will be present (having been scattered at 3) and it will be at 2. On the other hand if t_2 is less than t_3, for example, if $t_2 = t_1 < t_3$, the same expression gives the amplitude that a pair, electron at 1, positron at 2 will annihilate at 3, and subsequently no particles will be present. Likewise if t_2 and t_1 exceed t_3 we have (minus) the amplitude for finding a single pair, electron at 2, positron at 1 created by $A(3)$ from a vacuum. If $t_1 > t_3 > t_2$, (13) describes the scattering of a positron. All these amplitudes are relative to the amplitude that a vacuum will remain a vacuum, which is taken as unity. (This will be discussed more fully later.)

The analogue of (2) can be easily worked out.[8] It is,

$$\psi(2) = \int K_+(2, 1) N(1) \psi(1) d^3V_1, \tag{18}$$

where d^3V_1 is the volume element of the closed 3-dimensional surface of a region of space time containing

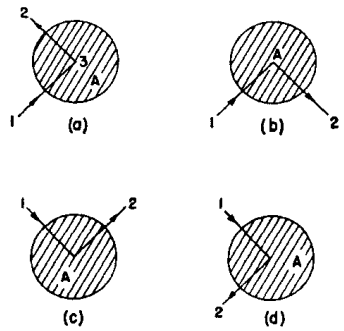


FIG. 3. Several different processes can be described by the same formula depending on the time relations of the variables t_2, t_1. Thus $P_v|K_+^{(A)}(2, 1)|^2$ is the probability that: (a) An electron at 1 will be scattered at 2 (and no other pairs form in vacuum). (b) Electron at 1 and positron at 2 annihilate leaving nothing. (c) A single pair at 1 and 2 is created from vacuum. (d) A positron at 2 is scattered to 1. ($K_+^{(A)}(2, 1)$ is the sum of the effects of scattering in the potential to all orders. P_v is a normalizing constant.)

[6] It has often been noted that the one-electron theory apparently gives the same matrix elements for this process as does hole theory. The problem is one of interpretation, especially in a way that will also give correct results for other processes, e.g., self-energy.

[7] The idea that positrons can be represented as electrons with proper time reversed relative to true time has been discussed by the author and others, particularly by Stückelberg. E. C. C. Stückelberg, Helv. Phys. Acta **15**, 23 (1942); R. P. Feynman, Phys. Rev. **74**, 939 (1948). The fact that classically the action (proper time) increases continuously as one follows a trajectory is reflected in quantum mechanics in the fact that the phase, which is $|E_n||t_2 - t_1|$, always increases as the particle proceeds from one scattering point to the next.

[8] By multiplying (12) on the right by $(-i\nabla_1 - m)$ and noting that $\nabla_1\delta(2, 1) = -\nabla_2\delta(2, 1)$ show that $K_+(2, 1)$ also satisfies $K_+(2, 1)(-i\nabla_1 - m) = i\delta(2, 1)$, where the ∇_1 operates on variable 1 in $K_+(2, 1)$ but is written after that function to keep the correct order of the γ matrices. Multiply this equation by $\psi(1)$ and Eq. (11) (with $A = 0$, calling the variables 1) by $K_+(2, 1)$, subtract and integrate over a region of space-time. The integral on the left-hand side can be transformed to an integral over the surface of the region. The right-hand side is $\psi(2)$ if the point 2 lies within the region, and is zero otherwise. (What happens when the 3-surface contains a light line and hence has no unique normal need not concern us as these points can be made to occur so far away from 2 that their contribution vanishes.)

point 2, and $N(1)$ is $N_\mu(1)\gamma_\mu$ where $N_\mu(1)$ is the *inward* drawn unit normal to the surface at the point 1. That is, the wave function $\psi(2)$ (in this case for a free particle) is determined at any point inside a four-dimensional region if its values on the surface of that region are specified.

To interpret this, consider the case that the 3-surface consists essentially of all space at some time say $t=0$ previous to t_2, and of all space at the time $T>t_2$. The cylinder connecting these to complete the closure of the surface may be very distant from $\mathbf{x}_2$ so that it gives no appreciable contribution (as $K_+(2,1)$ decreases exponentially in space-like directions). Hence, if $\gamma_4=\beta$, since the inward drawn normals N will be β and $-\beta$,

$$\psi(2)=\int K_+(2,1)\beta\psi(1)d^3\mathbf{x}_1 - \int K_+(2,1')\beta\psi(1')d^3\mathbf{x}_{1'}, \quad (19)$$

where $t_1=0$, $t_{1'}=T$. Only positive energy (electron) components in $\psi(1)$ contribute to the first integral and only negative energy (positron) components of $\psi(1')$ to the second. That is, the amplitude for finding a charge at 2 is determined both by the amplitude for finding an electron previous to the measurement and by the amplitude for finding a positron after the measurement. This might be interpreted as meaning that even in a problem involving but one charge the amplitude for finding the charge at 2 is not determined when the only thing known in the amplitude for finding an electron (or a positron) at an earlier time. There may have been no electron present initially but a pair was created in the measurement (or also by other external fields). The amplitude for this contingency is specified by the amplitude for finding a positron in the future.

We can also obtain expressions for transition amplitudes, like (5). For example if at $t=0$ we have an electron present in a state with (positive energy) wave function $f(\mathbf{x})$, what is the amplitude for finding it at $t=T$ with the (positive energy) wave function $g(\mathbf{x})$? The amplitude for finding the electron anywhere after $t=0$ is given by (19) with $\psi(1)$ replaced by $f(\mathbf{x})$, the second integral vanishing. Hence, the transition element to find it in state $g(\mathbf{x})$ is, in analogy to (5), just ($t_2=T$, $t_1=0$)

$$\int \bar{g}(\mathbf{x}_2)\beta K_+(2,1)\beta f(\mathbf{x}_1)d^3\mathbf{x}_1 d^3\mathbf{x}_2, \quad (20)$$

since $g^*=\bar{g}\beta$.

If a potential acts somewhere in the interval between 0 and T, K_+ is replaced by $K_+^{(A)}$. Thus the first order effect on the transition amplitude is, from (13),

$$-i\int \bar{g}(\mathbf{x}_2)\beta K_+(2,3)A(3)K_+(3,1)\beta f(\mathbf{x}_1)d^3\mathbf{x}_1 d^3\mathbf{x}_2. \quad (21)$$

Expressions such as this can be simplified and the 3-surface integrals, which are inconvenient for relativistic calculations, can be removed as follows. Instead of defining a state by the wave function $f(\mathbf{x})$, which it has at a given time $t_1=0$, we define the state by the function $f(1)$ of four variables $\mathbf{x}_1$, t_1 which is a solution of the free particle equation for all t_1 and is $f(\mathbf{x}_1)$ for $t_1=0$. The final state is likewise defined by a function $g(2)$ over-all space-time. Then our surface integrals can be performed since $\int K_+(3,1)\beta f(\mathbf{x}_1)d^3\mathbf{x}_1=f(3)$ and $\int \bar{g}(\mathbf{x}_2)\beta d^3\mathbf{x}_2 K_+(2,3)=\bar{g}(3)$. There results

$$-i\int \bar{g}(3)A(3)f(3)d\tau_3, \quad (22)$$

the integral now being over-all space-time. The transition amplitude to second order (from (14)) is

$$-\int\int \bar{g}(2)A(2)K_+(2,1)A(1)f(1)d\tau_1 d\tau_2, \quad (23)$$

for the particle arriving at 1 with amplitude $f(1)$ is scattered ($A(1)$), progresses to 2, ($K_+(2,1)$), and is scattered again ($A(2)$), and we then ask for the amplitude that it is in state $g(2)$. If $g(2)$ is a negative energy state we are solving a problem of annihilation of electron in $f(1)$, positron in $g(2)$, etc.

We have been emphasizing scattering problems, but obviously the motion in a fixed potential V, say in a hydrogen atom, can also be dealt with. If it is first viewed as a scattering problem we can ask for the amplitude, $\phi_k(1)$, that an electron with original free wave function was scattered k times in the potential V either forward or backward in time to arrive at 1. Then the amplitude after one more scattering is

$$\phi_{k+1}(2)=-i\int K_+(2,1)V(1)\phi_k(1)d\tau_1. \quad (24)$$

An equation for the total amplitude

$$\psi(1)=\sum_{k=0}^{\infty}\phi_k(1)$$

for arriving at 1 either directly or after any number of scatterings is obtained by summing (24) over all k from 0 to ∞;

$$\psi(2)=\phi_0(2)-i\int K_+(2,1)V(1)\psi(1)d\tau_1. \quad (25)$$

Viewed as a steady state problem we may wish, for example, to find that initial condition ϕ_0 (or better just the ψ) which leads to a periodic motion of ψ. This is most practically done, of course, by solving the Dirac equation,

$$(i\nabla-m)\psi(1)=V(1)\psi(1), \quad (26)$$

deduced from (25) by operating on both sides by $i\nabla_2-m$, thereby eliminating the ϕ_0, and using (12). This illustrates the relation between the points of view.

For many problems the total potential $A+V$ may be split conveniently into a fixed one, V, and another, A, considered as a perturbation. If $K_+^{(V)}$ is defined as in

(16) with V for A, expressions such as (23) are valid and useful with K_+ replaced by $K_+^{(V)}$ and the functions $f(1)$, $g(2)$ replaced by solutions for all space and time of the Dirac Eq. (26) in the potential V (rather than free particle wave functions).

4. PROBLEMS INVOLVING SEVERAL CHARGES

We wish next to consider the case that there are two (or more) distinct charges (in addition to pairs they may produce in virtual states). In a succeeding paper we discuss the interaction between such charges. Here we assume that they do not interact. In this case each particle behaves independently of the other. We can expect that if we have two particles a and b, the amplitude that particle a goes from $\mathbf{x}_1$ at t_1, to $\mathbf{x}_3$ at t_3 while b goes from $\mathbf{x}_2$ at t_2 to $\mathbf{x}_4$ at t_4 is the product

$$K(3,4;1,2)=K_{+a}(3,1)K_{+b}(4,2).$$

The symbols a, b simply indicate that the matrices appearing in the K_+ apply to the Dirac four component spinors corresponding to particle a or b respectively (the wave function now having 16 indices). In a potential K_{+a} and K_{+b} become $K_{+a}^{(A)}$ and $K_{+b}^{(A)}$ where $K_{+a}^{(A)}$ is defined and calculated as for a single particle. They commute. Hereafter the a, b can be omitted; the space time variable appearing in the kernels suffice to define on what they operate.

The particles are identical however and satisfy the exclusion principle. The principle requires only that one calculate $K(3,4;1,2)-K(4,3;1,2)$ to get the net amplitude for arrival of charges at 3, 4. (It is normalized assuming that when an integral is performed over points 3 and 4, for example, since the electrons represented are identical, one divides by 2.) This expression is correct for positrons also (Fig. 4). For example the amplitude that an electron and a positron found initially at $\mathbf{x}_1$ and $\mathbf{x}_4$ (say $t_1=t_4$) are later found at $\mathbf{x}_3$ and $\mathbf{x}_2$ (with $t_2=t_3>t_1$) is given by the same expression

$$K_+^{(A)}(3,1)K_+^{(A)}(4,2)-K_+^{(A)}(4,1)K_+^{(A)}(3,2). \quad (27)$$

The first term represents the amplitude that the electron proceeds from 1 to 3 and the positron from 4 to 2 (Fig. 4(c)), while the second term represents the interfering amplitude that the pair at 1, 4 annihilate and what is found at 3, 2 is a pair newly created in the potential. The generalization to several particles is clear. There is an additional factor $K_+^{(A)}$ for each particle, and antisymmetric combinations are always taken.

No account need be taken of the exclusion principle in intermediate states. As an example consider again expression (14) for $t_2>t_1$ and suppose $t_4<t_3$ so that the situation represented (Fig. 2(c)) is that a pair is made at 4 with the electron proceeding to 2, and the positron to 3 where it annihilates the electron arriving from 1. It may be objected that if it happens that the electron created at 4 is in the same state as the one coming from 1, then the process cannot occur because of the exclusion principle and we should not have included it in our term (14). We shall see, however, that considering the exclusion principle also requires another change which reinstates the quantity.

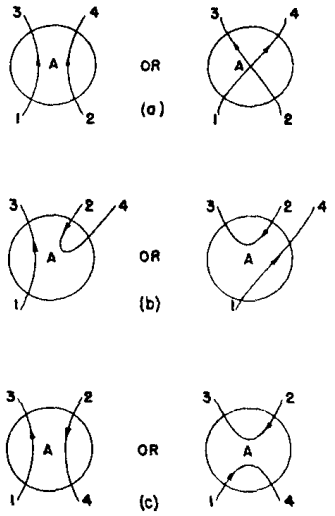


FIG. 4. Some problems involving two distinct charges (in addition to virtual pairs they may produce): $P_v|K_+^{(A)}(3,1)K_+^{(A)}(4,2)-K_+^{(A)}(4,1)K_+^{(A)}(3,2)|^2$ is the probability that: (a) Electrons at 1 and 2 are scattered to 3, 4 (and no pairs are formed). (b) Starting with an electron at 1 a single pair is formed, positron at 2, electrons at 3, 4. (c) A pair at 1, 4 is found at 3, 2, etc. The exclusion principle requires that the amplitudes for processes involving exchange of two electrons be subtracted.

For we are computing amplitudes relative to the amplitude that a vacuum at t_1 will still be a vacuum at t_2. We are interested in the alteration in this amplitude due to the presence of an electron at 1. Now one process that can be visualized as occurring in the vacuum is the creation of a pair at 4 followed by a re-annihilation of the *same* pair at 3 (a process which we shall call a closed loop path). But if a real electron is present in a certain state 1, those pairs for which the electron was created in state 1 in the vacuum must now be excluded. We must therefore subtract from our relative amplitude the term corresponding to this process. But this just reinstates the quantity which it was argued should not have been included in (14), the necessary minus sign coming automatically from the definition of K_+. It is obviously simpler to disregard the exclusion principle completely in the intermediate states.

All the amplitudes are relative and their squares give the relative probabilities of the various phenomena. Absolute probabilities result if one multiplies each of the probabilities by P_v, the true probability that if one has no particles present initially there will be none finally. This quantity P_v can be calculated by normalizing the relative probabilities such that the sum of the probabilities of all mutually exclusive alternatives is unity. (For example if one starts with a vacuum one can calculate the relative probability that there remains a

vacuum (unity), or one pair is created, or two pairs, etc. The sum is P_v^{-1}.) Put in this form the theory is complete and there are no divergence problems. Real processes are completely independent of what goes on in the vacuum.

When we come, in the succeeding paper, to deal with interactions between charges, however, the situation is not so simple. There is the possibility that virtual electrons in the vacuum may interact electromagnetically with the real electrons. For that reason processes occuring in the vacuum are analyzed in the next section, in which an independent method of obtaining P_v is discussed.

5. VACUUM PROBLEMS

An alternative way of obtaining absolute amplitudes is to multiply all amplitudes by C_v, the vacuum to vacuum amplitude, that is, the absolute amplitude that there be no particles both initially and finally. We can assume $C_v=1$ if no potential is present during the interval, and otherwise we compute it as follows. It differs from unity because, for example, a pair could be created which eventually annihilates itself again. Such a path would appear as a closed loop on a space-time diagram. The sum of the amplitudes resulting from all such single closed loops we call L. To a first approximation L is

$$L^{(1)}=-\frac{1}{2}\int\int Sp[K_+(2,1)A(1)\times K_+(1,2)A(2)]d\tau_1 d\tau_2. \quad (28)$$

For a pair could be created say at 1, the electron and positron could both go on to 2 and there annihilate. The spur, Sp, is taken since one has to sum over all possible spins for the pair. The factor $\frac{1}{2}$ arises from the fact that the same loop could be considered as starting at either potential, and the minus sign results since the interactors are each $-iA$. The next order term would be[9]

$$L^{(2)}=+(i/3)\int\int\int Sp[K_+(2,1)A(1)\times K_+(1,3)A(3)K_+(3,2)A(2)]d\tau_1 d\tau_2 d\tau_3,$$

etc. The sum of all such terms gives L.[10]

In addition to these single loops we have the possibility that two independent pairs may be created and each pair may annihilate itself again. That is, there may be formed in the vacuum two closed loops, and the contribution in amplitude from this alternative is just the product of the contribution from each of the loops considered singly. The total contribution from all such pairs of loops (it is still consistent to disregard the exclusion principle for these virtual states) is $L^2/2$ for in L^2 we count every pair of loops twice. The total vacuum-vacuum amplitude is then

$$C_v=1-L+L^2/2-L^3/6+\cdots=\exp(-L), \quad (30)$$

the successive terms representing the amplitude from zero, one, two, etc., loops. The fact that the contribution to C_v of single loops is $-L$ is a consequence of the Pauli principle. For example, consider a situation in which two pairs of particles are created. Then these pairs later destroy themselves so that we have two loops. The electrons could, at a given time, be interchanged forming a kind of figure eight which is a single loop. The fact that the interchange must change the sign of the contribution requires that the terms in C_v appear with alternate signs. (The exclusion principle is also responsible in a similar way for the fact that the amplitude for a pair creation is $-K_+$ rather than $+K_+$.) Symmetrical statistics would lead to

$$C_v=1+L+L^2/2=\exp(+L).$$

The quantity L has an infinite imaginary part (from $L^{(1)}$, higher orders are finite). We will discuss this in connection with vacuum polarization in the succeeding paper. This has no effect on the normalization constant for the probability that a vacuum remain vacuum is given by

$$P_v=|C_v|^2=\exp(-2\cdot\text{real part of } L),$$

from (30). This value agrees with the one calculated directly by renormalizing probabilities. The real part of L appears to be positive as a consequence of the Dirac equation and properties of K_+ so that P_v is less than one. Bose statistics gives $C_v=\exp(+L)$ and consequently a value of P_v greater than unity which appears meaningless if the quantities are interpreted as we have done here. Our choice of K_+ apparently requires the exclusion principle.

Charges obeying the Klein-Gordon equation can be equally well treated by the methods which are discussed here for the Dirac electrons. How this is done is discussed in more detail in the succeeding paper. The real part of L comes out negative for this equation so that in this case Bose statistics appear to be required for consistency.[3]

[9] This term actually vanishes as can be seen as follows. In any spur the sign of all γ matrices may be reversed. Reversing the sign of γ in $K_+(2,1)$ changes it to the transpose of $K_+(1,2)$ so that the order of all factors and variables is reversed. Since the integral is taken over all τ_1, τ_2, and τ_3 this has no effect and we are left with $(-1)^3$ from changing the sign of A. Thus the spur equals its negative. Loops with an odd number of potential interactors give zero. Physically this is because for each loop the electron can go around one way or in the opposite direction and we must add these amplitudes. But reversing the motion of an electron makes it behave like a positive charge thus changing the sign of each potential interaction, so that the sum is zero if the number of interactions is odd. This theorem is due to W. H. Furry, Phys. Rev. 51, 125 (1937).

[10] A closed expression for L in terms of $K_+^{(A)}$ is hard to obtain because of the factor $(1/n)$ in the nth term. However, the perturbation in L, ΔL due to a small change in potential ΔA, is easy to express. The $(1/n)$ is canceled by the fact that ΔA can appear in any of the n potentials. The result after summing over n by (13), (14) and using (16) is

$$\Delta L=-i\int Sp[(K_+^{(A)}(1,1)-K_+(1,1))\Delta A(1)]d\tau_1. \quad (29)$$

The term $K_+(1,1)$ actually integrates to zero.

6. ENERGY-MOMENTUM REPRESENTATION

The practical evaluation of the matrix elements in some problems is often simplified by working with momentum and energy variables rather than space and time. This is because the function $K_+(2,1)$ is fairly complicated but we shall find that its Fourier transform is very simple, namely $(i/4\pi^2)(\not{p}-m)^{-1}$ that is

$$K_+(2,1)=(i/4\pi^2)\int(\not{p}-m)^{-1}\exp(-ip\cdot x_{21})d^4p, \quad (31)$$

where $p\cdot x_{21}=p\cdot x_2-p\cdot x_1=p_\mu x_{2\mu}-p_\mu x_{1\mu}$, $\not{p}=p_\mu\gamma_\mu$, and d^4p means $(2\pi)^{-2}dp_1dp_2dp_3dp_4$, the integral over all p. That this is true can be seen immediately from (12), for the representation of the operator $i\nabla-m$ in energy (p_4) and momentum ($p_{1,2,3}$) space is $\not{p}-m$ and the transform of $\delta(2,1)$ is a constant. The reciprocal matrix $(\not{p}-m)^{-1}$ can be interpreted as $(\not{p}+m)(p^2-m^2)^{-1}$ for $\not{p}^2-m^2=(\not{p}-m)(\not{p}+m)$ is a pure number not involving γ matrices. Hence if one wishes one can write

$$K_+(2,1)=i(i\nabla_2+m)I_+(2,1),$$

where

$$I_+(2,1)=(2\pi)^{-2}\int(p^2-m^2)^{-1}\exp(-ip\cdot x_{21})d^4p, \quad (32)$$

is not a matrix operator but a function satisfying

$$\Box_2^2 I_+(2,1)-m^2I_+(2,1)=\delta(2,1), \quad (33)$$

where $-\Box_2^2=(\nabla_2)^2=(\partial/\partial x_{2\mu})(\partial/\partial x_{2\mu})$.

The integrals (31) and (32) are not yet completely defined for there are poles in the integrand when $p^2-m^2=0$. We can define how these poles are to be evaluated by the rule that *m is considered to have an infinitesimal negative imaginary part.* That is m, is replaced by $m-i\delta$ and the limit taken as $\delta\to 0$ from above. This can be seen by imagining that we calculate K_+ by integrating on p_4 first. If we call $E=+(m^2+p_1^2+p_2^2+p_3^2)^{\frac{1}{2}}$ then the integrals involve p_4 essentially as $\int\exp(-ip_4(t_2-t_1))dp_4(p_4^2-E^2)^{-1}$ which has poles at $p_4=+E$ and $p_4=-E$. The replacement of m by $m-i\delta$ means that E has a small negative imaginary part; the first pole is below, the second above the real axis. Now if $t_2-t_1>0$ the contour can be completed around the semicircle below the real axis thus giving a residue from the $p_4=+E$ pole, or $-(2E)^{-1}\exp(-iE(t_2-t_1))$. If $t_2-t_1<0$ the upper semicircle must be used, and $p_4=-E$ at the pole, so that the function varies in each case as required by the other definition (17).

Other solutions of (12) result from other prescriptions. For example if p_4 in the factor $(p^2-m^2)^{-1}$ is considered to have a positive imaginary part K_+ becomes replaced by K_0, the Dirac one-electron kernel, zero for $t_2<t_1$. Explicitly the function is[11] $(\mathbf{x}, t=x_{21\mu})$

$$I_+(\mathbf{x},t)=-(4\pi)^{-1}\delta(s^2)+(m/8\pi s)H_1^{(2)}(ms), \quad (34)$$

where $s=+(t^2-\mathbf{x}^2)^{\frac{1}{2}}$ for $t^2>\mathbf{x}^2$ and $s=-i(x^2-t^2)^{\frac{1}{2}}$ for $t^2<x$, $H_1^{(2)}$ is the Hankel function and $\delta(s^2)$ is the Dirac delta function of s^2. It behaves asymptotically as $\exp(-ims)$, decaying exponentially in space-like directions.[12]

By means of such transforms the matrix elements like (22), (23) are easily worked out. A free particle wave function for an electron of momentum $\not{p}_1$ is $u_1\exp(-ip_1\cdot x)$ where u_1 is a constant spinor satisfying the Dirac equation $\not{p}_1u_1=mu_1$ so that $p_1^2=m^2$. The matrix element (22) for going from a state $\not{p}_1$, u_1 to a state of momentum $\not{p}_2$, spinor u_2, is $-4\pi^2i(\bar{u}_2\not{a}(q)u_1)$ where we have imagined A expanded in a Fourier integral

$$A(1)=\int\not{a}(q)\exp(-iq\cdot x_1)d^4q,$$

and we select the component of momentum $q=p_2-p_1$.

The second order term (23) is the matrix element between u_1 and u_2 of

$$-4\pi^2i\int(\not{a}(\not{p}_2-\not{p}_1-\not{q}))(\not{p}_1+\not{q}-m)^{-1}\not{a}(q)d^4q, \quad (35)$$

since the electron of momentum $\not{p}_1$ may pick up q from the potential $\not{a}(q)$, propagate with momentum $\not{p}_1+\not{q}$ (factor $(\not{p}_1+\not{q}-m)^{-1}$) until it is scattered again by the potential, $\not{a}(\not{p}_2-\not{p}_1-\not{q})$, picking up the remaining momentum, $\not{p}_2-\not{p}_1-\not{q}$, to bring the total to $\not{p}_2$. Since all values of q are possible, one integrates over q.

These same matrices apply directly to positron problems, for if the time component of, say, $\not{p}_1$ is negative the state represents a positron of four-momentum $-\not{p}_1$, and we are describing pair production if $\not{p}_2$ is an electron, i.e., has positive time component, etc.

The probability of an event whose matrix element is $(\bar{u}_2Mu_1)$ is proportional to the absolute square. This may also be written $(\bar{u}_1\bar{M}u_2)(\bar{u}_2Mu_1)$, where $\bar{M}$ is M with the operators written in opposite order and explicit appearance of i changed to $-i(\bar{M}$ is β times the complex conjugate transpose of $\beta M)$. For many problems we are not concerned about the spin of the final state. Then we can sum the probability over the two u_2 corresponding to the two spin directions. This is not a complete set because $\not{p}_2$ has another eigenvalue, $-m$. To permit summing over all states we can insert the projection operator $(2m)^{-1}(\not{p}_2+m)$ and so obtain $(2m)^{-1}(\bar{u}_1\bar{M}(\not{p}_2+m)Mu_1)$ for the probability of transition from $\not{p}_1$, u_1, to $\not{p}_2$ with arbitrary spin. If the incident state is unpolarized we can sum on its spins too, and obtain

$$(2m)^{-2}Sp[(\not{p}_1+m)\bar{M}(\not{p}_2+m)M] \quad (36)$$

for (twice) the probability that an electron of arbitrary spin with momentum $\not{p}_1$ will make transition to $\not{p}_2$. The expressions are all valid for positrons when $\not{p}$'s with

[11] $I_+(\mathbf{x},t)$ is $(2i)^{-1}(D_1(\mathbf{x},t)-iD(\mathbf{x},t))$ where D_1 and D are the functions defined by W. Pauli, Rev. Mod. Phys. **13**, 203 (1941).

[12] If the $-i\delta$ is kept with m here too the function I_+ approaches zero for infinite positive and negative times. This may be useful in general analyses in avoiding complications from infinitely remote surfaces.

negative energies are inserted, and the situation interpreted in accordance with the timing relations discussed above. (We have used functions normalized to $(\bar{u}u)=1$ instead of the conventional $(\bar{u}\beta u)=(u^*u)=1$. On our scale $(\bar{u}\beta u)=$energy$/m$ so the probabilities must be corrected by the appropriate factors.)

The author has many people to thank for fruitful conversations about this subject, particularly H. A. Bethe and F. J. Dyson.

APPENDIX

a. Deduction from Second Quantization

In this section we shall show the equivalence of this theory with the hole theory of the positron.[2] According to the theory of second quantization of the electron field in a given potential,[13] the state of this field at any time is represented by a wave function χ satisfying

$$i\partial\chi/\partial t=H\chi,$$

where $H=\int\Psi^*(\mathbf{x})(\boldsymbol{\alpha}\cdot(-i\boldsymbol{\nabla}-\mathbf{A})+A_4+m\beta)\Psi(\mathbf{x})d^3\mathbf{x}$ and $\Psi(\mathbf{x})$ is an operator annihilating an electron at position $\mathbf{x}$, while $\Psi^*(\mathbf{x})$ is the corresponding creation operator. We contemplate a situation in which at $t=0$ we have present some electrons in states represented by ordinary spinor functions $f_1(\mathbf{x})$, $f_2(\mathbf{x})$, $\cdots$ assumed orthogonal, and some positrons. These are described as holes in the negative energy sea, the electrons which would normally fill the holes having wave functions $p_1(\mathbf{x})$, $p_2(\mathbf{x})$, $\cdots$. We ask, at time T what is the amplitude that we find electrons in states $g_1(\mathbf{x})$, $g_2(\mathbf{x})$, $\cdots$ and holes at $q_1(\mathbf{x})$, $q_2(\mathbf{x})$, $\cdots$. If the initial and final state vectors representing this situation are χ_i and χ_f respectively, we wish to calculate the matrix element

$$R=\left(\chi_f^*\exp\left(-i\int_0^T H dt\right)\chi_i\right)=(\chi_f^* S\chi_i). \tag{37}$$

We assume that the potential A differs from zero only for times between 0 and T so that a vacuum can be defined at these times. If χ_0 represents the vacuum state (that is, all negative energy states filled, all positive energies empty), the amplitude for having a vacuum at time T, if we had one at $t=0$, is

$$C_v=(\chi_0^*S\chi_0), \tag{38}$$

writing S for $\exp(-i\int_0^T Hdt)$. Our problem is to evaluate R and show that it is a simple factor times C_v, and that the factor involves the $K_+^{(A)}$ functions in the way discussed in the previous sections.

To do this we first express χ_i in terms of χ_0. The operator

$$\Phi^*=\int\Psi^*(\mathbf{x})\phi(\mathbf{x})d^3\mathbf{x}, \tag{39}$$

creates an electron with wave function $\phi(\mathbf{x})$. Likewise $\Phi=\int\phi^*(\mathbf{x})\times\Psi(\mathbf{x})d^3\mathbf{x}$ annihilates one with wave function $\phi(\mathbf{x})$. Hence state χ_i is $\chi_i=F_1^*F_2^*\cdots P_1P_2\cdots\chi_0$ while the final state is $G_1^*G_2^*\cdots\times Q_1Q_2\cdots\chi_0$ where F_i, G_i, P_i, Q_i are operators defined like Φ, in (39), but with f_i, g_i, p_i, q_i replacing ϕ; for the initial state would result from the vacuum if we created the electrons in f_1, f_2, $\cdots$ and annihilated those in p_1, p_2, $\cdots$. Hence we must find

$$R=(\chi_0^*\cdots Q_2^*Q_1^*\cdots G_2G_1SF_1^*F_2^*\cdots P_1P_2\cdots\chi_0). \tag{40}$$

To simplify this we shall have to use commutation relations between a Φ^* operator and S. To this end consider $\exp(-i\int_0^t Hdt')\Phi^*\times\exp(+i\int_0^t Hdt')$ and expand this quantity in terms of $\Psi^*(\mathbf{x})$, giving $\int\Psi^*(\mathbf{x})\phi(\mathbf{x},t)d^3\mathbf{x}$, (which defines $\phi(\mathbf{x},t)$). Now multiply this equation by $\exp(+i\int_0^t Hdt')\cdots\exp(-i\int_0^t Hdt')$ and find

$$\int\Psi^*(\mathbf{x})\phi(\mathbf{x})d^3\mathbf{x}=\int\Psi^*(\mathbf{x},t)\phi(\mathbf{x},t)d^3\mathbf{x}, \tag{41}$$

where we have defined $\Psi(\mathbf{x},t)$ by $\Psi(\mathbf{x},t)=\exp(+i\int_0^t Hdt')\Psi(\mathbf{x})\times\exp(-i\int_0^t Hdt')$. As is well known $\Psi(\mathbf{x},t)$ satisfies the Dirac equation, (differentiate $\Psi(\mathbf{x},t)$ with respect to t and use commutation relations of H and Ψ)

$$i\partial\Psi(\mathbf{x},t)/\partial t=(\boldsymbol{\alpha}\cdot(-i\boldsymbol{\nabla}-\mathbf{A})+A_4+m\beta)\Psi(\mathbf{x},t). \tag{42}$$

Consequently $\phi(\mathbf{x},t)$ must also satisfy the Dirac equation (differentiate (41) with respect to t, use (42) and integrate by parts).

That is, if $\phi(\mathbf{x},T)$ is that solution of the Dirac equation at time T which is $\phi(\mathbf{x})$ at $t=0$, and if we define $\Phi^*=\int\Psi^*(\mathbf{x})\phi(\mathbf{x})d^3\mathbf{x}$ and $\Phi'^*=\int\Psi^*(\mathbf{x})\phi(\mathbf{x},T)d^3\mathbf{x}$ then $\Phi'^*=S\Phi^*S^{-1}$, or

$$S\Phi^*=\Phi'^*S. \tag{43}$$

The principle on which the proof will be based can now be illustrated by a simple example. Suppose we have just one electron initially and finally and ask for

$$r=(\chi_0^*GSF^*\chi_0). \tag{44}$$

We might try putting F^* through the operator S using (43), $SF^*=F'^*S$, where f' in $F'^*=\int\Psi^*(\mathbf{x})f'(\mathbf{x})d^3\mathbf{x}$ is the wave function at T arising from $f(\mathbf{x})$ at 0. Then

$$r=(\chi_0^*GF'^*S\chi_0)=\int g^*(\mathbf{x})f'(\mathbf{x})d^3\mathbf{x}\cdot C_v-(\chi_0^*F'^*GS\chi_0), \tag{45}$$

where the second expression has been obtained by use of the definition (38) of C_v and the general commutation relation

$$GF^*+F^*G=\int g^*(\mathbf{x})f(\mathbf{x})d^3\mathbf{x},$$

which is a consequence of the properties of $\Psi(\mathbf{x})$ (the others are $FG=-GF$ and $F^*G^*=-G^*F^*$). Now $\chi_0^*F'^*$ in the last term in (45) is the complex conjugate of $F'\chi_0$. Thus if f' contained only positive energy components, $F'\chi_0$ would vanish and we would have reduced r to a factor times C_v. But F', as worked out here, does contain negative energy components created in the potential A and the method must be slightly modified.

Before putting F^* through the operator we shall add to it another operator F''^* arising from a function $f''(x)$ containing *only negative* energy components and so chosen that the resulting f' has *only positive* ones. That is we want

$$S(F_{\text{pos}}^*+F_{\text{neg}}''^*)=F_{\text{pos}}'^*S, \tag{46}$$

where the "pos" and "neg" serve as reminders of the sign of the energy components contained in the operators. This we can now use in the form

$$SF_{\text{pos}}^*=F_{\text{pos}}'^*S-SF_{\text{neg}}''^*. \tag{47}$$

In our one electron problem this substitution replaces r by two terms

$$r=(\chi_0^*GF_{\text{pos}}'^*S\chi_0)-(\chi_0^*GSF_{\text{neg}}''^*\chi_0).$$

The first of these reduces to

$$r=\int g^*(\mathbf{x})f_{\text{pos}}'(\mathbf{x})d^3\mathbf{x}\cdot C_v,$$

as above, for $F_{\text{pos}}'\chi_0$ is now zero, while the second is zero since the creation operator $F_{\text{neg}}''^*$ gives zero when acting on the vacuum state as all negative energies are full. This is the central idea of the demonstration.

The problem presented by (46) is this: Given a function $f_{\text{pos}}(\mathbf{x})$ at time 0, to find the amount, f_{neg}'', of negative energy component which must be added in order that the solution of Dirac's equation at time T will have only positive energy components, f_{pos}'. This is a boundary value problem for which the kernel $K_+^{(A)}$ is designed. We know the positive energy components initially, f_{pos}, and the negative ones finally (zero). The positive ones finally are therefore (using (19))

$$f_{\text{pos}}'(\mathbf{x}_2)=\int K_+^{(A)}(2,1)\beta f_{\text{pos}}(\mathbf{x}_1)d^3\mathbf{x}_1, \tag{48}$$

where $t_2=T$, $t_1=0$. Similarly, the negative ones initially are

$$f_{\text{neg}}''(\mathbf{x}_2)=\int K_+^{(A)}(2,1)\beta f_{\text{pos}}(\mathbf{x}_1)d^3\mathbf{x}_1-f_{\text{pos}}(\mathbf{x}_2), \tag{49}$$

where t_2 approaches zero from above, and $t_1=0$. The $f_{\text{pos}}(\mathbf{x}_2)$ is

[13] See, for example, G. Wentzel, *Einfuhrung in die Quantentheorie der Wellenfelder* (Franz Deuticke, Leipzig, 1943), Chapter V.

subtracted to keep in $f_{neg}''(\mathbf{x}_2)$ only those waves which return from the potential and not those arriving directly at t_2 from the $K_+(2,1)$ part of $K_+^{(A)}(2,1)$, as $t_2 \to 0$. We could also have written

$$f_{neg}''(\mathbf{x}_2) = \int [K_+^{(A)}(2,1) - K_+(2,1)]\beta f_{pos}(\mathbf{x}_1) d^3\mathbf{x}_1. \quad (50)$$

Therefore the one-electron problem, $r = \int g^*(\mathbf{x}) f_{pos}'(\mathbf{x}) d^3\mathbf{x} \cdot C_v$, gives by (48)

$$r = C_v \int g^*(\mathbf{x}_2) K_+^{(A)}(2,1) \beta f(\mathbf{x}_1) d^3\mathbf{x}_1 d^3\mathbf{x}_2,$$

as expected in accordance with the reasoning of the previous sections (i.e., (20) with $K_+^{(A)}$ replacing K_+).

The proof is readily extended to the more general expression R, (40), which can be analyzed by induction. First one replaces F_1^* by a relation such as (47) obtaining two terms

$$R = (\chi_0^* \cdots Q_2^* Q_1^* \cdots G_2 G_1 F_{1pos}'^* S F_2^* \cdots P_1 P_2 \cdots \chi_0) - (\chi_0^* \cdots Q_2^* Q_1^* \cdots G_2 G_1 S F_{1neg}''^* F_2^* \cdots P_1 P_2 \cdots \chi_0).$$

In the first term the order of $F_{1pos}'^*$ and G_1 is then interchanged, producing an additional term $\int g_1^*(\mathbf{x}) f_{1pos}'(\mathbf{x}) d^3\mathbf{x}$ times an expression with one less electron in initial and final state. Next it is exchanged with G_2 producing an addition $-\int g_2^*(\mathbf{x}) f_{1pos}'(\mathbf{x}) d^3\mathbf{x}$ times a similar term, etc. Finally on reaching the Q_1^* with which it anticommutes it can be simply moved over to juxtaposition with χ_0^* where it gives zero. The second term is similarly handled by moving $F_{1neg}''^*$ through anti commuting F_2^*, etc., until it reaches P_1. Then it is exchanged with P_1 to produce an additional simpler term with a factor $\mp \int p_1^*(\mathbf{x}) f_{1neg}''(\mathbf{x}) d^3\mathbf{x}$ or $\mp \int p_1^*(\mathbf{x}_2) K_+^{(A)}(2,1) \beta f_1(\mathbf{x}_1) d^3\mathbf{x}_1 d^3\mathbf{x}_2$ from (49), with $t_2 = t_1 = 0$ (the extra $f_1(\mathbf{x}_2)$ in (49) gives zero as it is orthogonal to $p_1(\mathbf{x}_2)$). This describes in the expected manner the annihilation of the pair, electron f_1, positron p_1. The $F_{neg}''^*$ is moved in this way successively through the P's until it gives zero when acting on χ_0. Thus R is reduced, with the expected factors (and with alternating signs as required by the exclusion principle), to simpler terms containing two less operators which may in turn be further reduced by using F_1^* in a similar manner, etc. After all the F^* are used the Q^*'s can be reduced in a similar manner. They are moved through the S in the opposite direction in such a manner as to produce a purely negative energy operator at time 0, using relations analogous to (46) to (49). After all this is done we are left simply with the expected factor times C_v (assuming the net charge is the same in initial and final state.)

In this way we have written the solution to the general problem of the motion of electrons in given potentials. The factor C_v is obtained by normalization. However for photon fields it is desirable to have an explicit form for C_v in terms of the potentials. This is given by (30) and (29) and it is readily demonstrated that this also is correct according to second quantization.

b. Analysis of the Vacuum Problem

We shall calculate C_v from second quantization by induction considering a series of problems each containing a potential distribution more nearly like the one we wish. Suppose we know C_v for a problem like the one we want and having the same potentials for time t between some t_0 and T, but having potential zero for times from 0 to t_0. Call this $C_v(t_0)$, the corresponding Hamiltonian H_{t_0} and the sum of contributions for all single loops, $L(t_0)$. Then for $t_0 = T$ we have zero potential at all times, no pairs can be produced, $L(T) = 0$ and $C_v(T) = 1$. For $t_0 = 0$ we have the complete problem, so that $C_v(0)$ is what is defined as C_v in (38). Generally we have,

$$C_v(t_0) = \left(\chi_0^* \exp\left(-i\int_0^T H_{t_0} dt\right)\chi_0\right) = \left(\chi_0^* \exp\left(-i\int_{t_0}^T H_{t_0} dt\right)\chi_0\right),$$

since H_{t_0} is identical to the constant vacuum Hamiltonian H_T for $t < t_0$ and χ_0 is an eigenfunction of H_T with an eigenvalue (energy of vacuum) which we can take as zero.

The value of $C_v(t_0 - \Delta t_0)$ arises from the Hamiltonian $H_{t_0 - \Delta t_0}$ which differs from H_{t_0} just by having an extra potential during the short interval Δt_0. Hence, to first order in Δt_0, we have

$$C_v(t_0 - \Delta t_0) = \left(\chi_0^* \exp\left(-i\int_{t_0 - \Delta t_0}^T H_{t_0 - \Delta t_0} dt\right)\chi_0\right) = \left(\chi_0^* \exp\left(-i\int_{t_0}^T H_{t_0} dt\right)\left[1 - i\Delta t_0 \int \Psi^*(\mathbf{x}) \times (-\boldsymbol{\alpha} \cdot \mathbf{A}(\mathbf{x}, t_0) + A_4(\mathbf{x}, t_0))\Psi(\mathbf{x}) d^3\mathbf{x}\right]\chi_0\right);$$

we therefore obtain for the derivative of C_v the expression

$$-dC_v(t_0)/dt_0 = -i\left(\chi_0^* \exp\left(-i\int_{t_0}^T H_{t_0} dt\right) \times \int \Psi^*(\mathbf{x}) \beta A(\mathbf{x}, t_0) \Psi(\mathbf{x}) d^3\mathbf{x} \chi_0\right), \quad (51)$$

which will be reduced to a simple factor times $C_v(t_0)$ by methods analogous to those used in reducing R. The operator Ψ can be imagined to be split into two pieces Ψ_{pos} and Ψ_{neg} operating on positive and negative energy states respectively. The Ψ_{pos} on χ_0 gives zero so we are left with two terms in the current density, $\Psi_{pos}^* \beta A \Psi_{neg}$ and $\Psi_{neg}^* \beta A \Psi_{neg}$. The latter $\Psi_{neg}^* \beta A \Psi_{neg}$ is just the expectation value of βA taken over all negative energy states (minus $\Psi_{neg} \beta A \Psi_{neg}^*$ which gives zero acting on χ_0). This is the effect of the vacuum expectation current of the electrons in the sea which we should have subtracted from our original Hamiltonian in the customary way.

The remaining term $\Psi_{pos}^* \beta A \Psi_{neg}$, or its equivalent $\Psi_{pos}^* \beta A \Psi$ can be considered as $\Psi^*(\mathbf{x}) \mathbf{f}_{pos}(\mathbf{x})$ where $\mathbf{f}_{pos}(\mathbf{x})$ is written for the positive energy component of the operator $\beta A \Psi(\mathbf{x})$. Now this operator, $\Psi^*(\mathbf{x}) \mathbf{f}_{pos}(\mathbf{x})$, or more precisely just the $\Psi^*(\mathbf{x})$ part of it, can be pushed through the $\exp(-i\int_{t_0}^T H dt)$ in a manner exactly analogous to (47) when f is a function. (An alternative derivation results from the consideration that the operator $\Psi(\mathbf{x}, t)$ which satisfies the Dirac equation also satisfies the linear integral equations which are equivalent to it.) That is, (51) can be written by (48), (50),

$$-dC_v(t_0)/dt_0 = -i\left(\chi_0^* \int\int \Psi^*(\mathbf{x}_2) K_+^{(A)}(2,1) \times \exp\left(-i\int_{t_0}^T H dt\right) A(1) \Psi(\mathbf{x}_1) d^3\mathbf{x}_1 d^3\mathbf{x}_2 \chi_0\right) + i\left(\chi_0^* \exp\left(-i\int_{t_0}^T H dt\right) \int\int \Psi^*(\mathbf{x}_2) [K_+^{(A)}(2,1) - K_+(2,1)] A(1) \Psi(\mathbf{x}_1) d^3\mathbf{x}_1 d^3\mathbf{x}_2 \chi_0\right),$$

where in the first term $t_2 = T$, and in the second $t_2 \to t_0 = t_1$. The (A) in $K_+^{(A)}$ refers to that part of the potential A after t_0. The first term vanishes for it involves (from the $K_+^{(A)}(2,1)$) only positive energy components of Ψ^*, which give zero operating into χ_0^*. In the second term only negative components of $\Psi^*(\mathbf{x}_2)$ appear. If, then $\Psi^*(\mathbf{x}_2)$ is interchanged in order with $\Psi(\mathbf{x}_1)$ it will give zero operating on χ_0, and only the term,

$$-dC_v(t_0)/dt_0 = +i \int Sp[(K_+^{(A)}(1,1) - K_+(1,1)) A(1)] d^3\mathbf{x}_1 \cdot C_v(t_0), \quad (52)$$

will remain, from the usual commutation relation of Ψ^* and Ψ.

The factor of $C_v(t_0)$ in (52) times $-\Delta t_0$ is, according to (29) (reference 10), just $L(t_0 - \Delta t_0) - L(t_0)$ since this difference arises from the extra potential $\Delta A = A$ during the short time interval Δt_0. Hence $-dC_v(t_0)/dt_0 = +(dL(t_0)/dt_0) C_v(t_0)$ so that integration from $t_0 = T$ to $t_0 = 0$ establishes (30).

Starting from the theory of the electromagnetic field in second quantization, a deduction of the equations for quantum electrodynamics which appear in the succeeding paper may be worked out using very similar principles. The Pauli-Weisskopf theory of the Klein-Gordon equation can apparently be analyzed in essentially the same way as that used here for Dirac electrons.

Space-Time Approach to Quantum Electrodynamics

R. P. FEYNMAN
Department of Physics, Cornell University, Ithaca, New York
(Received May 9, 1949)

In this paper two things are done. (1) It is shown that a considerable simplification can be attained in writing down matrix elements for complex processes in electrodynamics. Further, a physical point of view is available which permits them to be written down directly for any specific problem. Being simply a restatement of conventional electrodynamics, however, the matrix elements diverge for complex processes. (2) Electrodynamics is modified by altering the interaction of electrons at short distances. All matrix elements are now finite, with the exception of those relating to problems of vacuum polarization. The latter are evaluated in a manner suggested by Pauli and Bethe, which gives finite results for these matrices also. The only effects sensitive to the modification are changes in mass and charge of the electrons. Such changes could not be directly observed. Phenomena directly observable, are insensitive to the details of the modification used (except at extreme energies). For such phenomena, a limit can be taken as the range of the modification goes to zero. The results then agree with those of Schwinger. A complete, unambiguous, and presumably consistent, method is therefore available for the calculation of all processes involving electrons and photons.

The simplification in writing the expressions results from an emphasis on the over-all space-time view resulting from a study of the solution of the equations of electrodynamics. The relation of this to the more conventional Hamiltonian point of view is discussed. It would be very difficult to make the modification which is proposed if one insisted on having the equations in Hamiltonian form.

The methods apply as well to charges obeying the Klein-Gordon equation, and to the various meson theories of nuclear forces. Illustrative examples are given. Although a modification like that used in electrodynamics can make all matrices finite for all of the meson theories, for some of the theories it is no longer true that all directly observable phenomena are insensitive to the details of the modification used.

The actual evaluation of integrals appearing in the matrix elements may be facilitated, in the simpler cases, by methods described in the appendix.

THIS paper should be considered as a direct continuation of a preceding one[1] (I) in which the motion of electrons, neglecting interaction, was analyzed, by dealing directly with the *solution* of the Hamiltonian differential equations. Here the same technique is applied to include interactions and in that way to express in simple terms the solution of problems in quantum electrodynamics.

For most practical calculations in quantum electrodynamics the solution is ordinarily expressed in terms of a matrix element. The matrix is worked out as an expansion in powers of $e^2/\hbar c$, the successive terms corresponding to the inclusion of an increasing number of virtual quanta. It appears that a considerable simplification can be achieved in writing down these matrix elements for complex processes. Furthermore, each term in the expansion can be written down and understood directly from a physical point of view, similar to the space-time view in I. It is the purpose of this paper to describe how this may be done. We shall also discuss methods of handling the divergent integrals which appear in these matrix elements.

The simplification in the formulae results mainly from the fact that previous methods unnecessarily separated into individual terms processes that were closely related physically. For example, in the exchange of a quantum between two electrons there were two terms depending on which electron emitted and which absorbed the quantum. Yet, in the virtual states considered, timing relations are not significant. Olny the order of operators in the matrix must be maintained. We have seen (I), that in addition, processes in which virtual pairs are produced can be combined with others in which only positive energy electrons are involved. Further, the effects of longitudinal and transverse waves can be combined together. The separations previously made were on an unrelativistic basis (reflected in the circumstance that apparently momentum but not energy is conserved in intermediate states). When the terms are combined and simplified, the relativistic invariance of the result is self-evident.

We begin by discussing the solution in space and time of the Schrödinger equation for particles interacting instantaneously. The results are immediately generalizable to delayed interactions of relativistic electrons and we represent in that way the laws of quantum electrodynamics. We can then see how the matrix element for any process can be written down directly. In particular, the self-energy expression is written down.

So far, nothing has been done other than a restatement of conventional electrodynamics in other terms. Therefore, the self-energy diverges. A modification[2] in interaction between charges is next made, and it is shown that the self-energy is made convergent and corresponds to a correction to the electron mass. After the mass correction is made, other real processes are finite and insensitive to the "width" of the cut-off in the interaction.[3]

Unfortunately, the modification proposed is not completely satisfactory theoretically (it leads to some difficulties of conservation of energy). It does, however, seem consistent and satisfactory to define the matrix

[1] R. P. Feynman, Phys. Rev. 76, 749 (1949), hereafter called I.

[2] For a discussion of this modification in classical physics see R. P. Feynman, Phys. Rev. 74 939 (1948), hereafter referred to as A.

[3] A brief summary of the methods and results will be found in R. P. Feynman, Phys. Rev. 74, 1430 (1948), hereafter referred to as B.

element for all real processes as the limit of that computed here as the cut-off width goes to zero. A similar technique suggested by Pauli and by Bethe can be applied to problems of vacuum polarization (resulting in a renormalization of charge) but again a strict physical basis for the rules of convergence is not known.

After mass and charge renormalization, the limit of zero cut-off width can be taken for all real processes. The results are then equivalent to those of Schwinger[4] who does not make explicit use of the convergence factors. The method of Schwinger is to identify the terms corresponding to corrections in mass and charge and, previous to their evaluation, to remove them from the expressions for real processes. This has the advantage of showing that the results can be strictly independent of particular cut-off methods. On the other hand, many of the properties of the integrals are analyzed using formal properties of invariant propagation functions. But one of the properties is that the integrals are infinite and it is not clear to what extent this invalidates the demonstrations. A practical advantage of the present method is that ambiguities can be more easily resolved; simply by direct calculation of the otherwise divergent integrals. Nevertheless, it is not at all clear that the convergence factors do not upset the physical consistency of the theory. Although in the limit the two methods agree, neither method appears to be thoroughly satisfactory theoretically. Nevertheless, it does appear that we now have available a complete and definite method for the calculation of physical processes to any order in quantum electrodynamics.

Since we can write down the solution to any physical problem, we have a complete theory which could stand by itself. It will be theoretically incomplete, however, in two respects. First, although each term of increasing order in $e^2/\hbar c$ can be written down it would be desirable to see some way of expressing things in finite form to all orders in $e^2/\hbar c$ at once. Second, although it will be physically evident that the results obtained are equivalent to those obtained by conventional electrodynamics the mathematical proof of this is not included. Both of these limitations will be removed in a subsequent paper (see also Dyson[4]).

Briefly the genesis of this theory was this. The conventional electrodynamics was expressed in the Lagrangian form of quantum mechanics described in the Reviews of Modern Physics.[5] The motion of the field oscillators could be integrated out (as described in Section 13 of that paper), the result being an expression of the delayed interaction of the particles. Next the modification of the delta-function interaction could be made directly from the analogy to the classical case.[2] This was still not complete because the Lagrangian method had been worked out in detail only for particles obeying the non-relativistic Schrödinger equation. It was then modified in accordance with the requirements of the Dirac equation and the phenomenon of pair creation. This was made easier by the reinterpretation of the theory of holes (I). Finally for practical calculations the expressions were developed in a power series in $e^2/\hbar c$. It was apparent that each term in the series had a simple physical interpretation. Since the result was easier to understand than the derivation, it was thought best to publish the results first in this paper. Considerable time has been spent to make these first two papers as complete and as physically plausible as possible without relying on the Lagrangian method, because it is not generally familiar. It is realized that such a description cannot carry the conviction of truth which would accompany the derivation. On the other hand, in the interest of keeping simple things simple the derivation will appear in a separate paper.

The possible application of these methods to the various meson theories is discussed briefly. The formulas corresponding to a charge particle of zero spin moving in accordance with the Klein Gordon equation are also given. In an Appendix a method is given for calculating the integrals appearing in the matrix elements for the simpler processes.

The point of view which is taken here of the interaction of charges differs from the more usual point of view of field theory. Furthermore, the familiar Hamiltonian form of quantum mechanics must be compared to the over-all space-time view used here. The first section is, therefore, devoted to a discussion of the relations of these viewpoints.

1. COMPARISON WITH THE HAMILTONIAN METHOD

Electrodynamics can be looked upon in two equivalent and complementary ways. One is as the description of the behavior of a field (Maxwell's equations). The other is as a description of a direct interaction at a distance (albeit delayed in time) between charges (the solutions of Lienard and Wiechert). From the latter point of view light is considered as an interaction of the charges in the source with those in the absorber. This is an impractical point of view because many kinds of sources produce the same kind of effects. The field point of view separates these aspects into two simpler problems, production of light, and absorption of light. On the other hand, the field point of view is less practical when dealing with close collisions of particles (or their action on themselves). For here the source and absorber are not readily distinguishable, there is an intimate exchange of quanta. The fields are so closely determined by the motions of the particles that it is just as well not to separate the question into two problems but to consider the process as a direct interaction. Roughly, the field point of view is most practical for problems involv-

[4] J. Schwinger, Phys. Rev. **74**, 1439 (1948), Phys. Rev. **75**, 651 (1949). A proof of this equivalence is given by F. J. Dyson, Phys. Rev. **75**, 486 (1949).

[5] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948). The application to electrodynamics is described in detail by H. J. Groenewold, Koninklijke Nederlandsche Akademia van Weteschappen. Proceedings Vol. LII, 3 (226) 1949.

ing real quanta, while the interaction view is best for the discussion of the virtual quanta involved. We shall emphasize the interaction viewpoint in this paper, first because it is less familiar and therefore requires more discussion, and second because the important aspect in the problems with which we shall deal is the effect of virtual quanta.

The Hamiltonian method is not well adapted to represent the direct action at a distance between charges because that action is delayed. The Hamiltonian method represents the future as developing out of the present. If the values of a complete set of quantities are known now, their values can be computed at the next instant in time. If particles interact through a delayed interaction, however, one cannot predict the future by simply knowing the present motion of the particles. One would also have to know what the motions of the particles were in the past in view of the interaction this may have on the future motions. This is done in the Hamiltonian electrodynamics, of course, by requiring that one specify besides the present motion of the particles, the values of a host of new variables (the coordinates of the field oscillators) to keep track of that aspect of the past motions of the particles which determines their future behavior. The use of the Hamiltonian forces one to choose the field viewpoint rather than the interaction viewpoint.

In many problems, for example, the close collisions of particles, we are not interested in the precise temporal sequence of events. It is not of interest to be able to say how the situation would look at each instant of time during a collision and how it progresses from instant to instant. Such ideas are only useful for events taking a long time and for which we can readily obtain information during the intervening period. For collisions it is much easier to treat the process as a whole.[6] The Møller interaction matrix for the the collision of two electrons is not essentially more complicated than the nonrelativistic Rutherford formula, yet the mathematical machinery used to obtain the former from quantum electrodynamics is vastly more complicated than Schrödinger's equation with the e^2/r_{12} interaction needed to obtain the latter. The difference is only that in the latter the action is instantaneous so that the Hamiltonian method requires no extra variables, while in the former relativistic case it is delayed and the Hamiltonian method is very cumbersome.

We shall be discussing the solutions of equations rather than the time differential equations from which they come. We shall discover that the solutions, because of the over-all space-time view that they permit, are as easy to understand when interactions are delayed as when they are instantaneous.

As a further point, relativistic invariance will be self-evident. The Hamiltonian form of the equations develops the future from the instantaneous present. But for different observers in relative motion the instantaneous present is different, and corresponds to a different 3-dimensional cut of space-time. Thus the temporal analyses of different observers is different and their Hamiltonian equations are developing the process in different ways. These differences are irrelevant, however, for the solution is the same in any space time frame. By forsaking the Hamiltonian method, the wedding of relativity and quantum mechanics can be accomplished most naturally.

We illustrate these points in the next section by studying the solution of Schrödinger's equation for nonrelativistic particles interacting by an instantaneous Coulomb potential (Eq. 2). When the solution is modified to include the effects of delay in the interaction and the relativistic properties of the electrons we obtain an expression of the laws of quantum electrodynamics (Eq. 4).

2. THE INTERACTION BETWEEN CHARGES

We study by the same methods as in I, the interaction of two particles using the same notation as I. We start by considering the non-relativistic case described by the Schrödinger equation (I, Eq. 1). The wave function at a given time is a function $\psi(\mathbf{x}_a, \mathbf{x}_b, t)$ of the coordinates $\mathbf{x}_a$ and $\mathbf{x}_b$ of each particle. Thus call $K(\mathbf{x}_a, \mathbf{x}_b, t; \mathbf{x}_a', \mathbf{x}_b', t')$ the amplitude that particle a at $\mathbf{x}_a'$ at time t' will get to $\mathbf{x}_a$ at t while particle b at $\mathbf{x}_b'$ at t' gets to $\mathbf{x}_b$ at t. If the particles are free and do not interact this is

$$K(\mathbf{x}_a, \mathbf{x}_b, t; \mathbf{x}_a', \mathbf{x}_b', t') = K_{0a}(\mathbf{x}_a, t; \mathbf{x}_a', t')K_{0b}(\mathbf{x}_b, t; \mathbf{x}_b', t')$$

where K_{0a} is the K_0 function for particle a considered as free. In *this* case we can obviously define a quantity like K, but for which the time t need not be the same for particles a and b (likewise for t'); e.g.,

$$K_0(3, 4; 1, 2) = K_{0a}(3, 1)K_{0b}(4, 2) \tag{1}$$

can be thought of as the amplitude that particle a goes from $\mathbf{x}_1$ at t_1 to $\mathbf{x}_3$ at t_3 and that particle b goes from $\mathbf{x}_2$ at t_2 to $\mathbf{x}_4$ at t_4.

When the particles do interact, one can only define the quantity $K(3, 4; 1, 2)$ precisely if the interaction vanishes between t_1 and t_2 and also between t_3 and t_4. In a real physical system such is not the case. There is such an enormous advantage, however, to the concept that we shall continue to use it, imagining that we can neglect the effect of interactions between t_1 and t_2 and between t_3 and t_4. For practical problems this means choosing such long time intervals t_3-t_1 and t_4-t_2 that the extra interactions near the end points have small relative effects. As an example, in a scattering problem it may well be that the particles are so well separated initially and finally that the interaction at these times is negligible. Again energy values can be defined by the average rate of change of phase over such long time intervals that errors initially and finally can be neglected. Inasmuch as any physical problem can be defined in terms of scattering processes we do not lose much in

[6] This is the viewpoint of the theory of the S matrix of Heisenberg.

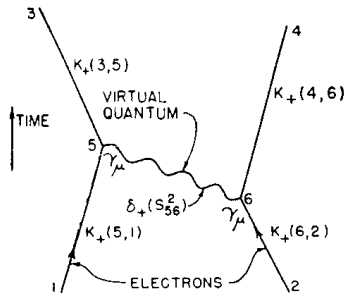


FIG. 1. The fundamental interaction Eq. (4). Exchange of one quantum between two electrons.

a general theoretical sense by this approximation. If it is not made it is not easy to study interacting particles relativistically, for there is nothing significant in choosing $t_1=t_3$ if $\mathbf{x}_1\neq\mathbf{x}_3$, as absolute simultaneity of events at a distance cannot be defined invariantly. It is essentially to avoid this approximation that the complicated structure of the older quantum electrodynamics has been built up. We wish to describe electrodynamics as a delayed interaction between particles. If we can make the approximation of assuming a meaning to $K(3,4;1,2)$ the results of this interaction can be expressed very simply.

To see how this may be done, imagine first that the interaction is simply that given by a Coulomb potential e^2/r where r is the distance between the particles. If this be turned on only for a very short time Δt_0 at time t_0, the first order correction to $K(3,4;1,2)$ can be worked out exactly as was Eq. (9) of I by an obvious generalization to two particles:

$$K^{(1)}(3,4;1,2)=-ie^2\int\int K_{0a}(3,5)K_{0b}(4,6)r_{56}^{-1}\times K_{0a}(5,1)K_{0b}(6,2)d^3\mathbf{x}_5d^3\mathbf{x}_6\Delta t_0,$$

where $t_5=t_6=t_0$. If now the potential were on at all times (so that strictly K is not defined unless $t_4=t_3$ and $t_1=t_2$), the first-order effect is obtained by integrating on t_0, which we can write as an integral over both t_5 and t_6 if we include a delta-function $\delta(t_5-t_6)$ to insure contribution only when $t_5=t_6$. Hence, the first-order effect of interaction is (calling $t_5-t_6=t_{56}$):

$$K^{(1)}(3,4;1,2)=-ie^2\int\int K_{0a}(3,5)K_{0b}(4,6)r_{56}^{-1}\times\delta(t_{56})K_{0a}(5,1)K_{0b}(6,2)d\tau_5d\tau_6, \quad (2)$$

where $d\tau=d^3\mathbf{x}dt$.

We know, however, in classical electrodynamics, that the Coulomb potential does not act instantaneously, but is delayed by a time r_{56}, taking the speed of light as unity. This suggests simply replacing $r_{56}^{-1}\delta(t_{56})$ in (2) by something like $r_{56}^{-1}\delta(t_{56}-r_{56})$ to represent the delay in the effect of b on a.

This turns out to be not quite right,[7] for when this interaction is represented by photons they must be of only positive energy, while the Fourier transform of $\delta(t_{56}-r_{56})$ contains frequencies of both signs. It should instead be replaced by $\delta_+(t_{56}-r_{56})$ where

$$\delta_+(x)=\int_0^\infty e^{-i\omega x}d\omega/\pi=\lim_{\epsilon\to 0}\frac{(\pi i)^{-1}}{x-i\epsilon}=\delta(x)+(\pi ix)^{-1}. \quad (3)$$

This is to be averaged with $r_{56}^{-1}\delta_+(-t_{56}-r_{56})$ which arises when $t_5<t_6$ and corresponds to a emitting the quantum which b receives. Since

$$(2r)^{-1}(\delta_+(t-r)+\delta_+(-t-r))=\delta_+(t^2-r^2),$$

this means $r_{56}^{-1}\delta(t_{56})$ is replaced by $\delta_+(s_{56}^2)$ where $s_{56}^2=t_{56}^2-r_{56}^2$ is the square of the relativistically invariant interval between points 5 and 6. Since in classical electrodynamics there is also an interaction through the vector potential, the complete interaction (see A, Eq. (1)) should be $(1-(\mathbf{v}_5\cdot\mathbf{v}_6)\delta_+(s_{56}^2)$, or in the relativistic case,

$$(1-\boldsymbol{\alpha}_a\cdot\boldsymbol{\alpha}_b)\delta_+(s_{56}^2)=\beta_a\beta_b\gamma_{a\mu}\gamma_{b\mu}\delta_+(s_{56}^2).$$

Hence we have for electrons obeying the Dirac equation,

$$K^{(1)}(3,4;1,2)=-ie^2\int\int K_{+a}(3,5)K_{+b}(4,6)\gamma_{a\mu}\gamma_{b\mu}\times\delta_+(s_{56}^2)K_{+a}(5,1)K_{+b}(6,2)d\tau_5d\tau_6, \quad (4)$$

where $\gamma_{a\mu}$ and $\gamma_{b\mu}$ are the Dirac matrices applying to the spinor corresponding to particles a and b, respectively (the factor $\beta_a\beta_b$ being absorbed in the definition, I Eq. (17), of K_+).

This is our fundamental equation for electrodynamics. It describes the effect of exchange of one quantum (therefore first order in e^2) between two electrons. It will serve as a prototype enabling us to write down the corresponding quantities involving the exchange of two or more quanta between two electrons or the interaction of an electron with itself. It is a consequence of conventional electrodynamics. Relativistic invariance is clear. Since one sums over μ it contains the effects of both longitudinal and transverse waves in a relativistically symmetrical way.

We shall now interpret Eq. (4) in a manner which will permit us to write down the higher order terms. It can be understood (see Fig. 1) as saying that the amplitude for "a" to go from 1 to 3 and "b" to go from 2 to 4 is altered to first order because they can exchange a quantum. Thus, "a" can go to 5 (amplitude $K_+(5,1)$)

[7] It, and a like term for the effect of a on b, leads to a theory which, in the classical limit, exhibits interaction through half-advanced and half-retarded potentials. Classically, this is equivalent to purely retarded effects within a closed box from which no light escapes (e.g., see A, or J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. 17, 157 (1945)). Analogous theorems exist in quantum mechanics but it would lead us too far astray to discuss them now.

emit a quantum (longitudinal, transverse, or scalar $\gamma_{a\mu}$) and then proceed to 3 ($K_+(3, 5)$). Meantime "b" goes to 6 ($K_+(6, 2)$), absorbs the quantum ($\gamma_{b\mu}$) and proceeds to 4 ($K_+(4, 6)$). The quantum meanwhile proceeds from 5 to 6, which it does with amplitude $\delta_+(s_{56}^2)$. We must sum over all the possible quantum polarizations μ and positions and times of emission 5, and of absorption 6. Actually if $t_5 > t_6$ it would be better to say that "a" absorbs and "b" emits but no attention need be paid to these matters, as all such alternatives are automatically contained in (4).

The correct terms of higher order in e^2 or involving larger numbers of electrons (interacting with themselves or in pairs) can be written down by the same kind of reasoning. They will be illustrated by examples as we proceed. In a succeeding paper they will all be deduced from conventional quantum electrodynamics.

Calculation, from (4), of the transition element between positive energy free electron states gives the Möller scattering of two electrons, when account is taken of the Pauli principle.

The exclusion principle for interacting charges is handled in exactly the same way as for non-interacting charges (I). For example, for two charges it requires only that one calculate $K(3, 4; 1, 2) - K(4, 3; 1, 2)$ to get the net amplitude for arrival of charges at 3 and 4. It is disregarded in intermediate states. The interference effects for scattering of electrons by positrons discussed by Bhabha will be seen to result directly in this formulation. The formulas are interpreted to apply to positrons in the manner discussed in I.

As our primary concern will be for processes in which the quanta are virtual we shall not include here the detailed analysis of processes involving real quanta in initial or final state, and shall content ourselves by only stating the rules applying to them.[8] The result of the analysis is, as expected, that they can be included by the same line of reasoning as is used in discussing the virtual processes, provided the quantities are normalized in the usual manner to represent single quanta. For example, the amplitude that an electron in going from 1 to 2 absorbs a quantum whose vector potential, suitably normalized, is $c_\mu \exp(-ik\cdot x) = C_\mu(x)$ is just the expression (I, Eq. (13)) for scattering in a potential with $\boldsymbol{A}$ (3) replaced by $\boldsymbol{C}$ (3). Each quantum interacts only once (either in emission or in absorption), terms like (I, Eq. (14)) occur only when there is more than one quantum involved. The Bose statistics of the quanta can, in all cases, be disregarded in intermediate states. The only effect of the statistics is to change the weight of initial or final states. If there are among quanta, in the initial state, some n which are identical then the weight of the state is $(1/n!)$ of what it would be if these quanta were considered as different (similarly for the final state).

3. THE SELF-ENERGY PROBLEM

Having a term representing the mutual interaction of a pair of charges, we must include similar terms to represent the interaction of a charge with itself. For under some circumstances what appears to be two distinct electrons may, according to I, be viewed also as a single electron (namely in case one electron was created in a pair with a positron destined to annihilate the other electron). Thus to the interaction between such electrons must correspond the possibility of the action of an electron on itself.[9]

This interaction is the heart of the self energy problem. Consider to first order in e^2 the action of an electron on itself in an otherwise force free region. The amplitude $K(2, 1)$ for a single particle to get from 1 to 2 differs from $K_+(2, 1)$ to first order in e^2 by a term

$$K^{(1)}(2, 1) = -ie^2 \int\int K_+(2, 4)\gamma_\mu K_+(4, 3)\gamma_\mu \times K_+(3, 1) d\tau_3 d\tau_4 \delta_+(s_{43}^2). \quad (6)$$

It arises because the electron instead of going from 1 directly to 2, may go (Fig. 2) first to 3, ($K_+(3, 1)$), emit a quantum (γ_μ), proceed to 4, ($K_+(4, 3)$), absorb it (γ_μ), and finally arrive at 2 ($K_+(2, 4)$). The quantum must go from 3 to 4 ($\delta_+(s_{43}^2)$).

This is related to the self-energy of a free electron in the following manner. Suppose initially, time t_1, we have an electron in state $f(1)$ which we imagine to be a positive energy solution of Dirac's equation for a free particle. After a long time $t_2 - t_1$ the perturbation will alter

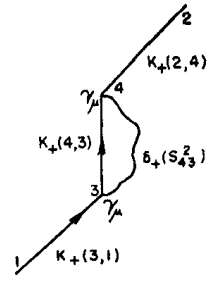


FIG. 2. Interaction of an electron with itself, Eq. (6).

[8] Although in the expressions stemming from (4) the quanta are virtual, this is not actually a theoretical limitation. One way to deduce the correct rules for real quanta from (4) is to note that in a closed system all quanta can be considered as virtual (i.e., they have a known source and are eventually absorbed) so that in such a system the present description is complete and equivalent to the conventional one. In particular, the relation of the Einstein A and B coefficients can be deduced. A more practical direct deduction of the expressions for real quanta will be given in the subsequent paper. It might be noted that (4) can be rewritten as describing the action on a, $K^{(1)}(3, 1) = i\int K_+(3, 5) \times \boldsymbol{A}(5) K_+(5, 1) d\tau_5$ of the potential $A_\mu(5) = e^2 \int K_+(4, 6)\delta_+(s_{56}^2)\gamma_\mu \times K_+(6, 2) d\tau_6$ arising from Maxwell's equations $-\Box^2 A_\mu = 4\pi j_\mu$ from a "current" $j_\mu(6) = e^2 K_+(4, 6)\gamma_\mu K_+(6, 2)$ produced by particle b in going from 2 to 4. This is virtue of the fact that δ_+ satisfies

$$-\Box_2^2 \delta_+(s_{21}^2) = 4\pi\delta(2, 1). \quad (5)$$

[9] These considerations make it appear unlikely that the contention of J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. 17, 157 (1945), that electrons do not act on themselves, will be a successful concept in quantum electrodynamics.

the wave function, which can then be looked upon as a superposition of free particle solutions (actually it only contains f). The amplitude that $g(2)$ is contained is calculated as in (I, Eq. (21)). The diagonal element ($g=f$) is therefore

$$\int\int \bar{f}(2)\beta K^{(1)}(2,\,1)\beta f(1)d^3\mathbf{x}_1 d^3\mathbf{x}_2. \qquad (7)$$

The time interval $T=t_2-t_1$ (and the spatial volume V over which one integrates) must be taken very large, for the expressions are only approximate (analogous to the situation for two interacting charges).[10] This is because, for example, we are dealing incorrectly with quanta emitted just before t_2 which would normally be reabsorbed at times after t_2.

If $K^{(1)}(2,\,1)$ from (6) is actually substituted into (7) the surface integrals can be performed as was done in obtaining I, Eq. (22) resulting in

$$-ie^2\int\int \bar{f}(4)\gamma_\mu K_+(4,\,3)\gamma_\mu f(3)\delta_+(s_{43}^2)d\tau_3 d\tau_4. \qquad (8)$$

Putting for $f(1)$ the plane wave $u\exp(-ip\cdot x_1)$ where p_μ is the energy (p_4) and momentum of the electron ($p^2=m^2$), and u is a constant 4-index symbol, (8) becomes

$$-ie^2\int\int (\bar{u}\gamma_\mu K_+(4,\,3)\gamma_\mu u) \times\exp(ip\cdot(x_1-x_3))\delta_+(s_{43}^2)d\tau_3 d\tau_4,$$

the integrals extending over the volume V and time interval T. Since $K_+(4,\,3)$ depends only on the difference of the coordinates of 4 and 3, $x_{43\mu}$, the integral on 4 gives a result (except near the surfaces of the region) independent of 3. When integrated on 3, therefore, the result is of order VT. The effect is proportional to V, for the wave functions have been normalized to unit volume. If normalized to volume V, the result would simply be proportional to T. This is expected, for if the effect were equivalent to a change in energy ΔE, the amplitude for arrival in f at t_2 is altered by a factor $\exp(-i\Delta E(t_2-t_1))$, or to first order by the difference $-i(\Delta E)T$. Hence, we have

$$\Delta E=e^2\int (\bar{u}\gamma_\mu K_+(4,\,3)\gamma_\mu u)\exp(ip\cdot x_{43})\delta_+(s_{43}^2)d\tau_4, \qquad (9)$$

integrated over all space-time $d\tau_4$. This expression will be simplified presently. In interpreting (9) we have tacitly assumed that the wave functions are normalized so that $(u^*u)=(\bar{u}\gamma_4 u)=1$. The equation may therefore be made independent of the normalization by writing the left side as $(\Delta E)(\bar{u}\gamma_4 u)$, or since $(\bar{u}\gamma_4 u)=(E/m)(\bar{u}u)$ and $m\Delta m=E\Delta E$, as $\Delta m(\bar{u}u)$ where Δm is an equivalent change in mass of the electron. In this form invariance is obvious.

One can likewise obtain an expression for the energy shift for an electron in a hydrogen atom. Simply replace K_+ in (8), by $K_+^{(V)}$, the exact kernel for an electron in the potential, $V=\beta e^2/r$, of the atom, and f by a wave function (of space and time) for an atomic state. In general the ΔE which results is not real. The imaginary part is negative and in $\exp(-i\Delta ET)$ produces an exponentially decreasing amplitude with time. This is because we are asking for the amplitude that an atom initially with no photon in the field, will still appear after time T with no photon. If the atom is in a state which can radiate, this amplitude must decay with time. The imaginary part of ΔE when calculated does indeed give the correct rate of radiation from atomic states. It is zero for the ground state and for a free electron.

In the non-relativistic region the expression for ΔE can be worked out as has been done by Bethe.[11] In the relativistic region (points 4 and 3 as close together as a Compton wave-length) the $K_+^{(V)}$ which should appear in (8) can be replaced to first order in V by K_+ plus $K_+^{(1)}(2,\,1)$ given in I, Eq. (13). The problem is then very similar to the radiationless scattering problem discussed below.

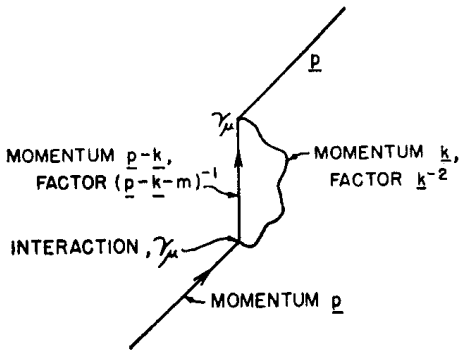


FIG. 3. Interaction of an electron with itself. Momentum space, Eq. (11).

4. EXPRESSION IN MOMENTUM AND ENERGY SPACE

The evaluation of (9), as well as all the other more complicated expressions arising in these problems, is very much simplified by working in the momentum and energy variables, rather than space and time. For this we shall need the Fourier Transform of $\delta_+(s_{21}^2)$ which is

$$-\delta_+(s_{21}^2)=\pi^{-1}\int \exp(-ik\cdot x_{21})k^{-2}d^4k, \qquad (10)$$

which can be obtained from (3) and (5) or from I, Eq. (32) noting that $I_+(2,\,1)$ for $m^2=0$ is $\delta_+(s_{21}^2)$ from

[10] This is discussed in reference 5 in which it is pointed out that the concept of a wave function loses accuracy if there are delayed self-actions.

[11] H. A. Bethe, Phys. Rev. 72, 339 (1947).

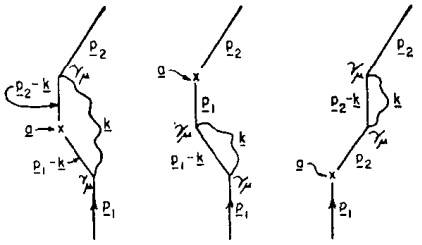


FIG. 4. Radiative correction to scattering, momentum space.

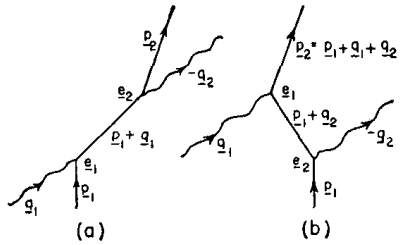


FIG. 5. Compton scattering, Eq. (15).

I, Eq. (34). The k^{-2} means $(k\cdot k)^{-1}$ or more precisely the limit as $\delta\to 0$ of $(k\cdot k+i\delta)^{-1}$. Further d^4k means $(2\pi)^{-2}dk_1dk_2dk_3dk_4$. If we imagine that quanta are particles of zero mass, then we can make the general rule that all poles are to be resolved by considering the masses of the particles and quanta to have infinitesimal negative imaginary parts.

Using these results we see that the self-energy (9) is the matrix element between $\bar{u}$ and u of the matrix

$$(e^2/\pi i)\int \gamma_\mu(p-k-m)^{-1}\gamma_\mu k^{-2}d^4k, \qquad (11)$$

where we have used the expression (I, Eq. (31)) for the Fourier transform of K_+. This form for the self-energy is easier to work with than is (9).

The equation can be understood by imagining (Fig. 3) that the electron of momentum p emits (γ_μ) a quantum of momentum k, and makes its way now with momentum $p-k$ to the next event (factor $(p-k-m)^{-1}$) which is to absorb the quantum (another γ_μ). The amplitude of propagation of quanta is k^{-2}. (There is a factor $e^2/\pi i$ for each virtual quantum). One integrates over all quanta. The reason an electron of momentum p propagates as $1/(p-m)$ is that this operator is the reciprocal of the Dirac equation operator, and we are simply solving this equation. Likewise light goes as $1/k^2$, for this is the reciprocal D'Alembertian operator of the wave equation of light. The first γ_μ represents the current which generates the vector potential, while the second is the velocity operator by which this potential is multiplied in the Dirac equation when an external field acts on an electron.

Using the same line of reasoning, other problems may be set up directly in momentum space. For example, consider the scattering in a potential $A=A_\mu\gamma_\mu$ varying in space and time as $a\exp(-iq\cdot x)$. An electron initially in state of momentum $p_1=p_{1\mu}\gamma_\mu$ will be deflected to state p_2 where $p_2=p_1+q$. The zero-order answer is simply the matrix element of a between states 1 and 2. We next ask for the first order (in e^2) radiative correction due to virtual radiation of one quantum. There are several ways this can happen. First for the case illustrated in Fig. 4(a), find the matrix:

$$(e^2/\pi i)\int \gamma_\mu(p_2-k-m)^{-1}a(p_1-k-m)^{-1}\gamma_\mu k^{-2}d^4k. \qquad (12)$$

For in this case, first[12] a quantum of momentum k is emitted (γ_μ), the electron then having momentum p_1-k and hence propagating with factor $(p_1-k-m)^{-1}$. Next it is scattered by the potential (matrix a) receiving additional momentum q, propagating on then (factor $(p_2-k-m)^{-1}$) with the new momentum until the quantum is reabsorbed (γ_μ). The quantum propagates from emission to absorption (k^{-2}) and we integrate over all quanta (d^4k), and sum on polarization μ. When this is integrated on k_4, the result can be shown to be exactly equal to the expressions (16) and (17) given in B for the same process, the various terms coming from residues of the poles of the integrand (12).

Or again if the quantum is both emitted and reabsorbed before the scattering takes place one finds (Fig. 4(b))

$$(e^2/\pi i)\int a(p_1-m)^{-1}\gamma_\mu(p_1-k-m)^{-1}\gamma_\mu k^{-2}d^4k, \qquad (13)$$

or if both emission and absorption occur after the scattering, (Fig. 4(c))

$$(e^2/\pi i)\int \gamma_\mu(p_2-k-m)^{-1}\gamma_\mu(p_2-m)^{-1}ak^{-2}d^4k. \qquad (14)$$

These terms are discussed in detail below.

We have now achieved our simplification of the form of writing matrix elements arising from virtual processes. Processes in which a number of real quanta is given initially and finally offer no problem (assuming correct normalization). For example, consider the Compton effect (Fig. 5(a)) in which an electron in state p_1 absorbs a quantum of momentum q_1, polarization vector $e_{1\mu}$ so that its interaction is $e_{1\mu}\gamma_\mu=e_1$, and emits a second quantum of momentum $-q_2$, polarization e_2 to arrive in final state of momentum p_2. The matrix for

[12] First, next, etc., here refer not to the order in true time but to the succession of events along the trajectory of the electron. That is, more precisely, to the order of appearance of the matrices in the expressions.

this process is $e_2(p_1+q_1-m)^{-1}e_1$. The total matrix for the Compton effect is, then,

$$e_2(p_1+q_1-m)^{-1}e_1+e_1(p_1+q_2-m)^{-1}e_2, \qquad (15)$$

the second term arising because the emission of e_2 may also precede the absorption of e_1 (Fig. 5(b)). One takes matrix elements of this between initial and final electron states ($p_1+q_1=p_2-q_2$), to obtain the Klein Nishina formula. Pair annihilation with emission of two quanta, etc., are given by the same matrix, positron states being those with negative time component of p. Whether quanta are absorbed or emitted depends on whether the time component of q is positive or negative.

5. THE CONVERGENCE OF PROCESSES WITH VIRTUAL QUANTA

These expressions are, as has been indicated, no more than a re-expression of conventional quantum electrodynamics. As a consequence, many of them are meaningless. For example, the self-energy expression (9) or (11) gives an infinite result when evaluated. The infinity arises, apparently, from the coincidence of the δ-function singularities in $K_+(4, 3)$ and $\delta_+(s_{43}^2)$. Only at this point is it necessary to make a real departure from conventional electrodynamics, a departure other than simply rewriting expressions in a simpler form.

We desire to make a modification of quantum electrodynamics analogous to the modification of classical electrodynamics described in a previous article, A. There the $\delta(s_{12}^2)$ appearing in the action of interaction was replaced by $f(s_{12}^2)$ where $f(x)$ is a function of small width and great height.

The obvious corresponding modification in the quantum theory is to replace the $\delta_+(s^2)$ appearing the quantum mechanical interaction by a new function $f_+(s^2)$. We can postulate that if the Fourier transform of the classical $f(s_{12}^2)$ is the integral over all k of $F(k^2)\exp(-ik\cdot x_{12})d^4k$, then the Fourier transform of $f_+(s^2)$ is the same integral taken over only positive frequencies k_4 for $t_2>t_1$ and over only negative ones for $t_2<t_1$ in analogy to the relation of $\delta_+(s^2)$ to $\delta(s^2)$. The function $f(s^2)=f(x\cdot x)$ can be written* as

$$f(x\cdot x)=(2\pi)^{-2}\int_{k_4=0}^{\infty}\int \sin(k_4|x_4|)\times\cos(\mathbf{K}\cdot\mathbf{x})dk_4d^3\mathbf{K}g(k\cdot k),$$

where $g(k\cdot k)$ is k_4^{-1} times the density of oscillators and may be expressed for positive k_4 as (A, Eq. (16))

$$g(k^2)=\int_0^{\infty}(\delta(k^2)-\delta(k^2-\lambda^2))G(\lambda)d\lambda,$$

where $\int_0^{\infty}G(\lambda)d\lambda=1$ and G involves values of λ large compared to m. This simply means that the amplitude for propagation of quanta of momentum k is

$$-F_+(k^2)=\pi^{-1}\int_0^{\infty}(k^{-2}-(k^2-\lambda^2)^{-1})G(\lambda)d\lambda,$$

rather than k^{-2}. That is, writing $F_+(k^2)=-\pi^{-1}k^{-2}C(k^2)$,

$$-f_+(s_{12}^2)=\pi^{-1}\int \exp(-ik\cdot x_{12})k^{-2}C(k^2)d^4k. \qquad (16)$$

Every integral over an intermediate quantum which previously involved a factor d^4k/k^2 is now supplied with a convergence factor $C(k^2)$ where

$$C(k^2)=\int_0^{\infty}-\lambda^2(k^2-\lambda^2)^{-1}G(\lambda)d\lambda. \qquad (17)$$

The poles are defined by replacing k^2 by $k^2+i\delta$ in the limit $\delta\to 0$. That is λ^2 may be assumed to have an infinitesimal negative imaginary part.

The function $f_+(s_{12}^2)$ may still have a discontinuity in value on the light cone. This is of no influence for the Dirac electron. For a particle satisfying the Klein Gordon equation, however, the interaction involves gradients of the potential which reinstates the δ function if f has discontinuities. The condition that f is to have no discontinuity in value on the light cone implies $k^2C(k^2)$ approaches zero as k^2 approaches infinity. In terms of $G(\lambda)$ the condition is

$$\int_0^{\infty}\lambda^2G(\lambda)d\lambda=0. \qquad (18)$$

This condition will also be used in discussing the convergence of vacuum polarization integrals.

The expression for the self-energy matrix is now

$$(e^2/\pi i)\int\gamma_\mu(p-k-m)^{-1}\gamma_\mu k^{-2}d^4kC(k^2), \qquad (19)$$

which, since $C(k^2)$ falls off at least as rapidly as $1/k^2$, converges. For practical purposes we shall suppose hereafter that $C(k^2)$ is simply $-\lambda^2/(k^2-\lambda^2)$ implying that some average (with weight $G(\lambda)d\lambda$) over values of λ may be taken afterwards. Since in all processes the quantum momentum will be contained in at least one extra factor of the form $(p-k-m)^{-1}$ representing propagation of an electron while that quantum is in the field, we can expect all such integrals with their convergence factors to converge and that the result of all such processes will now be finite and definite (excepting the processes with closed loops, discussed below, in which the diverging integrals are over the momenta of the electrons rather than the quanta).

The integral of (19) with $C(k^2)=-\lambda^2(k^2-\lambda^2)^{-1}$ noting that $p^2=m^2$, $\lambda\gg m$ and dropping terms of order m/λ, is (see Appendix A)

$$(e^2/2\pi)[4m(\ln(\lambda/m)+\tfrac{1}{2})-p(\ln(\lambda/m)+5/4)]. \qquad (20)$$

* This relation is given incorrectly in A, equation just preceding 16.

When applied to a state of an electron of momentum p satisfying $pu=mu$, it gives for the change in mass (as in B, Eq. (9))

$$\Delta m = m(e^2/2\pi)(3\ln(\lambda/m)+\tfrac{3}{4}). \tag{21}$$

6. RADIATIVE CORRECTIONS TO SCATTERING

We can now complete the discussion of the radiative corrections to scattering. In the integrals we include the convergence factor $C(k^2)$, so that they converge for large k. Integral (12) is also not convergent because of the well-known infra-red catastrophy. For this reason we calculate (as discussed in B) the value of the integral assuming the photons to have a small mass $\lambda_{\min}\ll m\ll\lambda$. The integral (12) becomes

$$(e^2/\pi i)\int \gamma_\mu(p_2-k-m)^{-1}a(p_1-k-m)^{-1} \times\gamma_\mu(k^2-\lambda_{\min}^2)^{-1}d^4kC(k^2-\lambda_{\min}^2),$$

which when integrated (see Appendix B) gives $(e^2/2\pi)$ times

$$\left[2\left(\ln\frac{m}{\lambda_{\min}}-1\right)\left(1-\frac{2\theta}{\tan2\theta}\right)+\theta\tan\theta +\frac{4}{\tan2\theta}\int_0^\theta \alpha\tan\alpha d\alpha\right]a +\frac{1}{4m}(qa-aq)\frac{2\theta}{\sin2\theta}+ra, \tag{22}$$

where $(q^2)^{\frac{1}{2}}=2m\sin\theta$ and we have assumed the matrix to operate between states of momentum p_1 and $p_2=p_1+q$ and have neglected terms of order $\lambda_{\min}/m$, m/λ, and q^2/λ^2. Here the only dependence on the convergence factor is in the term ra, where

$$r=\ln(\lambda/m)+9/4-2\ln(m/\lambda_{\min}). \tag{23}$$

As we shall see in a moment, the other terms (13), (14) give contributions which just cancel the ra term. The remaining terms give for small q,

$$(e^2/4\pi)\left(\frac{1}{2m}(qa-aq)+\frac{4q^2}{3m^2}a\left(\ln\frac{m}{\lambda_{\min}}-\frac{3}{8}\right)\right), \tag{24}$$

which shows the change in magnetic moment and the Lamb shift as interpreted in more detail in B.[13]

We must now study the remaining terms (13) and (14). The integral on k in (13) can be performed (after multiplication by $C(k^2)$) since it involves nothing but the integral (19) for the self-energy and the result is allowed to operate on the initial state u_1, (so that $p_1u_1=mu_1$). Hence the factor following $a(p_1-m)^{-1}$ will be just Δm. But, if one now tries to expand $1/(p_1-m)=(p_1+m)/(p_1^2-m^2)$ one obtains an infinite result, since $p_1^2=m^2$. This is, however, just what is expected physically. For the quantum can be emitted and absorbed at any time previous to the scattering. Such a process has the effect of a change in mass of the electron in the state 1. It therefore changes the energy by ΔE and the amplitude to first order in ΔE by $-i\Delta E\cdot t$ where t is the time it is acting, which is infinite. That is, the major effect of this term would be canceled by the effect of change of mass Δm.

The situation can be analyzed in the following manner. We suppose that the electron approaching the scattering potential a has not been free for an infinite time, but at some time far past suffered a scattering by a potential b. If we limit our discussion to the effects of Δm and of the virtual radiation of one quantum between two such scatterings each of the effects will be finite, though large, and their difference is determinate. The propagation from b to a is represented by a matrix

$$a(p'-m)^{-1}b, \tag{25}$$

in which one is to integrate possibly over p' (depending on details of the situation). (If the time is long between b and a, the energy is very nearly determined so that p'^2 is very nearly m^2.)

We shall compare the effect on the matrix (25) of the virtual quanta and of the change of mass Δm. The effect of a virtual quantum is

$$(e^2/\pi i)\int a(p'-m)^{-1}\gamma_\mu(p'-k-m)^{-1} \times\gamma_\mu(p'-m)^{-1}bk^{-2}d^4kC(k^2), \tag{26}$$

while that of a change of mass can be written

$$a(p'-m)^{-1}\Delta m(p'-m)^{-1}b, \tag{27}$$

and we are interested in the difference (26)–(27). A simple and direct method of making this comparison is just to evaluate the integral on k in (26) and subtract from the result the expression (27) where Δm is given in (21). The remainder can be expressed as a multiple $-r(p'^2)$ of the unperturbed amplitude (25);

$$-r(p'^2)a(p'-m)^{-1}b. \tag{28}$$

This has the same result (to this order) as replacing the potentials a and b in (25) by $(1-\frac{1}{2}r(p'^2))a$ and

[13] That the result given in B in Eq. (19) was in error was repeatedly pointed out to the author, in private communication, by V. F. Weisskopf and J. B. French, as their calculation, completed simultaneously with the author's early in 1948, gave a different result. French has finally shown that although the expression for the radiationless scattering B, Eq. (18) or (24) above is correct, it was incorrectly joined onto Bethe's non-relativistic result. He shows that the relation $\ln2k_{\max}-1=\ln\lambda_{\min}$ used by the author should have been $\ln2k_{\max}-5/6=\ln\lambda_{\min}$. This results in adding a term $-(1/6)$ to the logarithm in B, Eq. (19) so that the result now agrees with that of J. B. French and V. F. Weisskopf, Phys. Rev. 75, 1240 (1949) and N. H. Kroll and W. E. Lamb, Phys. Rev. 75, 388 (1949). The author feels unhappily responsible for the very considerable delay in the publication of French's result occasioned by this error. This footnote is appropriately numbered.

$(1-\frac{1}{2}r(p'^2))b$. In the limit, then, as $p'^2\to m^2$ the net effect on the scattering is $-\frac{1}{2}ra$ where r, the limit of $r(p'^2)$ as $p'^2\to m^2$ (assuming the integrals have an infrared cut-off), turns out to be just equal to that given in (23). An equal term $-\frac{1}{2}ra$ arises from virtual transitions after the scattering (14) so that the entire ra term in (22) is canceled.

The reason that r is just the value of (12) when $q^2=0$ can also be seen without a direct calculation as follows: Let us call p the vector of length m in the direction of p' so that if $p'^2=m(1+\epsilon)^2$ we have $p'=(1+\epsilon)p$ and we take ϵ as very small, being of order T^{-1} where T is the time between the scatterings b and a. Since $(p'-m)^{-1}=(p'+m)/(p'^2-m^2)\approx(p+m)/2m^2\epsilon$, the quantity (25) is of order ϵ^{-1} or T. We shall compute corrections to it only to its own order (ϵ^{-1}) in the limit $\epsilon\to 0$. The term (27) can be written approximately[14] as

$$(e^2/\pi i)\int a(p'-m)^{-1}\gamma_\mu(p-k-m)^{-1}\times\gamma_\mu(p'-m)^{-1}bk^{-2}d^4kC(k^2),$$

using the expression (19) for Δm. The net of the two effects is therefore approximately[15]

$$-(e^2/\pi i)\int a(p'-m)^{-1}\gamma_\mu(p-k-m)^{-1}\epsilon p(p-k-m)^{-1}\times\gamma_\mu(p'-m)^{-1}bk^{-2}d^4kC(k^2),$$

a term now of order $1/\epsilon$ (since $(p'-m)^{-1}\approx(p+m)\times(2m^2\epsilon)^{-1}$) and therefore the one desired in the limit. Comparison to (28) gives for r the expression

$$(p_1+m/2m)\int\gamma_\mu(p_1-k-m)^{-1}(p_1m^{-1})(p_1-k-m)^{-1}\times\gamma_\mu k^{-2}d^4kC(k^2). \quad (29)$$

The integral can be immediately evaluated, since it is the same as the integral (12), but with $q=0$, for a replaced by p_1/m. The result is therefore $r\cdot(p_1/m)$ which when acting on the state u_1 is just r, as $p_1u_1=mu_1$. For the same reason the term $(p_1+m)/2m$ in (29) is effectively 1 and we are left with $-r$ of (23).[16]

In more complex problems starting with a free electron the same type of term arises from the effects of a virtual emission and absorption both previous to the other processes. They, therefore, simply lead to the same factor r so that the expression (23) may be used directly and these renormalization integrals need not be computed afresh for each problem.

In this problem of the radiative corrections to scattering the net result is insensitive to the cut-off. This means, of course, that by a simple rearrangement of terms previous to the integration we could have avoided the use of the convergence factors completely (see for example Lewis[17]). The problem was solved in the manner here in order to illustrate how the use of such convergence factors, even when they are actually unnecessary, may facilitate analysis somewhat by removing the effort and ambiguities that may be involved in trying to rearrange the otherwise divergent terms.

The replacement of δ_+ by f_+ given in (16), (17) is not determined by the analogy with the classical problem. In the classical limit only the real part of δ_+ (i.e., just δ) is easy to interpret. But by what should the imaginary part, $1/(\pi is^2)$, of δ_+ be replaced? The choice we have made here (in defining, as we have, the location of the poles of (17)) is arbitrary and almost certainly incorrect. If the radiation resistance is calculated for an atom, as the imaginary part of (8), the result depends slightly on the function f_+. On the other hand the light radiated at very large distances from a source is independent of f_+. The total energy absorbed by distant absorbers will not check with the energy loss of the source. We are in a situation analogous to that in the classical theory if the entire f function is made to contain only retarded contributions (see A, Appendix). One desires instead the analogue of $\langle F\rangle_{\rm ret}$ of A. This problem is being studied.

One can say therefore, that this attempt to find a consistent modification of quantum electrodynamics is incomplete (see also the question of closed loops, below). For it could turn out that any correct form of f_+ which will guarantee energy conservation may at the same time not be able to make the self-energy integral finite. The desire to make the methods of simplifying the calculation of quantum electrodynamic processes more widely available has prompted this publication before an analysis of the correct form for f_+ is complete. One might try to take the position that, since the energy discrepancies discussed vanish in the limit $\lambda\to\infty$, the correct physics might be considered to be that obtained by letting $\lambda\to\infty$ after mass renormalization. I have no proof of the mathematical consistency of this procedure, but the presumption is very strong that it is satisfactory. (It is also strong that a satisfactory form for f_+ can be found.)

7. THE PROBLEM OF VACUUM POLARIZATION

In the analysis of the radiative corrections to scattering one type of term was not considered. The potential

[14] The expression is not exact because the substitution of Δm by the integral in (19) is valid only if p operates on a state such that p can be replaced by m. The error, however, is of order $a(p'-m)^{-1}(p-m)(p'-m)^{-1}b$ which is $a((1+\epsilon)p+m)(p-m)\times((1+\epsilon)p+m)p(2\epsilon+\epsilon^2)^{-2}m^{-4}$. But since $p^2=m^2$, we have $p(p-m)=-m(p-m)=(p-m)p$ so the net result is approximately $a(p-m)b/4m^2$ and is not of order $1/\epsilon$ but smaller, so that its effect drops out in the limit.

[15] We have used, to first order, the general expansion (valid for any operators A, B)

$$(A+B)^{-1}=A^{-1}-A^{-1}BA^{-1}+A^{-1}BA^{-1}BA^{-1}-\cdots$$

with $A=p-k-m$ and $B=p'-p=\epsilon p$ to expand the difference of $(p'-k-m)^{-1}$ and $(p-k-m)^{-1}$.

[16] The renormalization terms appearing B, Eqs. (14), (15) when translated directly into the present notation do not give twice (29) but give this expression with the central p_1m^{-1} factor replaced by $m\gamma_4/E_1$ where $E_1=p_{1\mu}$ for $\mu=4$. When integrated it therefore gives $ra((p_1+m)/2m)(m\gamma_4/E_1)$ or $ra-ra(m\gamma_4/E_1)(p_1-m)/2m$. (Since $p_1\gamma_4+\gamma_4p_1=2E_1$) which gives just ra, since $p_1u_1=mu_1$.

[17] H. W. Lewis, Phys. Rev. **73**, 173 (1948).

which we can assume to vary as $a_\mu \exp(-iq\cdot x)$ creates a pair of electrons (see Fig. 6), momenta p_a, $-p_b$. This pair then reannihilates, emitting a quantum $q=p_b-p_a$, which quantum scatters the original electron from state 1 to state 2. The matrix element for this process (and the others which can be obtained by rearranging the order in time of the various events) is

$$-(e^2/\pi i)(\bar{u}_2\gamma_\mu u_1)\int Sp[(p_a+q-m)^{-1}\times\gamma_\nu(p_a-m)^{-1}\gamma_\mu]d^4p_a q^{-2}C(q^2)a_\nu. \quad (30)$$

This is because the potential produces the pair with amplitude proportional to $a_\nu\gamma_\nu$, the electrons of momenta p_a and $-(p_a+q)$ proceed from there to annihilate, producing a quantum (factor γ_μ) which propagates (factor $q^{-2}C(q^2)$) over to the other electron, by which it is absorbed (matrix element of γ_μ between states 1 and 2 of the original electron $(\bar{u}_2\gamma_\mu u_1)$). All momenta p_a and spin states of the virtual electron are admitted, which means the spur and the integral on d^4p_a are calculated.

One can imagine that the closed loop path of the positron-electron produces a current

$$4\pi j_\mu = J_{\mu\nu}a_\nu, \quad (31)$$

which is the source of the quanta which act on the second electron. The quantity

$$J_{\mu\nu} = -(e^2/\pi i)\int Sp[(p+q-m)^{-1}\times\gamma_\nu(p-m)^{-1}\gamma_\mu]d^4p, \quad (32)$$

is then characteristic for this problem of polarization of the vacuum.

One sees at once that $J_{\mu\nu}$ diverges badly. The modification of δ to f alters the amplitude with which the current j_μ will affect the scattered electron, but it can do nothing to prevent the divergence of the integral (32) and of its effects.

One way to avoid such difficulties is apparent. From one point of view we are considering all routes by which a given electron can get from one region of space-time to another, i.e., from the source of electrons to the apparatus which measures them. From this point of view the closed loop path leading to (32) is unnatural. It might be assumed that the only paths of meaning are those which start from the source and work their way in a continuous path (possibly containing many time reversals) to the detector. Closed loops would be excluded. We have already found that this may be done for electrons moving in a fixed potential.

Such a suggestion must meet several questions, however. The closed loops are a consequence of the usual hole theory in electrodynamics. Among other things, they are required to keep probability conserved. The probability that no pair is produced by a potential is not unity and its deviation from unity arises from the imaginary part of $J_{\mu\nu}$. Again, with closed loops excluded, a pair of electrons once created cannot annihilate one another again, the scattering of light by light would be zero, etc. Although we are not experimentally sure of these phenomena, this does seem to indicate that the closed loops are necessary. To be sure, it is always possible that these matters of probability conservation, etc., will work themselves out as simply in the case of interacting particles as for those in a fixed potential. Lacking such a demonstration the presumption is that the difficulties of vacuum polarization are not so easily circumvented.[18]

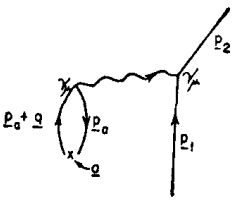


FIG. 6. Vacuum polarization effect on scattering, Eq. (30).

An alternative procedure discussed in B is to assume that the function $K_+(2, 1)$ used above is incorrect and is to be replaced by a modified function K_+' having no singularity on the light cone. The effect of this is to provide a convergence factor $C(p^2-m^2)$ for *every* integral over electron momenta.[19] This will multiply the integrand of (32) by $C(p^2-m^2)C((p+q)^2-m^2)$, since the integral was originally $\delta(p_a-p_b+q)d^4p_a d^4p_b$ and both p_a and p_b get convergence factors. The integral now converges but the result is unsatisfactory.[20]

One expects the current (31) to be conserved, that is $q_\mu j_\mu=0$ or $q_\mu J_{\mu\nu}=0$. Also one expects no current if a_ν is a gradient, or $a_\nu=q_\nu$ times a constant. This leads to the condition $J_{\mu\nu}q_\nu=0$ which is equivalent to $q_\mu J_{\mu\nu}=0$ since $J_{\mu\nu}$ is symmetrical. But when the expression (32) is integrated with such convergence factors it does not satisfy this condition. By altering the kernel from K to another, K', which does not satisfy the Dirac equation we have lost the gauge invariance, its consequent current conservation and the general consistency of the theory.

One can see this best by calculating $J_{\mu\nu}q_\nu$ directly from (32). The expression within the spur becomes $(p+q-m)^{-1}q(p-m)^{-1}\gamma_\mu$ which can be written as the difference of two terms: $(p-m)^{-1}\gamma_\mu-(p+q-m)^{-1}\gamma_\mu$. Each of these terms would give the same result if the integration d^4p were without a convergence factor, for

[18] It would be very interesting to calculate the Lamb shift accurately enough to be sure that the 20 megacycles expected from vacuum polarization are actually present.

[19] This technique also makes self-energy and radiationless scattering integrals finite even without the modification of δ_+ to f_+ for the radiation (and the consequent convergence factor $C(k^2)$ for the quanta). See B.

[20] Added to the terms given below (33) there is a term $\frac{1}{4}(\lambda^2-2\mu^2+\frac{1}{3}q^2)\delta_{\mu\nu}$ for $C(k^2)=-\lambda^2(k^2-\lambda^2)^{-1}$, which is not gauge invariant. (In addition the charge renormalization has $-7/6$ added to the logarithm.)

the first can be converted into the second by a shift of the origin of p, namely $p'=p+q$. This does not result in cancelation in (32) however, for the convergence factor is altered by the substitution.

A method of making (32) convergent without spoiling the gauge invariance has been found by Bethe and by Pauli. The convergence factor for light can be looked upon as the result of superposition of the effects of quanta of various masses (some contributing negatively). Likewise if we take the factor $C(p^2-m^2)=-\lambda^2(p^2-m^2-\lambda^2)^{-1}$ so that $(p^2-m^2)^{-1}C(p^2-m^2)=(p^2-m^2)^{-1}-(p^2-m^2-\lambda^2)^{-1}$ we are taking the difference of the result for electrons of mass m and mass $(\lambda^2+m^2)^{\frac{1}{2}}$. But we have taken this difference for *each* propagation between interactions with photons. They suggest instead that once created with a certain mass the electron should continue to propagate with this mass through all the potential interactions until it closes its loop. That is if the quantity (32), integrated over some finite range of p, is called $J_{\mu\nu}(m^2)$ and the corresponding quantity over the same range of p, but with m replaced by $(m^2+\lambda^2)^{\frac{1}{2}}$ is $J_{\mu\nu}(m^2+\lambda^2)$ we should calculate

$$J_{\mu\nu}{}^P=\int_0^\infty [J_{\mu\nu}(m^2)-J_{\mu\nu}(m^2+\lambda^2)]G(\lambda)d\lambda, \quad (32')$$

the function $G(\lambda)$ satisfying $\int_0^\infty G(\lambda)d\lambda=1$ and $\int_0^\infty G(\lambda)\lambda^2 d\lambda=0$. Then in the expression for $J_{\mu\nu}{}^P$ the range of p integration can be extended to infinity as the integral now converges. The result of the integration using this method is the integral on $d\lambda$ over $G(\lambda)$ of (see Appendix C)

$$J_{\mu\nu}{}^P=-\frac{e^2}{\pi}(q_\mu q_\nu-\delta_{\mu\nu}q^2)\left(-\frac{1}{3}\ln\frac{\lambda^2}{m^2}-\left[\frac{4m^2+2q^2}{3q^2}\left(1-\frac{\theta}{\tan\theta}\right)-\frac{1}{9}\right]\right), \quad (33)$$

with $q^2=4m^2\sin^2\theta$.

The gauge invariance is clear, since $q_\mu(q_\mu q_\nu-q^2\delta_{\mu\nu})=0$. Operating (as it always will) on a potential of zero divergence the $(q_\mu q_\nu-\delta_{\mu\nu}q^2)a_\nu$ is simply $-q^2a_\mu$, the D'Alembertian of the potential, that is, the current producing the potential. The term $-\frac{1}{3}(\ln(\lambda^2/m^2))(q_\mu q_\nu-q^2\delta_{\mu\nu})$ therefore gives a current proportional to the current producing the potential. This would have the same effect as a change in charge, so that we would have a difference $\Delta(e^2)$ between e^2 and the experimentally observed charge, $e^2+\Delta(e^2)$, analogous to the difference between m and the observed mass. This charge depends logarithmically on the cut-off, $\Delta(e^2)/e^2=-(2e^2/3\pi)\ln(\lambda/m)$. After this renormalization of charge is made, no effects will be sensitive to the cut-off.

After this is done the final term remaining in (33), contains the usual effects[21] of polarization of the vacuum. It is zero for a free light quantum ($q^2=0$). For small q^2 it behaves as $(2/15)q^2$ (adding $-\frac{1}{5}$ to the logarithm in the Lamb effect). For $q^2>(2m)^2$ it is complex, the imaginary part representing the loss in amplitude required by the fact that the probability that no quanta are produced by a potential able to produce pairs ($(q^2)^{\frac{1}{2}}>2m$) decreases with time. (To make the necessary analytic continuation, imagine m to have a small negative imaginary part, so that $(1-q^2/4m^2)^{\frac{1}{2}}$ becomes $-i(q^2/4m^2-1)^{\frac{1}{2}}$ as q^2 goes from below to above $4m^2$. Then $\theta=\pi/2+iu$ where $\sinh u=+(q^2/4m^2-1)^{\frac{1}{2}}$, and $-1/\tan\theta=i\tanh u=+i(q^2-4m^2)^{\frac{1}{2}}(q^2)^{-\frac{1}{2}}$.)

Closed loops containing a number of quanta or potential interactions larger than two produce no trouble. Any loop with an odd number of interactions gives zero (I, reference 9). Four or more potential interactions give integrals which are convergent even without a convergence factor as is well known. The situation is analogous to that for self-energy. Once the simple problem of a single closed loop is solved there are no further divergence difficulties for more complex processes.[22]

8. LONGITUDINAL WAVES

In the usual form of quantum electrodynamics the longitudinal and transverse waves are given separate treatment. Alternately the condition $(\partial A_\mu/\partial x_\mu)\Psi=0$ is carried along as a supplementary condition. In the present form no such special considerations are necessary for we are dealing with the solutions of the equation $-\Box^2 A_\mu=4\pi j_\mu$ with a current j_μ which is conserved $\partial j_\mu/\partial x_\mu=0$. That means at least $\Box^2(\partial A_\mu/\partial x_\mu)=0$ and in fact our solution also satisfies $\partial A_\mu/\partial x_\mu=0$.

To show that this is the case we consider the amplitude for emission (real or virtual) of a photon and show that the divergence of this amplitude vanishes. The amplitude for emission for photons polarized in the μ direction involves matrix elements of γ_μ. Therefore what we have to show is that the corresponding matrix elements of $q_\mu\gamma_\mu=q$ vanish. For example, for a first order effect we would require the matrix element of q between two states p_1 and $p_2=p_1+q$. But since $q=p_2-p_1$ and $(\bar{u}_2 p_1 u_1)=m(\bar{u}_2 u_1)=(\bar{u}_2 p_2 u_1)$ the matrix element vanishes, which proves the contention in this case. It also vanishes in more complex situations (essentially because of relation (34), below) (for example, try putting $e_2=q_2$ in the matrix (15) for the Compton Effect).

To prove this in general, suppose a_i, $i=1$ to N are a set of plane wave disturbing potentials carrying momenta q_i (e.g., some may be emissions or absorptions of the same or different quanta) and consider a matrix for the transition from a state of momentum p_0 to p_N such

[21] E. A. Uehling, Phys. Rev. **48**, 55 (1935), R. Serber, Phys. Rev. **48**, 49 (1935).

[22] There are loops completely without external interactions. For example, a pair is created virtually along with a photon. Next they annihilate, absorbing this photon. Such loops are disregarded on the grounds that they do not interact with anything and are thereby completely unobservable. Any indirect effects they may have via the exclusion principle have already been included.

as $a_N\prod_{i=1}^{N-1}(p_i-m)^{-1}a_i$ where $p_i=p_{i-1}+q_i$ (and in the product, terms with larger i are written to the left). The most general matrix element is simply a linear combination of these. Next consider the matrix between states p_0 and p_N+q in a situation in which not only are the a_i acting but also another potential $a\exp(-iq\cdot x)$ where $a=q$. This may act previous to all a_i, in which case it gives $a_N\prod(p_i+q-m)^{-1}a_i(p_0+q-m)^{-1}q$ which is equivalent to $+a_N\prod(p_i+q-m)^{-1}a_i$ since $+(p_0+q-m)^{-1}q$ is equivalent to $(p_0+q-m)^{-1}\times(p_0+q-m)$ as p_0 is equivalent to m acting on the initial state. Likewise if it acts after all the potentials it gives $q(p_N-m)^{-1}a_N\prod(p_i-m)^{-1}a_i$ which is equivalent to $-a_N\prod(p_i-m)^{-1}a_i$ since p_N+q-m gives zero on the final state. Or again it may act between the potential a_k and a_{k+1} for each k. This gives

$$\sum_{k=1}^{N-1}a_N\prod_{i=k+1}^{N-1}(p_i+q-m)^{-1}a_i(p_k+q-m)^{-1}\times q(p_k-m)^{-1}a_k\prod_{j=1}^{k-1}(p_j-m)^{-1}a_j.$$

However,

$$(p_k+q-m)^{-1}q(p_k-m)^{-1}=(p_k-m)^{-1}-(p_k+q-m)^{-1}, \quad (34)$$

so that the sum breaks into the difference of two sums, the first of which may be converted to the other by the replacement of k by $k-1$. There remain only the terms from the ends of the range of summation,

$$+a_N\prod_{i=1}^{N-1}(p_i-m)^{-1}a_i-a_N\prod_{i=1}^{N-1}(p_i+q-m)^{-1}a_i.$$

These cancel the two terms originally discussed so that the entire effect is zero. Hence any wave emitted will satisfy $\partial A_\mu/\partial x_\mu=0$. Likewise longitudinal waves (that is, waves for which $A_\mu=\partial\phi/\partial x_\mu$ or $a=q$) cannot be absorbed and will have no effect, for the matrix elements for emission and absorption are similar. (We have said little more than that a potential $A_\mu=\partial\varphi/\partial x_\mu$ has no effect on a Dirac electron since a transformation $\psi'=\exp(-i\phi)\psi$ removes it. It is also easy to see in coordinate representation using integrations by parts.)

This has a useful practical consequence in that in computing probabilities for transition for unpolarized light one can sum the squared matrix over all four directions rather than just the two special polarization vectors. Thus suppose the matrix element for some process for light polarized in direction e_μ is $e_\mu M_\mu$. If the light has wave vector q_μ we know from the argument above that $q_\mu M_\mu=0$. For unpolarized light progressing in the z direction we would ordinarily calculate $M_x^2+M_y^2$. But we can as well sum $M_x^2+M_y^2+M_z^2-M_t^2$ for $q_\mu M_\mu$ implies $M_t=M_z$ since $q_t=q_z$ for free quanta. This shows that unpolarized light is a relativistically invariant concept, and permits some simplification in computing cross sections for such light.

Incidentally, the virtual quanta interact through terms like $\gamma_\mu\cdots\gamma_\mu k^{-2}d^4k$. Real processes correspond to poles in the formulae for virtual processes. The pole occurs when $k^2=0$, but it looks at first as though in the sum on all four values of μ, of $\gamma_\mu\cdots\gamma_\mu$ we would have four kinds of polarization instead of two. Now it is clear that only two perpendicular to k are effective.

The usual elimination of longitudinal and scalar virtual photons (leading to an instantaneous Coulomb potential) can of course be performed here too (although it is not particularly useful). A typical term in a virtual transition is $\gamma_\mu\cdots\gamma_\mu k^{-2}d^4k$ where the $\cdots$ represent some intervening matrices. Let us choose for the values of μ, the time l, the direction of vector part $\mathbf{K}$, of k, and two perpendicular directions 1, 2. We shall not change the expression for these two 1, 2 for these are represented by transverse quanta. But we must find $(\gamma_l\cdots\gamma_l)-(\gamma_\mathbf{K}\cdots\gamma_\mathbf{K})$. Now $k=k_4\gamma_l-K\gamma_\mathbf{K}$, where $K=(\mathbf{K}\cdot\mathbf{K})^{\frac{1}{2}}$, and we have shown above that k replacing the γ_μ gives zero.[23] Hence $K\gamma_\mathbf{K}$ is equivalent to $k_4\gamma_l$ and

$$(\gamma_l\cdots\gamma_l)-(\gamma_\mathbf{K}\cdots\gamma_\mathbf{K})=((K^2-k_4^2)/K^2)(\gamma_l\cdots\gamma_l),$$

so that on multiplying by $k^{-2}d^4k=d^4k(k_4^2-K^2)^{-1}$ the net effect is $-(\gamma_l\cdots\gamma_l)d^4k/K^2$. The γ_l means just scalar waves, that is, potentials produced by charge density. The fact that $1/K^2$ does not contain k_4 means that k_4 can be integrated first, resulting in an instantaneous interaction, and the $d^3\mathbf{K}/K^2$ is just the momentum representation of the Coulomb potential, $1/r$.

9. KLEIN GORDON EQUATION

The methods may be readily extended to particles of spin zero satisfying the Klein Gordon equation,[24]

$$\Box^2\psi-m^2\psi=i\partial(A_\mu\psi)/\partial x_\mu+iA_\mu\partial\psi/\partial x_\mu-A_\mu A_\mu\psi. \quad (35)$$

[23] A little more care is required when both γ_μ's act on the same particle. Define $x=k_4\gamma_l+K\gamma_\mathbf{K}$, and consider $(k\cdots x)+(x\cdots k)$. Exactly this term would arise if a system, each one by potential x carrying momentum $-k$, is disturbed by an added potential k of *momentum* $+k$ (*the reversed sign of the momenta in the intermediate factors in the second term* $x\cdots k$ has no effect since we will later integrate over all k). Hence as shown above the result is zero, but since $(k\cdots x)+(x\cdots k)=k_4^2(\gamma_l\cdots\gamma_l)-K^2(\gamma_\mathbf{K}\cdots\gamma_\mathbf{K})$ we can still conclude $(\gamma_\mathbf{K}\cdots\gamma_\mathbf{K})=k_4^2K^{-2}(\gamma_l\cdots\gamma_l)$.

[24] The equations discussed in this section were deduced from the formulation of the Klein Gordon equation given in reference 5, Section 14. The function ψ in this section has only one component and is not a spinor. An alternative formal method of making the equations valid for spin zero and also for spin 1 is (presumably) by use of the Kemmer-Duffin matrices β_μ, satisfying the commutation relation

$$\beta_\mu\beta_\nu\beta_\sigma+\beta_\sigma\beta_\nu\beta_\mu=\delta_{\mu\nu}\beta_\sigma+\delta_{\sigma\nu}\beta_\mu.$$

If we interpret a to mean $a_\mu\beta_\mu$, rather than $a_\mu\gamma_\mu$, for any a_μ, all of the equations in momentum space will remain formally identical to those for the spin 1/2; with the exception of those in which a denominator $(p-m)^{-1}$ has been rationalized to $(p+m)(p^2-m^2)^{-1}$ since p^2 is no longer equal to a number, $p\cdot p$. But p^3 does equal $(p\cdot p)p$ so that $(p-m)^{-1}$ may now be interpreted as $(mp+m^2+p^2-p\cdot p)(p\cdot p-m^2)^{-1}m^{-1}$. This implies that equations in coordinate space will be valid of the function $K_+(2,1)$ is given as $K_+(2,1)=[(i\nabla_2+m)-m^{-1}(\nabla_2^2+\Box_2^2)]iI_+(2,1)$ with $\nabla_2=\beta_\mu\partial/\partial x_{2\mu}$. This is all in virtue of the fact that the many component wave function ψ (5 components for spin 0, 10 for spin 1) satisfies $(i\nabla-m)\psi=A\psi$ which is formally identical to the Dirac Equation. See W. Pauli, Rev. Mod. Phys. 13, 203 (1940).

The important kernel is now $I_+(2,1)$ defined in (I, Eq. (32)). For a free particle, the wave function $\psi(2)$ satisfies $+\Box^2\psi - m^2\psi = 0$. At a point, 2, inside a space time region it is given by

$$\psi(2) = \int [\psi(1)\partial I_+(2,1)/\partial x_{1\mu} - (\partial\psi/\partial x_{1\mu}) I_+(2,1)] N_\mu(1) d^3 1',$$

(as is readily shown by the usual method of demonstrating Green's theorem) the integral being over an entire 3-surface boundary of the region (with normal vector N_μ). Only the positive frequency components of ψ contribute from the surface preceding the time corresponding to 2, and only negative frequencies from the surface future to 2. These can be interpreted as electrons and positrons in direct analogy to the Dirac case.

The right-hand side of (35) can be considered as a source of new waves and a series of terms written down to represent matrix elements for processes of increasing order. There is only one new point here, the term in $A_\mu A_\mu$ by which two quanta can act at the same time. As an example, suppose three quanta or potentials, $a_\mu \exp(-iq_a\cdot x)$, $b_\mu \exp(-iq_b\cdot x)$, and $c_\mu \exp(-iq_c\cdot x)$ are to act in that order on a particle of original momentum $p_{0\mu}$ so that $p_a = p_0 + q_a$ and $p_b = p_a + q_b$; the final momentum being $p_c = p_b + q_c$. The matrix element is the sum of three terms ($p^2 = p_\mu p_\mu$) (illustrated in Fig. 7)

$$\begin{aligned}&(p_c\cdot c + p_b\cdot c)(p_b^2 - m^2)^{-1}(p_b\cdot b + p_a\cdot b)\\ &\qquad\times(p_a^2 - m^2)^{-1}(p_a\cdot a + p_0\cdot a)\\ &-(p_c\cdot c + p_b\cdot c)(p_b^2 - m^2)^{-1}(b\cdot a)\\ &-(c\cdot b)(p_a^2 - m^2)^{-1}(p_a\cdot a + p_0\cdot a).\end{aligned} \tag{36}$$

The first comes when each potential acts through the perturbation $i\partial(A_\mu\psi)/\partial x_\mu + iA_\mu\partial\psi/\partial x_\mu$. These gradient operators in momentum space mean respectively the momentum after and before the potential A_μ operates. The second term comes from b_μ and a_μ acting at the same instant and arises from the $A_\mu A_\mu$ term in (a). Together b_μ and a_μ carry momentum $q_{b\mu} + q_{a\mu}$ so that after $b\cdot a$ operates the momentum is $p_0 + q_a + q_b$ or p_b. The final term comes from c_μ and b_μ operating together in a similar manner. The term $A_\mu A_\mu$ thus permits a new type of process in which two quanta can be emitted (or absorbed, or one absorbed, one emitted) at the same time. There is no $a\cdot c$ term for the order a, b, c we have assumed. In an actual problem there would be other terms like (36) but with alterations in the order in which the quanta a, b, c act. In these terms $a\cdot c$ would appear.

As a further example the self-energy of a particle of momentum p_μ is

$$(e^2/2\pi i m)\int [(2p-k)_\mu((p-k)^2 - m^2)^{-1} \times (2p-k)_\mu - \delta_{\mu\mu}] d^4k k^{-2} C(k^2),$$

where the $\delta_{\mu\mu} = 4$ comes from the $A_\mu A_\mu$ term and represents the possibility of the simultaneous emission and absorption of the same virtual quantum. This integral without the $C(k^2)$ diverges quadratically and would not converge if $C(k^2) = -\lambda^2/(k^2 - \lambda^2)$. Since the interaction occurs through the gradients of the potential, we must use a stronger convergence factor, for example $C(k^2) = \lambda^4(k^2-\lambda^2)^{-2}$, or in general (17) with $\int_0^\infty \lambda^2 G(\lambda) d\lambda = 0$. In this case the self-energy converges but depends quadratically on the cut-off λ and is not necessarily small compared to m. The radiative corrections to scattering after mass renormalization are insensitive to the cut-off just as for the Dirac equation.

When there are several particles one can obtain Bose statistics by the rule that if two processes lead to the same state but with two electrons exchanged, their amplitudes are to be added (rather than subtracted as for Fermi statistics). In this case equivalence to the second quantization treatment of Pauli and Weisskopf should be demonstrable in a way very much like that given in I (appendix) for Dirac electrons. The Bose statistics mean that the sign of contribution of a closed loop to the vacuum polarization is the opposite of what it is for the Fermi case (see I). It is ($p_b = p_a + q$)

$$J_{\mu\nu} = \frac{e^2}{2\pi i m}\int [(p_{b\mu} + p_{a\mu})(p_{b\nu} + p_{a\nu})(p_a^2 - m^2)^{-1} \times (p_b^2 - m^2)^{-1} - \delta_{\mu\nu}(p_a^2 - m^2)^{-1} - \delta_{\mu\nu}(p_b^2 - m^2)^{-1}] d^4 p_a$$

giving,

$$J_{\mu\nu}{}^P = \frac{e^2}{\pi}(q_\mu q_\nu - \delta_{\mu\nu}q^2)\left[\frac{1}{6}\ln\frac{\lambda^2}{m^2} + \frac{1}{9} - \frac{4m^2 - q^2}{3q^2}\left(1 - \frac{\theta}{\tan\theta}\right)\right],$$

the notation as in (33). The imaginary part for $(q^2)^{\frac{1}{2}} > 2m$ is again positive representing the loss in the probability of finding the final state to be a vacuum, associated with the possibilities of pair production. Fermi statistics would give a gain in probability (and also a charge renormalization of opposite sign to that expected).

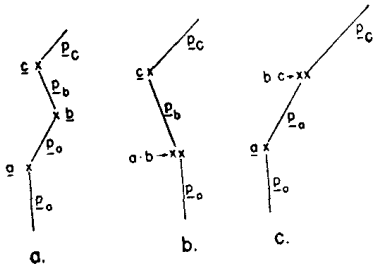


FIG. 7. Klein-Gordon particle in three potentials, Eq. (36). The coupling to the electromagnetic field is now, for example, $p_0\cdot a + p_a\cdot a$, and a new possibility arises, (b), of simultaneous interaction with two quanta $a\cdot b$. The propagation factor is now $(p\cdot p - m^2)^{-1}$ for a particle of momentum p_μ.

10. APPLICATION TO MESON THEORIES

The theories which have been developed to describe mesons and the interaction of nucleons can be easily expressed in the language used here. Calculations, to lowest order in the interactions can be made very easily for the various theories, but agreement with experimental results is not obtained. Most likely all of our present formulations are quantitatively unsatisfactory. We shall content ourselves therefore with a brief summary of the methods which can be used.

The nucleons are usually assumed to satisfy Dirac's equation so that the factor for propagation of a nucleon of momentum p is $(p-M)^{-1}$ where M is the mass of the nucleon (which implies that nucleons can be created in pairs). The nucleon is then assumed to interact with mesons, the various theories differing in the form assumed for this interaction.

First, we consider the case of neutral mesons. The theory closest to electrodynamics is the theory of vector mesons with vector coupling. Here the factor for emission or absorption of a meson is $g\gamma_\mu$ when this meson is "polarized" in the μ direction. The factor g, the "mesonic charge," replaces the electric charge e. The amplitude for propagation of a meson of momentum q in intermediate states is $(q^2-\mu^2)^{-1}$ (rather than q^{-2} as it is for light) where μ is the mass of the meson. The necessary integrals are made finite by convergence factors $C(q^2-\mu^2)$ as in electrodynamics. For scalar mesons with scalar coupling the only change is that one replaces the γ_μ by 1 in emission and absorption. There is no longer a direction of polarization, μ, to sum upon. For pseudoscalar mesons, pseudoscalar coupling replace γ_μ by $\gamma_5=i\gamma_x\gamma_y\gamma_z\gamma_t$. For example, the self-energy matrix of a nucleon of momentum p in this theory is

$$(g^2/\pi i)\int \gamma_5(p-k-M)^{-1}\gamma_5 d^4k(k^2-\mu^2)^{-1}C(k^2-\mu^2).$$

Other types of meson theory result from the replacement of γ_μ by other expressions (for example by $\frac{1}{2}(\gamma_\mu\gamma_\nu-\gamma_\nu\gamma_\mu)$ with a subsequent sum over all μ and ν for virtual mesons). Scalar mesons with vector coupling result from the replacement of γ_μ by $\mu^{-1}q$ where q is the final momentum of the nucleon minus its initial momentum, that is, it is the momentum of the meson if absorbed, or the negative of the momentum of a meson emitted. As is well known, this theory with neutral mesons gives zero for all processes, as is proved by our discussion on longitudinal waves in electrodynamics. Pseudoscalar mesons with pseudo-vector coupling corresponds to γ_μ being replaced by $\mu^{-1}\gamma_5 q$ while vector mesons with tensor coupling correspond to using $(2\mu)^{-1}(\gamma_\mu q-q\gamma_\mu)$. These extra gradients involve the danger of producing higher divergencies for real processes. For example, $\gamma_5 q$ gives a logarithmically divergent interaction of neutron and electron.[25] Although these divergencies can be held by strong enough convergence factors, the results then are sensitive to the method used for convergence and the size of the cut-off values of λ. For low order processes $\mu^{-1}\gamma_5 q$ is equivalent to the pseudoscalar interaction $2M\mu^{-1}\gamma_5$ because if taken between free particle wave functions of the nucleon of momenta p_1 and $p_2=p_1+q$, we have

$$(\bar{u}_2\gamma_5 q u_1)=(\bar{u}_2\gamma_5(p_2-p_1)u_1)=-(\bar{u}_2 p_2\gamma_5 u_1)-(\bar{u}_2\gamma_5 p_1 u_1)=-2M(\bar{u}_2\gamma_5 u_1)$$

since γ_5 anticommutes with p_2 and p_2 operating on the state 2 equivalent to M as is p_1 on the state 1. This shows that the γ_5 interaction is unusually weak in the non-relativistic limit (for example the expected value of γ_5 for a free nucleon is zero), but since $\gamma_5^2=1$ is not small, pseudoscalar theory gives a more important interaction in second order than it does in first. Thus the pseudoscalar coupling constant should be chosen to fit nuclear forces including these important second order processes.[26] The equivalence of pseudoscalar and pseudovector coupling which holds for low order processes therefore does not hold when the pseudoscalar theory is giving its most important effects. These theories will therefore give quite different results in the majority of practical problems.

In calculating the corrections to scattering of a nucleon by a neutral vector meson field (γ_μ) due to the effects of virtual mesons, the situation is just as in electrodynamics, in that the result converges without need for a cut-off and depends only on gradients of the meson potential. With scalar (1) or pseudoscalar (γ_5) neutral mesons the result diverges logarithmically and so must be cut off. The part sensitive to the cut-off, however, is directly proportional to the meson potential. It may thereby be removed by a renormalization of mesonic charge g. After this renormalization the results depend only on gradients of the meson potential and are essentially independent of cut-off. This is in addition to the mesonic charge renormalization coming from the production of virtual nucleon pairs by a meson, analogous to the vacuum polarization in electrodynamics. But here there is a further difference from electrodynamics for scalar or pseudoscalar mesons in that the polarization also gives a term in the induced current proportional to the meson potential representing therefore an additional renormalization of the *mass of the meson* which usually depends quadratically on the cut-off.

Next consider charged mesons in the absence of an electromagnetic field. One can introduce isotopic spin operators in an obvious way. (Specifically replace the neutral γ_5, say, by $\tau_i\gamma_5$ and sum over $i=1, 2$ where $\tau_1=\tau_++\tau_-$, $\tau_2=i(\tau_+-\tau_-)$ and τ_+ changes neutron to proton (τ_+ on proton$=0$) and τ_- changes proton to neutron.) It is just as easy for practical problems simply to keep track of whether the particle is a proton or a neutron on a diagram drawn to help write down the

[25] M. Slotnick and W. Heitler, Phys. Rev. 75, 1645 (1949).

[26] H. A. Bethe, Bull. Am. Phys. Soc. 24, 3, Z3 (Washington, 1949).

matrix element. This excludes certain processes. For example in the scattering of a negative meson from q_1 to q_2 by a neutron, the meson q_2 must be emitted first (in order of operators, not time) for the neutron cannot absorb the negative meson q_1 until it becomes a proton. That is, in comparison to the Klein Nishina formula (15), only the analogue of second term (see Fig. 5(b)) would appear in the scattering of negative mesons by neutrons, and only the first term (Fig. 5(a)) in the neutron scattering of positive mesons.

The source of mesons of a given charge is not conserved, for a neutron capable of emitting negative mesons may (on emitting one, say) become a proton no longer able to do so. The proof that a perturbation q gives zero, discussed for longitudinal electromagnetic waves, fails. This has the consequence that vector mesons, if represented by the interaction γ_μ would not satisfy the condition that the divergence of the potential is zero. The interaction is to be taken[27] as $\gamma_\mu - \mu^{-2}q_\mu q$ in emission and as γ_μ in absorption if the real emission of mesons with a non-zero divergence of potential is to be avoided. (The correction term $\mu^{-2}q_\mu q$ gives zero in the neutral case.) The asymmetry in emission and absorption is only apparent, as this is clearly the same thing as subtracting from the original $\gamma_\mu \cdots \gamma_\mu$, a term $\mu^{-2}q\cdots q$. That is, if the term $-\mu^{-2}q_\mu q$ is omitted the resulting theory describes a combination of mesons of spin one and spin zero. The spin zero mesons, coupled by vector coupling q, are removed by subtracting the term $\mu^{-2}q\cdots q$.

The two extra gradients $q\cdots q$ make the problem of diverging integrals still more serious (for example the interaction between two protons corresponding to the exchange of two charged vector mesons depends quadratically on the cut-off if calculated in a straightforward way). One is tempted in this formulation to choose simply $\gamma_\mu \cdots \gamma_\mu$ and accept the admixture of spin zero mesons. But it appears that this leads in the conventional formalism to negative energies for the spin zero component. This shows one of the advantages of the method of second quantization of meson fields over the present formulation. There such errors of sign are obvious while here we seem to be able to write seemingly innocent expressions which can give absurd results. Pseudovector mesons with pseudovector coupling correspond to using $\gamma_5(\gamma_\mu - \mu^{-2}q_\mu q)$ for absorption and $\gamma_5\gamma_\mu$ for emission for both charged and neutral mesons.

In the presence of an electromagnetic field, whenever the nucleon is a proton it interacts with the field in the way described for electrons. The meson interacts in the scalar or pseudoscalar case as a particle obeying the Klein-Gordon equation. It is important here to use the method of calculation of Bethe and Pauli, that is, a virtual meson is assumed to have the same "mass" during all its interactions with the electromagnetic field. The result for mass μ and for $(\mu^2+\lambda^2)^{\frac{1}{2}}$ are subtracted and the difference integrated over the function $G(\lambda)d\lambda$. A separate convergence factor is not provided for each meson propagation between electromagnetic interactions, otherwise gauge invariance is not insured. When the coupling involves a gradient, such as $\gamma_5 q$ where q is the final minus the initial momentum of the nucleon, the vector potential A must be subtracted from the momentum of the proton. That is, there is an additional coupling $\pm\gamma_5 A$ (plus when going from proton to neutron, minus for the reverse) representing the new possibility of a simultaneous emission (or absorption) of meson and photon.

Emission of positive or absorption of negative virtual mesons are represented in the same term, the sign of the charge being determined by temporal relations as for electrons and positrons.

Calculations are very easily carried out in this way to lowest order in g^2 for the various theories for nucleon interaction, scattering of mesons by nucleons, meson production by nuclear collisions and by gamma-rays, nuclear magnetic moments, neutron electron scattering, etc., However, no good agreement with experiment results, when these are available, is obtained. Probably all of the formulations are incorrect. An uncertainty arises since the calculations are only to first order in g^2, and are not valid if g^2/hc is large.

The author is particularly indebted to Professor H. A. Bethe for his explanation of a method of obtaining finite and gauge invariant results for the problem of vacuum polarization. He is also grateful for Professor Bethe's criticisms of the manuscript, and for innumerable discussions during the development of this work. He wishes to thank Professor J. Ashkin for his careful reading of the manuscript.

APPENDIX

In this appendix a method will be illustrated by which the simpler integrals appearing in problems in electrodynamics can be directly evaluated. The integrals arising in more complex processes lead to rather complicated functions, but the study of the relations of one integral to another and their expression in terms of simpler integrals may be facilitated by the methods given here.

[27] The vector meson field potentials φ_μ satisfy

$$-\partial/\partial x_\nu(\partial\varphi_\mu/\partial x_\nu - \partial\varphi_\nu/\partial x_\mu) - \mu^2\varphi_\mu = -4\pi s_\mu,$$

where s_μ, the source for such mesons, is the matrix element of γ_μ between states of neutron and proton. By taking the divergence $\partial/\partial x_\mu$ of both sides, conclude that $\partial\varphi_\nu/\partial x_\nu = 4\pi\mu^{-2}\partial s_\nu/\partial x_\nu$ so that the original equation can be rewritten as

$$\Box^2\varphi_\mu - \mu^2\varphi_\mu = -4\pi(s_\mu + \mu^{-2}\partial/\partial x_\mu(\partial s_\nu/\partial x_\nu)).$$

The right hand side gives in momentum representation $\gamma_\mu - \mu^{-2}q_\mu q_\nu\gamma_\nu$ the left yields the $(q^2-\mu^2)^{-1}$ and finally the interaction $s_\mu\varphi_\mu$ in the Lagrangian gives the γ_μ on absorption.

Proceeding in this way find generally that particles of spin one can be represented by a four-vector u_μ (which, for a free particle of momentum q satisfies $q\cdot u = 0$). The propagation of virtual particles of momentum q from state ν to μ is represented by multiplication by the 4-4 matrix (or tensor) $P_{\mu\nu} = (\delta_{\mu\nu} - \mu^{-2}q_\mu q_\nu) \times (q^2-\mu^2)^{-1}$. The first-order interaction (from the Proca equation) with an electromagnetic potential $a\exp(-ik\cdot x)$ corresponds to multiplication by the matrix $E_{\mu\nu} = (q_2\cdot a + q_1\cdot a)\delta_{\mu\nu} - q_{2\nu}a_\mu - q_{1\mu}a_\nu$ where q_1 and $q_2 = q_1 + k$ are the momenta before and after the interaction. Finally, two potentials a, b may act simultaneously, with matrix $E'_{\mu\nu} = -(a\cdot b)\delta_{\mu\nu} + b_\mu a_\nu$.

As a typical problem consider the integral (12) appearing in the first order radiationless scattering problem:

$$\int \gamma_\mu(\not p_2-\not k-m)^{-1}a(\not p_1-\not k-m)^{-1}\gamma_\mu k^{-2}d^4kC(k^2), \qquad (1a)$$

where we shall take $C(k^2)$ to be typically $-\lambda^2(k^2-\lambda^2)^{-1}$ and d^4k means $(2\pi)^{-2}dk_1dk_2dk_3dk_4$. We first rationalize the factors $(\not p-\not k-m)^{-1}=(\not p-\not k+m)((p-k)^2-m^2)^{-1}$ obtaining,

$$\int \gamma_\mu(\not p_2-\not k+m)a(\not p_1-\not k+m)\gamma_\mu k^{-2}d^4kC(k^2) \times((p_1-k)^2-m^2)^{-1}((p_2-k)^2-m^2)^{-1}. \qquad (2a)$$

The matrix expression may be simplified. It appears to be best to do so *after* the integrations are performed. Since $AB=2A\cdot B-BA$ where $A\cdot B=A_\mu B_\mu$ is a number commuting with all matrices, find, if R is any expression, and A a vector, since $\gamma_\mu A=-A\gamma_\mu+2A_\mu$,

$$\gamma_\mu AR\gamma_\mu=-A\gamma_\mu R\gamma_\mu+2RA. \qquad (3a)$$

Expressions between two γ_μ's can be thereby reduced by induction. Particularly useful are

$$\begin{aligned}\gamma_\mu\gamma_\mu&=4\\ \gamma_\mu A\gamma_\mu&=-2A\\ \gamma_\mu AB\gamma_\mu&=2(AB+BA)=4A\cdot B\\ \gamma_\mu ABC\gamma_\mu&=-2CBA\end{aligned} \qquad (4a)$$

where A, B, C are any three vector-matrices (i.e., linear combinations of the four γ's).

In order to calculate the integral in (2a) the integral may be written as the sum of three terms (since $\not k=k_\sigma\gamma_\sigma$),

$$\gamma_\mu(\not p_2+m)a(\not p_1+m)\gamma_\mu J_1-[\gamma_\mu\gamma_\sigma a(\not p_1+m)\gamma_\mu +\gamma_\mu(\not p_2+m)a\gamma_\sigma\gamma_\mu]J_2+\gamma_\mu\gamma_\sigma a\gamma_\tau\gamma_\mu J_3, \qquad (5a)$$

where

$$J_{(1;2;3)}=\int(1;k_\sigma;k_\sigma k_\tau)k^{-2}d^4kC(k^2) \times((p_2-k)^2-m^2)^{-1}((p_1-k)^2-m^2)^{-1}. \qquad (6a)$$

That is for J_1 the $(1;k_\sigma;k_\sigma k_\tau)$ is replaced by 1, for J_2 by k_σ, and for J_3 by $k_\sigma k_\tau$.

More complex processes of the first order involve more factors like $((p_3-k)^2-m^2)^{-1}$ and a corresponding increase in the number of k's which may appear in the numerator, as $k_\sigma k_\tau k_\nu\cdots$. Higher order processes involving two or more virtual quanta involve similar integrals but with factors possibly involving $k+k'$ instead of just k, and the integral extending on $k^{-2}d^4kC(k^2)k'^{-2}d^4k'C(k'^2)$. They can be simplified by methods analogous to those used on the first order integrals.

The factors $(\not p-\not k)^2-m^2$ may be written

$$(\not p-\not k)^2-m^2=k^2-2p\cdot k-\Delta, \qquad (7a)$$

where $\Delta=m^2-p^2$, $\Delta_1=m_1^2-p_1^2$, etc., and we can consider dealing with cases of greater generality in that the different denominators need not have the same value of the mass m. In our specific problem (6a), $p_1^2=m^2$ so that $\Delta_1=0$, but we desire to work with greater generality.

Now for the factor $C(k^2)/k^2$ we shall use $-\lambda^2(k^2-\lambda^2)^{-1}k^{-2}$. This can be written as

$$-\lambda^2/(k^2-\lambda^2)k^2=k^{-2}C(k^2)=-\int_0^{\lambda^2}dL(k^2-L)^{-2}. \qquad (8a)$$

Thus we can replace $k^{-2}C(k^2)$ by $(k^2-L)^{-2}$ and at the end integrate the result with respect to L from zero to λ^2. We can for many practical purposes consider λ^2 very large relative to m^2 or p^2. When the original integral converges even without the convergence factor, it will be obvious since the L integration will then be convergent to infinity. If an infra-red catastrophe exists in the integral one can simply assume quanta have a small mass $\lambda_{\min}$ and extend the integral on L from $\lambda^2_{\min}$ to λ^2, rather than from zero to λ^2.

We then have to do integrals of the form

$$\int(1;k_\sigma;k_\sigma k_\tau)d^4k(k^2-L)^{-2}(k^2-2p_1\cdot k-\Delta_1)^{-1} \times(k^2-2p_2\cdot k-\Delta_2)^{-1}, \qquad (9a)$$

where by $(1;k_\sigma;k_\sigma k_\tau)$ we mean that in the place of this symbol either 1, or k_σ, or $k_\sigma k_\tau$ may stand in different cases. In more complicated problems there may be more factors $(k^2-2p_i\cdot k-\Delta_i)^{-1}$ or other powers of these factors (the $(k^2-L)^{-2}$ may be considered as a special case of such a factor with $p_i=0$, $\Delta_i=L$) and further factors like $k_\sigma k_\tau k_\rho\cdots$ in the numerator. The poles in all the factors are made definite by the assumption that L, and the Δ's have infinitesimal negative imaginary parts.

We shall do the integrals of successive complexity by induction. We start with the simplest convergent one, and show

$$\int d^4k(k^2-L)^{-3}=(8iL)^{-1}. \qquad (10a)$$

For this integral is $\int(2\pi)^{-2}dk_4d^3\mathbf{K}(k_4^2-\mathbf{K}\cdot\mathbf{K}-L)^{-3}$ where the vector $\mathbf{K}$, of magnitude $K=(\mathbf{K}\cdot\mathbf{K})^{\frac12}$ is k_1, k_2, k_3. The integral on k_4 shows third order poles at $k_4=+(K^2+L)^{\frac12}$ and $k_4=-(K^2+L)^{\frac12}$. Imagining, in accordance with our definitions, that L has a small negative imaginary part only the first is below the real axis. The contour can be closed by an infinite semi-circle below this axis, without change of the value of the integral since the contribution from the semi-circle vanishes in the limit. Thus the contour can be shrunk about the pole $k_4=+(K^2+L)^{\frac12}$ and the resulting k_4 integral is $-2\pi i$ times the residue at this pole. Writing $k_4=(K^2+L)^{\frac12}+\epsilon$ and expanding $(k_4^2-K^2-L)^{-3}=\epsilon^{-3}(\epsilon+2(K^2+L)^{\frac12})^{-3}$ in powers of ϵ, the residue, being the coefficient of the term ϵ^{-1}, is seen to be $6(2(K^2+L)^{\frac12})^{-5}$ so our integral is

$$-(3i/32\pi)\int_0^\infty 4\pi K^2dK(K^2+L)^{-5/2}=(3/8i)(1/3L)$$

establishing (10a).

We also have $\int k_\sigma d^4k(k^2-L)^{-3}=0$ from the symmetry in the k space. We write these results as

$$(8i)\int(1;k_\sigma)d^4k(k^2-L)^{-3}=(1;0)L^{-1}, \qquad (11a)$$

where in the brackets $(1;k_\sigma)$ and $(1;0)$ corresponding entries are to be used.

Substituting $k=k'-p$ in (11a), and calling $L-p^2=\Delta$ shows that

$$(8i)\int(1;k_\sigma)d^4k(k^2-2p\cdot k-\Delta)^{-3}=(1;p_\sigma)(p^2+\Delta)^{-1}. \qquad (12a)$$

By differentiating both sides of (12a) with respect to Δ, or with respect to p_τ there follows directly

$$(24i)\int(1;k_\sigma;k_\sigma k_\tau)d^4k(k^2-2p\cdot k-\Delta)^{-4} =-(1;p_\sigma;p_\sigma p_\tau-\tfrac12\delta_{\sigma\tau}(p^2+\Delta))(p^2+\Delta)^{-2}. \qquad (13a)$$

Further differentiations give directly successive integrals including more k factors in the numerator and higher powers of $(k^2-2p\cdot k-\Delta)$ in the denominator.

The integrals so far only contain one factor in the denominator. To obtain results for two factors we make use of the identity

$$a^{-1}b^{-1}=\int_0^1 dx(ax+b(1-x))^{-2}, \qquad (14a)$$

(suggested by some work of Schwinger's involving Gaussian integrals). This represents the product of two reciprocals as a parametric integral over one and will therefore permit integrals with two factors to be expressed in terms of one. For other powers of a, b, we make use of all of the identities, such as

$$a^{-2}b^{-1}=\int_0^1 2xdx(ax+b(1-x))^{-3}, \qquad (15a)$$

deducible from (14a) by successive differentiations with respect to a or b.

To perform an integral, such as

$$(8i)\int(1;k_\sigma)d^4k(k^2-2p_1\cdot k-\Delta_1)^{-2}(k^2-2p_2\cdot k-\Delta_2)^{-1}, \qquad (16a)$$

write, using (15a),

$$(k^2-2p_1\cdot k-\Delta_1)^{-2}(k^2-2p_2\cdot k-\Delta_2)^{-1}=\int_0^1 2x\,dx(k^2-2p_x\cdot k-\Delta_x)^{-3},$$

where

$$p_x=xp_1+(1-x)p_2 \quad \text{and} \quad \Delta_x=x\Delta_1+(1-x)\Delta_2, \tag{17a}$$

(note that Δ_x is *not* equal to $m^2-p_x^2$) so that the expression (16a) is $(8i)\int_0^1 2x\,dx\int(1;k_\sigma)d^4k(k^2-2p_x\cdot k-\Delta_x)^{-3}$ which may now be evaluated by (12a) and is

$$(16a)=\int_0^1(1;p_{x\sigma})2x\,dx(p_x^2+\Delta_x)^{-1}, \tag{18a}$$

where p_x, Δ_x are given in (17a). The integral in (18a) is elementary, being the integral of ratio of polynomials, the denominator of second degree in x. The general expression although readily obtained is a rather complicated combination of roots and logarithms.

Other integrals can be obtained again by parametric differentiation. For example differentiation of (16a), (18a) with respect to Δ_2 or $p_{2\tau}$ gives

$$(8i)\int(1;k_\sigma;k_\sigma k_\tau)d^4k(k^2-2p_1\cdot k-\Delta_1)^{-2}(k^2-2p_2\cdot k-\Delta_2)^{-2}$$
$$=-\int_0^1(1;p_{x\sigma};p_{x\sigma}p_{x\tau}-\tfrac{1}{2}\delta_{\sigma\tau}(p_x^2+\Delta_x))\times 2x(1-x)dx(p_x^2+\Delta_x)^{-2}, \tag{19a}$$

again leading to elementary integrals.

As an example, consider the case that the second factor is just $(k^2-L)^{-2}$ and in the first put $p_1=p$, $\Delta_1=\Delta$. Then $p_x=xp$, $\Delta_x=x\Delta+(1-x)L$. There results

$$(8i)\int(1;k_\sigma;k_\sigma k_\tau)d^4k(k^2-L)^{-2}(k^2-2p\cdot k-\Delta)^{-2}$$
$$=-\int_0^1(1;xp_\sigma;x^2p_\sigma p_\tau-\tfrac{1}{2}\delta_{\sigma\tau}(x^2p^2+\Delta_x))\times 2x(1-x)dx(x^2p^2+\Delta_x)^{-2}. \tag{20a}$$

Integrals with three factors can be reduced to those involving two by using (14a) again. They, therefore, lead to integrals with two parameters (e.g., see application to radiative correction to scattering below).

The methods of calculation given in this paper are deceptively simple when applied to the lower order processes. For processes of increasingly higher orders the complexity and difficulty increases rapidly, and these methods soon become impractical in their present form.

A. Self-Energy

The self-energy integral (19) is

$$(e^2/\pi i)\int\gamma_\mu(\not p-\not k-m)^{-1}\gamma_\mu k^{-2}d^4kC(k^2), \tag{19}$$

so that it requires that we find (using the principle of (8a)) the integral on L from 0 to λ^2 of

$$\int\gamma_\mu(\not p-\not k+m)\gamma_\mu d^4k(k^2-L)^{-2}(k^2-2p\cdot k)^{-1},$$

since $(p-k)^2-m^2=k^2-2p\cdot k$, as $p^2=m^2$. This is of the form (16a) with $\Delta_1=L$, $p_1=0$, $\Delta_2=0$, $p_2=p$ so that (18a) gives, since $p_x=(1-x)p$, $\Delta_x=xL$,

$$(8i)\int(1;k_\sigma)d^4k(k^2-L)^{-2}(k^2-2p\cdot k)^{-1}=\int_0^1(1;(1-x)p_\sigma)2x\,dx((1-x)^2m^2+xL)^{-1},$$

or performing the integral on L, as in (8),

$$(8i)\int(1;k_\sigma)d^4k\,k^{-2}C(k^2)(k^2-2p\cdot k)^{-1}=\int_0^1(1;(1-x)p_\sigma)2dx\,\ln\frac{x\lambda^2+(1-x)^2m^2}{(1-x)^2m^2}.$$

Assuming now that $\lambda^2\gg m^2$ we neglect $(1-x)^2m^2$ relative to $x\lambda^2$ in the argument of the logarithm, which then becomes $(\lambda^2/m^2)(x/(1-x)^2)$. Then since $\int_0^1 dx\,\ln(x(1-x)^{-2})=1$ and $\int_0^1(1-x)dx\,\ln(x(1-x)^{-2})=-(1/4)$ find

$$(8i)\int(1;k_\sigma)k^{-2}C(k^2)d^4k(k^2-2p\cdot k)^{-1}=\left(2\ln\frac{\lambda^2}{m^2}+2;\ p_\sigma\left(\ln\frac{\lambda^2}{m^2}-\frac{1}{2}\right)\right),$$

so that substitution into (19) (after the $(p-k-m)^{-1}$ in (19) is replaced by $(p-k+m)(k^2-2p\cdot k)^{-1}$) gives

$$(19)=(e^2/8\pi)\gamma_\mu[(\not p+m)(2\ln(\lambda^2/m^2)+2)-\not p(\ln(\lambda^2/m^2)-\tfrac{1}{2})]\gamma_\mu$$
$$=(e^2/8\pi)[8m(\ln(\lambda^2/m^2)+1)-\not p(2\ln(\lambda^2/m^2)+5)], \tag{20}$$

using (4a) to remove the γ_μ's. This agrees with Eq. (20) of the text, and gives the self-energy (21) when $\not p$ is replaced by m.

B. Corrections to Scattering

The term (12) in the radiationless scattering, after rationalizing the matrix denominators and using $p_1^2=p_2^2=m^2$ requires the integrals (9a), as we have discussed. This is an integral with three denominators which we do in two stages. First the factors $(k^2-2p_1\cdot k)$ and $(k^2-2p_2\cdot k)$ are combined by a parameter y;

$$(k^2-2p_1\cdot k)^{-1}(k^2-2p_2\cdot k)^{-1}=\int_0^1 dy(k^2-2p_y\cdot k)^{-2},$$

from (14a) where

$$p_y=yp_1+(1-y)p_2. \tag{21a}$$

We therefore need the integrals

$$(8i)\int(1;k_\sigma;k_\sigma k_\tau)d^4k(k^2-L)^{-2}(k^2-2p_y\cdot k)^{-2}, \tag{22a}$$

which we will then integrate with respect to y from 0 to 1. Next we do the integrals (22a) immediately from (20a) with $p=p_y$, $\Delta=0$:

$$(22a)=-\int_0^1\int_0^1(1;xp_{y\sigma};x^2p_{y\sigma}p_{y\tau}-\tfrac{1}{2}\delta_{\sigma\tau}(x^2p_y^2+(1-x)L))2x(1-x)dx(x^2p_y^2+L(1-x))^{-2}dy.$$

We now turn to the integrals on L as required in (8a). The first term, (1), in $(1;k_\sigma;k_\sigma k_\tau)$ gives no trouble for large L, but if L is put equal to zero there results $x^{-2}p_y^{-2}$ which leads to a diverging integral on x as $x\to 0$. This infra-red catastrophe is analyzed by using $\lambda_{\min}^2$ for the lower limit of the L integral. For the last term the upper limit of L must be kept as λ^2. Assuming $\lambda_{\min}^2\ll p_y^2\ll\lambda^2$ the x integrals which remain are trivial, as in the self-energy case. One finds

$$-(8i)\int(k^2-\lambda_{\min}^2)^{-1}d^4kC(k^2-\lambda_{\min}^2)(k^2-2p_1\cdot k)^{-1}(k^2-2p_2\cdot k)^{-1}=\int_0^1 p_y^{-2}dy\,\ln(p_y^2/\lambda_{\min}^2) \tag{23a}$$

$$-(8i)\int k_\sigma k^{-2}d^4kC(k^2)(k^2-2p_1\cdot k)^{-1}(k^2-2p_2\cdot k)^{-1}=2\int_0^1 p_{y\sigma}p_y^{-2}dy, \tag{24a}$$

$$-(8i)\int k_\sigma k_\tau k^{-2}d^4kC(k^2)(k^2-2p_1\cdot k)^{-1}(k^2-2p_2\cdot k)^{-1}=\int_0^1 p_{y\sigma}p_{y\tau}p_y^{-2}dy-\tfrac{1}{2}\delta_{\sigma\tau}\int_0^1 dy\,\ln(\lambda^2p_y^{-2})+\tfrac{1}{2}\delta_{\sigma\tau}. \tag{25a}$$

The integrals on y give,

$$\int_0^1 p_y^{-2}dy\,\ln(p_y^2\lambda_{\min}^{-2})=4(m^2\sin2\theta)^{-1}\left[\theta\ln(m\lambda_{\min}^{-1})-\int_0^\theta\alpha\tan\alpha\,d\alpha\right], \tag{26a}$$

$$\int_0^1 p_{y\sigma}p_y^{-2}dy=\theta(m^2\sin2\theta)^{-1}(p_{1\sigma}+p_{2\sigma}), \tag{27a}$$

$$\int_0^1 p_{y\sigma}p_{y\tau}p_y^{-2}dy=\theta(2m^2\sin2\theta)^{-1}(p_{1\sigma}+p_{1\tau})(p_{2\sigma}+p_{2\tau})+q^{-2}q_\sigma q_\tau(1-\theta\,\mathrm{ctn}\theta), \tag{28a}$$

$$\int_0^1 dy\,\ln(\lambda^2p_y^{-2})=\ln(\lambda^2/m^2)+2(1-\theta\,\mathrm{ctn}\theta). \tag{29a}$$

These integrals on y were performed as follows. Since $p_2=p_1+q$ where q is the momentum carried by the potential, it follows from $p_2^2=p_1^2=m^2$ that $2p_1\cdot q=-q^2$ so that since $p_u=p_1+q(1-y)$, $p_u^2=m^2-q^2y(1-y)$. The substitution $2y-1=\tan\alpha/\tan\theta$ where θ is defined by $4m^2\sin^2\theta\approx q^2$ is useful for it means $p_u^2=m^2\sec^2\alpha/\sec^2\theta$ and $p_u^{-2}dy=(m^2\sin 2\theta)^{-1}d\alpha$ where α goes from $-\theta$ to $+\theta$.

These results are substituted into the original scattering formula (2a), giving (22). It has been simplified by frequent use of the fact that p_1 operating on the initial state is m, and likewise p_2 when it appears at the left is replacable by m. (Thus, to simplify:

$$\gamma_\mu p_2 a p_1\gamma_\mu=-2p_1ap_2 \text{ by (4a)},$$
$$=-2(p_2-q)a(p_1+q)=-2(m-q)a(m+q).$$

A term like $qaq=-q^2a+2(a\cdot q)q$ is equivalent to just $-q^2a$ since $q=p_2-p_1=m-m$ has zero matrix element.) The renormalization term requires the corresponding integrals for the special case $q=0$.

C. Vacuum Polarization

The expressions (32) and (32′) for $J_{\mu\nu}$ in the vacuum polarization problem require the calculation of the integral

$$J_{\mu\nu}(m^2)=-\frac{e^2}{\pi i}\int Sp[\gamma_\mu(p-\tfrac{1}{2}q+m)\gamma_\nu(p+\tfrac{1}{2}q+m)]d^4p$$
$$\times((p-\tfrac{1}{2}q)^2-m^2)^{-1}((p+\tfrac{1}{2}q)^2-m^2)^{-1}, \quad (32)$$

where we have replaced p by $p-\frac{1}{2}q$ to simplify the calculation somewhat. We shall indicate the method of calculation by studying the integral,

$$I(m^2)=\int p_\sigma p_\tau d^4p((p-\tfrac{1}{2}q)^2-m^2)^{-1}((p+\tfrac{1}{2}q)^2-m^2)^{-1}.$$

The factors in the denominator, $p^2-p\cdot q-m^2+\frac{1}{4}q^2$ and $p^2+p\cdot q-m^2+\frac{1}{4}q^2$ are combined as usual by (8a) but for symmetry we substitute $x=\frac{1}{2}(1+\eta)$, $(1-x)=\frac{1}{2}(1-\eta)$ and integrate η from -1 to $+1$:

$$I(m^2)=\int_{-1}^{+1}p_\sigma p_\tau d^4p(p^2-\eta p\cdot q-m^2+\tfrac{1}{4}q^2)^{-2}d\eta/2. \quad (30a)$$

But the integral on p will not be found in our list for it is badly divergent. However, as discussed in Section 7, Eq. (32′) we do not wish $I(m^2)$ but rather $\int_0^\infty[I(m^2)-I(m^2+\lambda^2)]G(\lambda)d\lambda$. We can calculate the difference $I(m^2)-I(m^2+\lambda^2)$ by first calculating the derivative $I'(m^2+L)$ of I with respect to m^2 at m^2+L and later integrating L from zero to λ^2. By differentiating (30a), with respect to m^2 find,

$$I'(m^2+L)=\int_{-1}^{+1}p_\sigma p_\tau d^4p(p^2-\eta p\cdot q-m^2-L+\tfrac{1}{4}q^2)^{-3}d\eta.$$

This still diverges, but we can differentiate again to get

$$I''(m^2+L)=3\int_{-1}^{+1}p_\sigma p_\tau d^4p(p^2-\eta p\cdot q-m^2-L+\tfrac{1}{4}q^2)^{-4}d\eta$$
$$=-(8i)^{-1}\int_{-1}^{+1}(\tfrac{1}{4}\eta^2q_\sigma q_\tau D^{-2}-\tfrac{1}{2}\delta_{\sigma\tau}D^{-1})d\eta \quad (31a)$$

(where $D=\frac{1}{4}(\eta^2-1)q^2+m^2+L$), which now converges and has been evaluated by (13a) with $p=\frac{1}{2}\eta q$ and $\Delta=m^2+L-\frac{1}{4}q^2$. Now to get I' we may integrate I'' with respect to L as an indefinite integral and *we may choose any convenient arbitrary constant*. This is because a constant C in I' will mean a term $-C\lambda^2$ in $I(m^2)-I(m^2+\lambda^2)$ which vanishes since we will integrate the results times $G(\lambda)d\lambda$ and $\int_0^\infty\lambda^2G(\lambda)d\lambda=0$. This means that the logarithm appearing on integrating L in (31a) presents no problem. We may take

$$I'(m^2+L)=(8i)^{-1}\int_{-1}^{+1}[\tfrac{1}{4}\eta^2q_\sigma q_\tau D^{-1}+\tfrac{1}{2}\delta_{\sigma\tau}\ln D]d\eta+C\delta_{\sigma\tau},$$

a subsequent integral on L and finally on η presents no new problems. There results

$$-(8i)\int p_\sigma p_\tau d^4p((p-\tfrac{1}{2}q)^2-m^2)^{-1}((p+\tfrac{1}{2}q)^2-m^2)^{-1}$$
$$=(q_\sigma q_\tau-\delta_{\sigma\tau}q^2)\left[\frac{1}{0}-\frac{4m^2-q^2}{3q^2}\left(1-\frac{\theta}{\tan\theta}\right)+\frac{1}{6}\ln\frac{\lambda^2}{m^2}\right]$$
$$+\delta_{\sigma\tau}[(\lambda^2+m^2)\ln(\lambda^2m^{-2}+1)-C'\lambda^2], \quad (32a)$$

where we assume $\lambda^2\gg m^2$ and have put some terms into the arbitrary constant C' which is independent of λ^2 (but in principle could *depend on* q^2) and which drops out in the integral on $G(\lambda)d\lambda$. We have set $q^2=4m^2\sin^2\theta$.

In a very similar way the integral with m^2 in the numerator can be worked out. It is, of course, necessary to differentiate this m^2 also when calculating I' and I''. There results

$$-(8i)\int m^2d^4p((p-\tfrac{1}{2}q)^2-m^2)^{-1}((p+\tfrac{1}{2}q)^2-m^2)^{-1}$$
$$=4m^2(1-\theta\operatorname{ctn}\theta)-q^2/3+2(\lambda^2+m^2)\ln(\lambda^2m^{-2}+1)-C''\lambda^2), \quad (33a)$$

with another unimportant constant C''. The complete problem requires the further integral,

$$-(8i)\int(1;p_\sigma)d^4p((p-\tfrac{1}{2}q)^2-m^2)^{-1}((p+\tfrac{1}{2}q)^2-m^2)^{-1}$$
$$=(1,0)(4(1-\theta\operatorname{ctn}\theta)+2\ln(\lambda^2m^{-2})). \quad (34a)$$

The value of the integral (34a) times m^2 differs from (33a), of course, because the results on the right are not actually the integrals on the left, but rather equal their actual value minus their value for $m^2=m^2+\lambda^2$.

Combining these quantities, as required by (32), dropping the constants C', C'' and evaluating the spur gives (33). The spurs are evaluated in the usual way, noting that the spur of any odd number of γ matrices vanishes and $Sp(AB)=Sp(BA)$ for arbitrary A, B. The $Sp(1)=4$ and we also have

$$\tfrac{1}{4}Sp[(p_1+m_1)(p_2-m_2)]=p_1\cdot p_2-m_1m_2, \quad (35a)$$

$$\tfrac{1}{4}Sp[(p_1+m_1)(p_2-m_2)(p_3+m_3)(p_4-m_4)]$$
$$=(p_1\cdot p_2-m_1m_2)(p_3\cdot p_4-m_3m_4)$$
$$-(p_1\cdot p_3-m_1m_3)(p_2\cdot p_4-m_2m_4)$$
$$+(p_1\cdot p_4-m_1m_4)(p_2\cdot p_3-m_2m_3), \quad (36a)$$

where p_i, m_i are arbitrary four-vectors and constants.

It is interesting that the terms of order $\lambda^2\ln\lambda^2$ go out, so that the charge renormalization depends only logarithmically on λ^2. This is not true for some of the meson theories. Electrodynamics is suspiciously unique in the mildness of its divergence.

D. More Complex Problems

Matrix elements for complex problems can be set up in a manner analogous to that used for the simpler cases. We give three illustrations; higher order corrections to the Møller scatter-

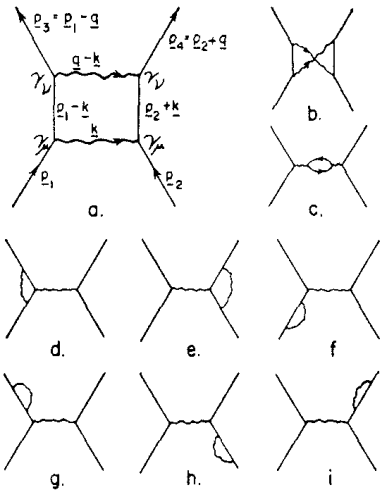


FIG. 8. The interaction between two electrons to order $(e^2/\hbar c)^2$. One adds the contribution of every figure involving two virtual quanta, Appendix D.

ing, to the Compton scattering, and the interaction of a neutron with an electromagnetic field.

For the Møller scattering, consider two electrons, one in state u_1 of momentum p_1 and the other in state u_2 of momentum p_2. Later they are found in states u_3, p_3 and u_4, p_4. This may happen (first order in $e^2/\hbar c$) because they exchange a quantum of momentum $q=p_1-p_3=p_4-p_2$ in the manner of Eq. (4) and Fig. 1. The matrix element for this process is proportional to (translating (4) to momentum space)

$$(\bar{u}_4\gamma_\mu u_2)(\bar{u}_3\gamma_\mu u_1)q^{-2}. \tag{37a}$$

We shall discuss corrections to (37a) to the next order in $e^2/\hbar c$. (There is also the possibility that it is the electron at 2 which finally arrives at 3, the electron at 1 going to 4 through the exchange of quantum of momentum p_3-p_2. The amplitude for this process, $(\bar{u}_4\gamma_\mu u_1)(\bar{u}_3\gamma_\mu u_2)(p_3-p_2)^{-2}$, must be subtracted from (37a) in accordance with the exclusion principle. A similar situation exists to each order so that we need consider in detail only the corrections to (37a), reserving to the last the subtraction of the same terms with 3, 4 exchanged.)

One reason that (37a) is modified is that two quanta may be exchanged, in the manner of Fig. 8a. The total matrix element for all exchanges of this type is

$$(e^2/\pi i)\int(\bar{u}_3\gamma_\nu(p_1-k-m)^{-1}\gamma_\mu u_1)(\bar{u}_4\gamma_\nu(p_2+k-m)^{-1}\gamma_\mu u_2)\,k^{-2}(q-k)^{-2}d^4k, \tag{38a}$$

as is clear from the figure and the general rule that electrons of momentum p contribute in amplitude $(p-m)^{-1}$ between interactions γ_μ, and that quanta of momentum k contribute k^{-2}. In integrating on d^4k and summing over μ and ν, we add all alternatives of the type of Fig. 8a. If the time of absorption, γ_μ, of the quantum k by electron 2 is later than the absorption, γ_ν, of $q-k$, this corresponds to the virtual state p_2+k being a positron (so that (38a) contains over thirty terms of the conventional method of analysis).

In integrating over all these alternatives we have considered all possible distortions of Fig. 8a which preserve the order of events along the trajectories. We have not included the possibilities corresponding to Fig. 8b, however. Their contribution is

$$(e^2/\pi i)\int(\bar{u}_3\gamma_\nu(p_1-k-m)^{-1}\gamma_\mu u_1)\times(\bar{u}_4\gamma_\mu(p_2+q-k-m)^{-1}\gamma_\nu u_2)k^{-2}(q-k)^{-2}d^4k, \tag{39a}$$

as is readily verified by labeling the diagram. The contributions of all possible ways that an event can occur are to be added. This means that one adds with equal weight the integrals corresponding to each topologically distinct figure.

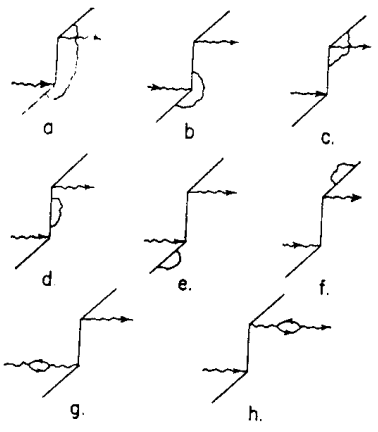


FIG. 9. Radiative correction to the Compton scattering term (a) of Fig. 5. Appendix D.

To this same order there are also the possibilities of Fig. 8d which give

$$(e^2/\pi i)\int(\bar{u}_3\gamma_\nu(p_3-k-m)^{-1}\gamma_\mu(p_1-k-m)^{-1}\gamma_\nu u_1)\times(\bar{u}_4\gamma_\mu u_2)k^{-2}q^{-2}d^4k.$$

This integral on k will be seen to be precisely the integral (12) for the radiative corrections to scattering, which we have worked out. The term may be combined with the renormalization terms resulting from the difference of the effects of mass change and the terms, Figs. 8f and 8g. Figures 8e, 8h, and 8i are similarly analyzed.

Finally the term Fig. 8c is clearly related to our vacuum polarization problem, and when integrated gives a term proportional to $(\bar{u}_4\gamma_\mu u_2)(\bar{u}_3\gamma_\nu u_1)J_{\mu\nu}q^{-4}$. If the charge is renormalized the term $\ln(\lambda/m)$ in $J_{\mu\nu}$ in (33) is omitted so there is no remaining dependence on the cut-off.

The only new integrals we require are the convergent integrals (38a) and (39a). They can be simplified by rationalizing the denominators and combining them by (14a). For example (38a) involves the factors $(k^2-2p_1\cdot k)^{-1}(k^2+2p_2\cdot k)^{-1}k^{-2}(q^2+k^2-2q\cdot k)^{-2}$. The first two may be combined by (14a) with a parameter x, and the second pair by an expression obtained by differentiation (15a) with respect to b and calling the parameter y. There results a factor $(k^2-2p_x\cdot k)^{-2}(k^2+yq^2-2yq\cdot k)^{-1}$ so that the integrals on d^4k now involve two factors and can be performed by the methods given earlier in the appendix. The subsequent integrals on the parameters x and y are complicated and have not been worked out in detail.

Working with charged mesons there is often a considerable reduction of the number of terms. For example, for the interaction between protons resulting from the exchange of two mesons only the term corresponding to Fig. 8b remains. Term 8a, for example, is impossible, for if the first proton emits a positive meson the second cannot absorb it directly for only neutrons can absorb positive mesons.

As a second example, consider the radiative correction to the Compton scattering. As seen from Eq. (15) and Fig. 5 this scattering is represented by two terms, so that we can consider the corrections to each one separately. Figure 9 shows the types of terms arising from corrections to the term of Fig. 5a. Calling k the momentum of the virtual quantum, Fig. 9a gives an integral

$$\int\gamma_\mu(p_2-k-m)^{-1}e_2(p_1+q_1-k-m)^{-1}e_1(p_1-k-m)^{-1}\gamma_\mu k^{-2}d^4k,$$

convergent without cut-off and reducible by the methods outlined in this appendix.

The other terms are relatively easy to evaluate. Terms b and c of Fig. 9 are closely related to radiative corrections (although somewhat more difficult to evaluate, for one of the states is not that of a free electron, $(p_1+q)^2\neq m^2$). Terms e, f are renormalization terms. From term d must be subtracted explicitly the effect of mass Δm, as analyzed in Eqs. (26) and (27) leading to (28) with $p'=p_1+q$, $a=e_2$, $b=e_1$. Terms g, h give zero since the vacuum polarization has zero effect on free light quanta, $q_1^2=0$, $q_2^2=0$. The total is insensitive to the cut-off λ.

The result shows an infra-red catastrophe, the largest part of the effect. When cut-off at $\lambda_{\min}$, the effect proportional to $\ln(m/\lambda_{\min})$ goes as

$$(e^2/\pi)\ln(m/\lambda_{\min})(1-2\theta\operatorname{ctn}2\theta), \tag{40a}$$

times the uncorrected amplitude, where $(p_2-p_1)^2=4m^2\sin^2\theta$. This is the same as for the radiative correction to scattering for a deflection p_2-p_1. This is physically clear since the long wave quanta are not effected by short-lived intermediate states. The infra-red effects arise[28] from a final adjustment of the field from the asymptotic coulomb field characteristic of the electron of

[28] F. Bloch and A. Nordsieck, Phys. Rev. **52**, 54 (1937).

momentum p_1 before the collision to that characteristic of an electron moving in a new direction p_2 after the collision.

The complete expression for the correction is a very complicated expression involving transcendental integrals.

As a final example we consider the interaction of a neutron with an electromagnetic field in virtue of the fact that the neutron may emit a virtual negative meson. We choose the example of pseudoscalar mesons with pseudovector coupling. The change in amplitude due to an electromagnetic field $A=a\exp(-iq\cdot x)$ determines the scattering of a neutron by such a field. In the limit of small q it will vary as $qa-aq$ which represents the interaction of a particle possessing a magnetic moment. The first-order interaction between an electron and a neutron is given by the same calculation by considering the exchange of a quantum between the electron and the nucleon. In this case a_μ is q^{-2} times the matrix element of γ_μ between the initial and final states of the electron, the states differing in momentum by q.

The interaction may occur because the neutron of momentum p_1 emits a negative meson becoming a proton which proton interacts with the field and then reabsorbs the meson (Fig. 10a). The matrix for this process is ($p_2=p_1+q$),

$$\int(\gamma_5 k)(p_2-k-M)^{-1}a(p_1-k-M)^{-1}(\gamma_5 k)(k^2-\mu^2)^{-1}d^4k. \quad (41a)$$

Alternatively it may be the meson which interacts with the field. We assume that it does this in the manner of a scalar potential satisfying the Klein Gordon Eq. (35), (Fig. 10b)

$$-\int(\gamma_5 k_2)(p_1-k_1-M)^{-1}(\gamma_5 k_1)(k_2^2-\mu^2)^{-1}\times(k_2\cdot a+k_1\cdot a)(k_1^2-\mu^2)^{-1}d^4k_1, \quad (42a)$$

where we have put $k_2=k_1+q$. The change in sign arises because the virtual meson is negative. Finally there are two terms arising from the $\gamma_5 a$ part of the pseudovector coupling (Figs. 10c, 10d)

$$\int(\gamma_5 k)(p_2-k-M)^{-1}(\gamma_5 a)(k^2-\mu^2)^{-1}d^4k, \quad (43a)$$

and

$$\int(\gamma_5 a)(p_1-k-M)^{-1}(\gamma_5 k)(k^2-\mu^2)^{-1}d^4k. \quad (44a)$$

Using convergence factors in the manner discussed in the section on meson theories each integral can be evaluated and the results combined. Expanded in powers of q the first term gives the magnetic moment of the neutron and is insensitive to the cut-off, the next gives the scattering amplitude of slow electrons on neutrons, and depends logarithmically on the cut-off.

The expressions may be simplified and combined somewhat before integration. This makes the integrals a little easier and also shows the relation to the case of pseudoscalar coupling. For example in (41a) the final $\gamma_5 k$ can be written as $\gamma_5(k-p_1+M)$ since $p_1=M$ when operating on the initial neutron state. This is

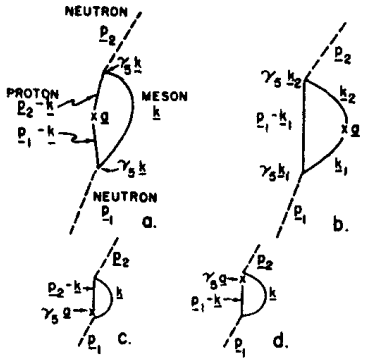


FIG. 10. According to the meson theory a neutron interacts with an electromagnetic potential a by first emitting a virtual charged meson. The figure illustrates the case for a pseudoscalar meson with pseudovector coupling. Appendix D.

$(p_1-k-M)\gamma_5+2M\gamma_5$ since γ_5 anticommutes with p_1 and k. The first term cancels the $(p_1-k-M)^{-1}$ and gives a term which just cancels (43a). In a like manner the leading factor $\gamma_5 k$ in (41a) is written as $-2M\gamma_5-\gamma_5(p_2-k-M)$, the second term leading to a simpler term containing no $(p_2-k-M)^{-1}$ factor and combining with a similar one from (44a). One simplifies the $\gamma_5 k_1$ and $\gamma_5 k_2$ in (42a) in an analogous way. There finally results terms like (41a), (42a) but with pseudoscalar coupling $2M\gamma_5$ instead of $\gamma_5 k$, no terms like (43a) or (44a) and a remainder, representing the difference in effects of pseudovector and pseudoscalar coupling. The pseudoscalar terms do not depend sensitively on the cut-off, but the difference term depends on it logarithmically. The difference term affects the electron-neutron interaction but not the magnetic moment of the neutron.

Interaction of a proton with an electromagnetic potential can be similarly analyzed. There is an effect of virtual mesons on the electromagnetic properties of the proton even in the case that the mesons are neutral. It is analogous to the radiative corrections to the scattering of electrons due to virtual photons. The sum of the magnetic moments of neutron and proton for charged mesons is the same as the proton moment calculated for the corresponding neutral mesons. In fact it is readily seen by comparing diagrams, that for arbitrary q, the scattering matrix to *first order in the electromagnetic potential* for a proton according to neutral meson theory is equal, if the mesons were charged, to the sum of the matrix for a neutron and the matrix for a proton. This is true, for any type or mixtures of meson coupling, to all orders in the coupling (neglecting the mass difference of neutron and proton).

PHYSICAL REVIEW VOLUME 85, NUMBER 2 JANUARY 15, 1952

Radiative Corrections to Compton Scattering

L. M. BROWN*
Cornell University, Ithaca, New York
AND
R. P. FEYNMAN†
California Institute of Technology, Pasadena, California
(Received August 6, 1951)

Corrections of order e^6 to the differential cross section for Compton scattering of unpolarized radiation by electrons are computed. The results for corrections ascribable to virtual photons are finite, relativistically invariant, and valid at all energies, but contain a term which depends logarithmically on an assumed small photon mass λ. A cross section of the same order has also been obtained for double Compton scattering in which one of the emitted photons has an energy small compared to the rest mass of the electron (with the electron initially at rest). This contains a term depending on $\ln\lambda$ which exactly compensates the similar term arising from virtual quanta in all observable cases. Approximations for low and high energies, as well as numerical results, are given. These disagree with results obtained previously by Schafroth.

THE object of this paper is to obtain the correction to the differential cross section for Compton scattering (Klein-Nishina formula) arising from the possibility that the electron may emit and reabsorb a virtual photon in connection with the scattering process. We shall apply the methods developed by one of us[1] to obtain an explicit cross section to order e^6 for unpolarized radiation, valid (in so far as the theory is valid) at all energies.

Previous workers have shown that the high frequency divergences which enter in the straightforward application of perturbation theory to this problem can be removed by charge and mass renormalization. Schafroth[2,3] has obtained a finite e^4-order matrix element in relativistic and gauge invariant form. He also showed, following the treatment of the analogous problem for scalar particles by Corinaldesi and Jost,[4] that the infrared divergence which occurs can be removed by addition of the double Compton cross section in which the incoming photon produces two photons on interacting with the electron, and he made explicit evaluation of the cross section (but not of the double scattering) in the nonrelativistic and extreme relativistic approximations. His results, however, disagree with ours in both limits.

Since the interpretation of any experiment to measure the radiative corrections requires a knowledge of the double Compton cross section, we have computed this also, for the case that one of the emitted photons has an energy in the laboratory system which is small compared to the electron rest energy.

After a brief introduction, we shall in Sec. II write down and discuss the matrix element for the corrections. Section III will detail the evaluation of the differential cross section. Section IV will be concerned with the infrared catastrophe and the double Compton effect. Sections V and VI will discuss limiting cases and some numerical results. Mathematical details will be reserved for the appendices.

The method of calculating this effect is given by Feynman,[5] and for brevity we will not repeat the discussion here but will simply carry out the explicit evaluation of the matrix elements involved. Our notation is that of reference 1.

Some improvement has been made in the method of computing matrix elements given in reference 1(b). This is described here in detail in Appendix Y.

I. THE KLEIN-NISHINA FORMULA FOR UNPOLARIZED RADIATION

The direct Compton effect, in which a photon of momentum q_1, polarization e_1, impinges on an electron of initial momentum p_1, to be scattered as a new photon of momentum q_2, polarization e_2, is represented by a matrix element

$$W=R+S \tag{1a}$$

with

$$R=\boldsymbol{e}_2(\boldsymbol{p}_1+\boldsymbol{q}_1-m)^{-1}\boldsymbol{e}_1, \quad S=\boldsymbol{e}_1(\boldsymbol{p}_1-\boldsymbol{q}_2-m)^{-1}\boldsymbol{e}_2. \tag{1b}$$

The final momentum of the electron is, of course, $p_2=p_1+q_1-q_2$. The terms correspond to the diagrams of Fig. 1.

We shall call

$$p_3=p_1+q_1=p_2+q_2, \quad p_4=p_1-q_2=p_2-q_1, \tag{2}$$

and define the important invariants κ, τ by

$$\begin{aligned} m^2\kappa &= m^2-p_3^2=-2p_1\cdot q_1=-2p_2\cdot q_2, \\ m^2\tau &= m^2-p_4^2=2p_1\cdot q_2=2p_2\cdot q_1. \end{aligned} \tag{3}$$

In the laboratory system, with ω_1 and ω_2 the energies of the incoming and outgoing photons, κ is $-2\omega_1/m$ and τ is $2\omega_2/m$. In terms of the quantities defined in

* Now at Northwestern University. Part of this work was done in partial fulfillment of the requirements for the Ph.D. degree at Cornell University.

† Now on leave at Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil.

[1] R. P. Feynman, Phys. Rev. 76, 749 (1949); and Phys. Rev. 76, 769 (1949).

[2] M. R. Schafroth, Helv. Phys. Acta 22, 501 (1949).

[3] M. R. Schafroth, Helv. Phys. Acta 23, 542 (1950).

[4] E. Corinaldesi and R. Jost, Helv. Phys. Acta 21, 183 (1948).

[5] This problem is discussed in reference 1, Appendix D, p. 788.

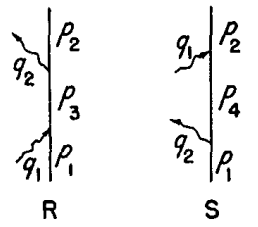


FIG. 1. Momentum diagrams for direct Compton effect.

(2) and (3) we have

$$m^2\kappa R=-e_2(p_3+m)e_1, \quad m^2\tau S=-e_1(p_4+m)e_2. \quad (4)$$

The differential cross section for the final photon to go into solid angle $d\Omega$, if the initial electron is at rest (laboratory system, $p_1=m\gamma_1$) is

$$d\sigma=e^4d\Omega(\omega_2^2/\omega_1^2)F \quad (5)$$

where F is the square of the matrix element of $W(1a)$,

$$F=|{}_2\langle W\rangle_1|^2. \quad (6)$$

If we are uninterested in the spin states of the electron, F may be replaced by $(2m^2)^{-1}U$ where

$$U=\tfrac{1}{4}\,\mathrm{Sp}[(p_2+m)W(p_1+m)\bar{W}]. \quad (7)$$

If, in addition, unpolarized radiation is used and the sum over polarization directions is required, e_1 can be replaced by γ_α and e_2 by γ_β in the spur and half the sum over α, β taken (reference 1(b), Sec. 8). Then the term in (7) which is second order in R is

$$\tfrac{1}{2}(2m^2\kappa)^{-2}\,\mathrm{Sp}[(p_2+m)\gamma_\beta(p_3+m)\gamma_\alpha(p_1+m)\gamma_\alpha \times(p_3+m)\gamma_\beta]=4/\kappa^2-\tau/\kappa-2/\kappa. \quad (8)$$

The reduction can be accomplished by Eqs. (4a) and (36a) of reference 1(b). The term of second order in S is (8) with κ, τ interchanged, since S is obtained from R by replacing p_3 by p_4 after the average is taken on photon polarization. The cross term is

$$(2m^2\kappa)^{-1}(2m^2\tau)^{-1}\,\mathrm{Sp}[(p_2+m)\gamma_\beta(p_3+m)\gamma_\alpha \times(p_1+m)\gamma_\beta(p_4+m)\gamma_\alpha]=8/\kappa\tau-2/\tau-2/\kappa. \quad (9)$$

The sum gives for $U=m^2\sum_{\text{spin}}\sum_{\text{pol}}|{}_2\langle W\rangle_1|^2$:

$$U=4(\kappa^{-1}+\tau^{-1})^2-4(\kappa^{-1}+\tau^{-1})-(\kappa/\tau+\tau/\kappa) \quad (10)$$

and for the Klein-Nishina formula in terms of κ, τ we have:

$$d\sigma=\frac{2\pi e_4}{m^2}\left(\frac{\tau^2}{\kappa^2}\right)\left(\frac{d\tau}{\tau^2}+\frac{d\kappa}{\kappa^2}\right)U. \quad (11)$$

In the laboratory system, in view of $\kappa=-2\omega_1/m$, $\tau=2\omega_2/m$ and the Compton relation

$$\omega_1\omega_2(1-\cos\varphi)=m(\omega_1-\omega_2), \quad (12)$$

(11) can be written in the usual way

$$d\sigma=(e^4/2m^2)d\Omega(\omega_2^2/\omega_1^2)(\omega_1/\omega_2+\omega_2/\omega_1-\sin^2\varphi). \quad (13)$$

II. THE e^4-ORDER MATRIX ELEMENT

The diagrams of the first radiative corrections to term R of the Compton effect are given in Fig. 2. (See reference 1(b), Fig. 9.) The terms containing the analogous modifications of S can be obtained throughout by the interchange of e_1 and e_2, of q_1 and $-q_2$, and of p_3 and p_4. In the final result this means simply an interchange of κ and τ. Hence we need study only R, the S terms being obtained from the R terms immediately.

Terms N' and N'' give zero since there are no vacuum polarization effects for free photons.

Terms M' and M'' together give a factor[6] $r/2i$ times R, where

$$r=\ln(\Lambda/m)+9/4-2\ln(m/\lambda) \quad (14)$$

as shown in reference 1(b), Sec. 6. The quantity Λ is a temporary high frequency cutoff, introduced so that each diagram can be separately evaluated. The final result will become independent of Λ as $\Lambda\to\infty$. The "infrared catastrophe" discussed in Sec. IV is treated, at this point, by assuming the photons to have a small rest mass λ.

The term L is

$$L=\int e_2(p_3-m)^{-1}\gamma_\mu(p_3-k-m)^{-1} \times\gamma_\mu(p_3-m)^{-1}e_1k^{-2}d^4kC(k^2). \quad (15)$$

From this must be subtracted the mass correction for an electron travelling between the absorption and emission of the virtual quantum. Since (to order Δm)

$$(p-m-\Delta m)^{-1}=(p-m)^{-1}+(p-m)^{-1}\Delta m(p-m)^{-1},$$

this gives just the expression for L except that Δm replaces

$$\int\gamma_\mu(p_3-k-m)^{-1}\gamma_\mu k^{-2}d^4kC(k^2),$$

where Δm is the mass correction for cutoff Λ [reference 1(b), Eq. (21)]:

$$\Delta m=im[\tfrac{3}{8}+\tfrac{3}{2}\ln(\Lambda/m)]. \quad (16)$$

Since this diagram occurs for problems other than the one we consider here, we give the result in a general way. Each $(p-m)^{-1}$ propagation factor has, as a consequence of diagrams like L, a correction to the first order in e^2 given by

$$(p-m)^{-1}\int\gamma_\mu(p-k-m)^{-1}\gamma_\mu k^{-2}d^4kC(k^2)(p-m)^{-1} -\Delta m(p-m)^{-2}$$
$$=(4i)^{-1}\left\{(p-m)^{-1}\left[\ln(\Lambda^2/m^2)+\frac{5}{2}-\frac{\eta}{\eta-1}+\frac{\eta(2-\eta)}{(\eta-1)^2}\ln\eta\right]-m(p-m)^{-2}\left[\frac{\eta}{\eta-1}-\frac{\eta(3\eta-2)}{(\eta-1)^2}\ln\eta\right]\right\}, \quad (17)$$

where $m^2\eta=m^2-p^2$.

[6] The factor obtained in reference 1(b) is $-(e^2/2\pi)r$, but we have reserved a factor $e^2/\pi i$ for later inclusion.

Terms K' and K'' again possess a feature common to several problems, and we will therefore first discuss it in a general way. In all problems in which an electron interacts with a potential or a free or virtual photon there will be a piece of the diagram like Fig. 3. That is, there will be a partial factor in one of the matrix elements:

$$T=\int \gamma_\mu(\not p+\not q-\not k-m)^{-1}e(\not p-\not k-m)^{-1}\times\gamma_\mu k^{-2}d^4kC(k^2). \quad (18)$$

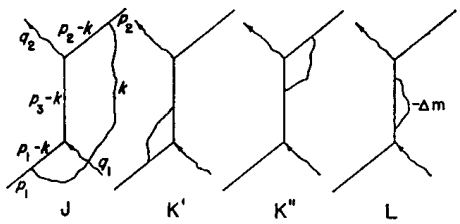
It would be most convenient to have this evaluated in the general case of arbitrary p and q. However, we have evaluated it only in the special case that $q^2=0$, $p^2=m^2$, with the matrix operating on a state u such that $pu=mu$. Calling $m^2\kappa=-2p\cdot q$ it is (Appendix Z):

$$\begin{aligned}8iT=&4\kappa^{-1}[m^2e+2\kappa^{-1}(e\cdot p)q]\int_{1-\kappa}^{1}\ln(1-v)dv/v\\&+2[(2m^2+pq-qp)e\\&+2\kappa^{-1}(e\cdot p)(q+m\kappa)(3\kappa-2)(\kappa-1)^{-1}](\kappa-1)^{-1}\ln\kappa\\&+[2\ln(m^2/\Lambda^2)-1]m^2e\\&-4(e\cdot p)[(q+m)(\kappa-1)^{-1}+q\kappa^{-1}]. \quad (19)\end{aligned}$$

If the final, rather than the initial, state is a free electron, the matrix required is $\bar T$, so the result is obtained directly from (19). For term K', this T for the case $p=p_1$, $q=q_1$, $e=e_1$ is to be multiplied on the left by $e_2(\not p_1+\not q_1-m)^{-1}=e_2(\not p_3-m)^{-1}$. Therefore, K' and the corresponding term K'' together give

$$\begin{aligned}K=K'+K''=(8i)^{-1}[&e_2(\not p_3-m)^{-1}T(p_1,q_1,e_1)\\&+\bar T(p_2,q_2,e_2)(\not p_3-m)^{-1}e_1]. \quad (20)\end{aligned}$$

If we now examine the coefficients of the term $\ln(m^2/\Lambda^2)$ in K, L, and M, that is in (20), (17), and



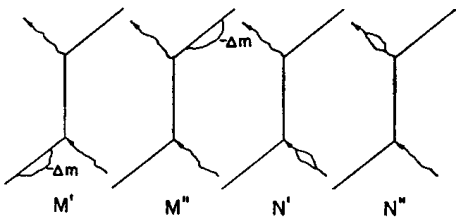


FIG. 2. Corrections to term R of Compton scattering.

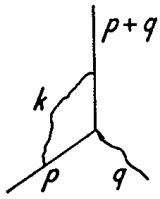


FIG. 3. Diagram for the expression T.

(14), we observe that K gives $(4m^2/8i)R$, L gives $(-2m^2/8i)R$, and M gives $(-2m^2/8i)R$. Therefore the terms dependent on Λ vanish. Since we shall find that the J integral is finite without cutoff, we note that the complete result is insensitive to Λ.

The term J is given by

$$J=\int \gamma_\mu(\not p_2-\not k-m)^{-1}e_2(\not p_3-\not k-m)^{-1}\times e_1(\not p_1-\not k-m)^{-1}\gamma_\mu k^{-2}d_4k. \quad (21)$$

For large k the factors in the integrand vary as k^{-n} with $n\geqq 5$ and the integration over k-space therefore converges. If we had included the convergence factor $C(k^2)$, the result would be independent of Λ as $\Lambda\to\infty$.

When the reciprocals are rationalized (e.g., $(\not p_2-\not k-m)^{-1}=(\not p_2-\not k+m)\cdot[(p_2-k)^2-m^2]^{-1}$), powers of k_μ up to the third appear in the numerator of the integrand. Therefore we shall have to evaluate integrals of the form:

$$\begin{aligned}J_{(0;\sigma;\sigma\tau;\sigma\tau\nu)}&=\int(1;k_\sigma;k_\sigma k_\tau;k_\sigma k_\tau k_\nu)[(p_2-k)^2-m^2]^{-1}\\&\times[(p_3-k)^2-m^2]^{-1}[(p_1-k)^2-m^2]^{-1}k^{-2}d_4k. \quad (22)\end{aligned}$$

That is, for J_0 the factor $(1;k_\sigma;\cdots$ etc.) is replaced by unity, for J_σ by k_σ, for $J_{\sigma\tau}$ by $k_\sigma k_\tau$, and for $J_{\sigma\tau\nu}$ by $k_\sigma k_\tau k_\nu$. The manner in which J can be expressed in terms of these integrals is illustrated, for the case of matrix T in Appendix Z.

The J integrals can be worked out by the parametric methods described in reference 1(b) (Appendix). They involve integrals having four factors in the denominator and will lead, therefore, to integrals over three parameters. (J_0 is integrated in this manner in Appendix Y.) Generally these are very difficult to evaluate, although J_0 is particularly simple. This fact makes it possible to circumvent some of the difficulties of J_σ, $J_{\sigma\tau}$, and $J_{\sigma\tau\nu}$.

It is possible to express these other J integrals as linear combinations of the integral J_0 and of other integrals, all of which involve only three quadratic factors in the denominator. These latter, in parametric form, require only two parameters (and are much more easily evaluated than a direct attack on J_σ, say, would

indicate). This technique is useful in other problems also[7] and is described in detail in Appendix Y.

III. CROSS SECTION FOR UNPOLARIZED LIGHT

If we call the sum $J+K+L+M=R^{(1)}$, then $(e^2/\pi i)R^{(1)}$ will be the correction to the matrix R of the direct effect (1). If the corresponding correction to the term S is called $(e^2/\pi i)S^{(1)}$, the corrected matrix for the Compton effect is

$$W'=W+(e^2/\pi i)W^{(1)}=R+S+(e^2/\pi i)(R^{(1)}+S^{(1)}). \quad (23)$$

The absolute square of the matrix element of W', taken between the initial and final electron states, gives the probability of transition correct to one order in e^2 higher than (6). We shall calculate in this paper only the cross section averaged over spin directions of the electron and polarization directions of the photons.

We need the spur:

$$\tfrac{1}{4}\,\mathrm{Sp}[(\boldsymbol{p}_2+m)W'(\boldsymbol{p}_1+m)\bar{W}'] \quad (24)$$

as in (7). Considering terms up to the first order in e^2 (which are all that are valid), (24) is

$$U-\tfrac{1}{4}\{(e^2/\pi i)\,\mathrm{Sp}[(\boldsymbol{p}_2+m)W(\boldsymbol{p}_1+m)\bar{W}^{(1)}] - (e^2/\pi i)\,\mathrm{Sp}[(\boldsymbol{p}_2+m)W^{(1)}(\boldsymbol{p}_1+m)\bar{W}]\}. \quad (25)$$

In evaluating (25) for unpolarized light we have replaced $\boldsymbol{e}_1$ by γ_α and $\boldsymbol{e}_2$ by γ_β and taken one-half of the resulting sum as discussed in connection with (8). Some algebraic details are discussed in Appendix Z.[8]

The last two spurs in (25) are complex conjugates, so that the correction to U is $-e^2/\pi$ times the real part of

$$U^{(1)}=-(4i)^{-1}\,\mathrm{Sp}[(\boldsymbol{p}_2+m)W^{(1)}(\boldsymbol{p}_1+m)\bar{W}]. \quad (26)$$

That is, U is to be replaced in (11) by

$$U'=U-(e^2/\pi)\mathbf{R.P.}U^{(1)}. \quad (27)$$

If we let

$$P(\kappa,\tau)=-(4i)^{-1}\,\mathrm{Sp}[(\boldsymbol{p}_2+m)R^{(1)}(\boldsymbol{p}_1+m)\bar{W}] \quad (28)$$

then

$$U^{(1)}=P(\kappa,\tau)+P(\tau,\kappa) \quad (29)$$

since the $S^{(1)}$ diagrams are obtained from the $R^{(1)}$ diagrams (for unpolarized light) by the interchange of $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$ and of $\boldsymbol{q}_1$ and $-\boldsymbol{q}_2$; hence the final result, simply by interchange of κ and τ.

The final result obtained in this way is:

$$\begin{aligned}P(\kappa,\tau)=&(1-2y\,\mathrm{ctnh}2y)\,\ln\lambda\cdot U\\ &-2y\,\mathrm{ctnh}2y[2h(y)-h(2y)]U\\ &+[-4y\sinh 2y(\kappa\tau)^{-1}(2-\cosh 2y)\\ &+2y\,\mathrm{ctnh}y]h(y)+\ln\kappa\Big\{4y\,\mathrm{ctnh}2y\Big[\frac{4}{\kappa\tau}\cosh^2 y\\ &+\frac{\kappa-6}{2\tau}\,\mathrm{sech}2y+\frac{4}{\kappa^2}-\frac{1}{\kappa}-\frac{\tau}{2\kappa}-\frac{\kappa}{\tau}-1\Big]\\ &+\frac{3\tau}{2\kappa^2}+\frac{3\tau}{2\kappa}+\frac{3}{\tau}+1-\frac{7}{\kappa\tau}+\frac{8}{\kappa}-\frac{8}{\kappa^2}+\frac{2\kappa-\tau^2-\kappa^2\tau}{2\kappa^2\tau(\kappa-1)}\\ &-\frac{1}{2\tau}\frac{2\kappa^2+\tau}{(\kappa-1)^2}\Big\}+y^2\,\mathrm{csch}^2 y\Big[\frac{2}{\kappa}-\frac{7}{4}-\kappa-\frac{3}{4}\frac{\tau^2}{\kappa}\Big]\\ &-4y\tanh y\Big(\frac{1}{2}-\frac{1}{\kappa}\Big)+4\Big(\frac{1}{\kappa}+\frac{1}{\tau}\Big)^2\\ &-\frac{12}{\kappa}-\frac{3\kappa}{2\tau}-2\frac{\kappa}{\tau^2}+\frac{1}{\kappa-1}\Big(\frac{\kappa}{\tau}+\frac{1}{2}\Big)\\ &+G_0(\kappa)\Big[\frac{\kappa^2}{\tau}+\frac{\tau}{\kappa^2}+\frac{\kappa}{\tau}+\kappa+\frac{1}{2}\tau+\frac{2}{\kappa}+\frac{3}{\tau}-1\Big]\\ &+\text{terms antisymmetric in }\kappa,\tau, \quad (30)\end{aligned}$$

where

$$4\sinh^2 y=-(\kappa+\tau) \quad (30a)$$

$$h(y)=y^{-1}\int_0^y u\,du\,\mathrm{ctnh}u \quad (30b)$$

$$G_0(\kappa)=-2\kappa^{-1}\int_{1-\kappa}^{1}\ln(1-u)\,du/u. \quad (30c)$$

This is to be added to the same expression with κ and τ interchanged (29) and the real part taken to get the correction to the Klein-Nishina formula (11). We discuss this result in the following sections.

We might note here, however, that the real part of $P(\kappa,\tau)$ is obtained by writing $\ln|\kappa|$ for $\ln\kappa$ and by writing for $G_0(\kappa)$ expression (30c) with $\ln(1-u)$ replaced by $\ln(u-1)$. Since τ is always positive, on the other hand, $P(\tau,\kappa)$ is always real. This is discussed further in Appendix W.

The imaginary part of $P(\kappa,\tau)$ is not without interest, as we shall show. This is given by π times the coefficient of $\ln\kappa$ in (30) plus $\pi\ln(1-\kappa)$ times the coefficient of $G_0(\kappa)$.

The loss of total intensity of a beam of photons is of course proportional to the total cross section for a photon to be scattered out of the beam. But this

[7] It has been applied by G. R. Lomanitz to completely evaluate the e^6 corrections to the Möller scattering cross section of electrons in his thesis *Second Order Effects in the Electron-Electron Interaction*, Cornell, 1950. Again, in the problem of scattering of light by light, the integral with unit numerator is easily done, and the other integrals can be reduced to it and simpler integrals algebraically. But here the algebraic complexity makes the problem extremely tedious.

[8] In actual evaluation it was found easier to take the spur first and perform the integrals later. Thus, in place of the expression T (Eq. 18), the expression $\boldsymbol{T}$ (Eq. A41) was substituted and the values of the integrals from Appendix X substituted after taking the spur. This has the advantage that some of the integrals do not appear, or appear only in simpler combinations.

decrease in forward intensity is the result of an interference between the incident photon and a photon scattered exactly in the forward direction. Therefore, as is well known, the imaginary part of the forward scattering amplitude is proportional to the total cross section (formally this is referred to as the unitary property of the S-matrix). We can use this relation to check the imaginary part of $P(\kappa, \tau)$ for the case of zero scattering angle (for which, of course, $p_1=p_2$, $q_1=q_2$, $\kappa=-\tau$).

We write $P(\kappa, \tau)$ again as a sum

$$P(\kappa, \tau)=(4m^2i)\tfrac{1}{4}\sum_{\text{spin}}\sum_{\text{pol}}{}_1\langle R^{(1)}\rangle_2\,{}_2\langle \bar{W}\rangle_1, \quad (31)$$

and can show easily that ${}_2\langle W\rangle_1=i/m$ if there is no spin change and no polarization change, and zero otherwise. This can be seen, aside from the phase factor i, from the fact that for small scattering angles the Klein-Nishina formula (13) is $d\sigma=r_0^2d\Omega$ in the laboratory system. Since ${}_1\langle R^{(1)}\rangle_2$ also vanishes when ${}_2\langle W\rangle_1$ does,

$$P(\kappa, \tau)=m\sum_{\text{spin}}\sum_{\text{pol}}{}_1\langle R^{(1)}\rangle_2. \quad (32)$$

But, including all factors, the complete e^4-order matrix element of $R^{(1)}$ is (according to reference 1(b)):

$$X=\frac{e^2}{\pi i}\cdot\frac{2\pi e^2}{(m\tau/2)}\cdot\frac{1}{4}\sum_{\text{spin}}\sum_{\text{pol}}{}_1\langle R^{(1)}+S^{(1)}\rangle_2 \quad (33)$$

and from the unitary property referred to above, it follows that the *total* cross section for Compton scattering to order e^4 is just twice the real part of X. Therefore,

$$\sigma_{\text{total}}(\text{to order } e^4)=2\mathbf{R.P.}X=(2r_0^2/\tau)\mathbf{I.P.}P(\kappa, \tau) \quad (34)$$

since $S^{(1)}$ has no imaginary part. That (30) satisfies this identity can be readily verified.[9]

IV. THE INFRARED CASTASTROPHE AND THE DOUBLE COMPTON EFFECT

In Sec. II we have derived the differential cross section for Compton scattering for unpolarized light, including radiative corrections, to order e^6. The cross section took the form

$$d\sigma=d\sigma_{\text{K.N.}}[1+(e^2/\pi)\delta] \quad (35)$$

with

$$\delta=-U^{(1)}/U. \quad (35a)$$

There are two reasons why this result cannot be compared directly with experiment. In the first place $U^{(1)}$ depends on the quantity λ to which no experimental significance has been attached. In the second place, it is impossible in principle to design an experiment which will guarantee that one and only one photon is emitted by the electron in the scattering process. The best one can do in an experiment is to require that if a second photon is emitted, its energy is less than some value $k_{\max}$. This can be done, for example, by measuring the energies of the final electron and photon to some specified accuracy, the sum of the errors in the measurement being less than $k_{\max}$. In such an experiment one would be measuring the cross section (35) plus the cross section for the double Compton effect, $d\sigma_D$, integrated over all possible directions of the second quantum and over its energy up to $k_{\max}$.

These two difficulties, both related to quanta of low energy (if $k_{\max}$ is small), in one case virtual, in the other real, are actually related. That this should be so, can be seen physically from the fact that it is difficult to distinguish between virtual and real quanta of extremely low energy since, by the uncertainty principle, a measurement made during a finite time interval will introduce an uncertainty in the energy of the quantum, which may enable a virtual quantum to be detected as a real one. It turns out in fact that $d\sigma_D$, integrated to $k_{\max}$, also contains an infrared divergence which just cancels the similar divergence in the radiative corrections. We are computing, of course, only to order e^6, but the multiple Compton scattering of a given higher order will also cancel all the radiative infrared catastrophes of the same order.

The problem is analogous to the perturbation theory treatment of the scattering of an electron by a potential, which has been considered by many workers, except that in our case the primary process is the Compton scattering considered in Sec. I. The cross section for emission of an additional photon q of energy ω goes for small ω as $(d^3q/\omega)(p_2/p_2\cdot q-p_1/p_1\cdot q)^2$ times the Klein-Nishina formula. Since this diverges as ω approaches zero, the probability of a single Compton process unaccompanied by such emission is zero. What is experimentally measured, however, is the probability that a Compton process occurs and that no other free photon is emitted except for a class of photons inaccessible to the experiment. This is equal, to our order of calculation, to the probability of the single process plus the probability of a double process in which one of the photons emitted is in the inaccessible class. This class is, of course, determined by the design of the experiment and a single calculation cannot suffice for all experiments. However, one feature common to all experiments will be a finite energy resolution, so that a part of the excluded class must consist of photons whose energy is less than some energy $k_{\max}$.

We will first therefore find that part of the differential cross section for double Compton scattering which gives rise to an infrared divergence. This will be integrated over all the directions of one of the photons, and over its energy from zero to a value $k_{\max}$, which we shall assume is small compared to the electron mass, and added to the previously obtained corrected cross section for single Compton scattering. It has already been pointed out that Schafroth[2] has demonstrated that a cancellation of the infrared divergence occurs in order e^6 when the double Compton cross section is added to

[9] W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, London, 1944), p. 157, Eq. (53). Our Eq. (33) agrees with this result with τ replacing Heitler's 2γ.

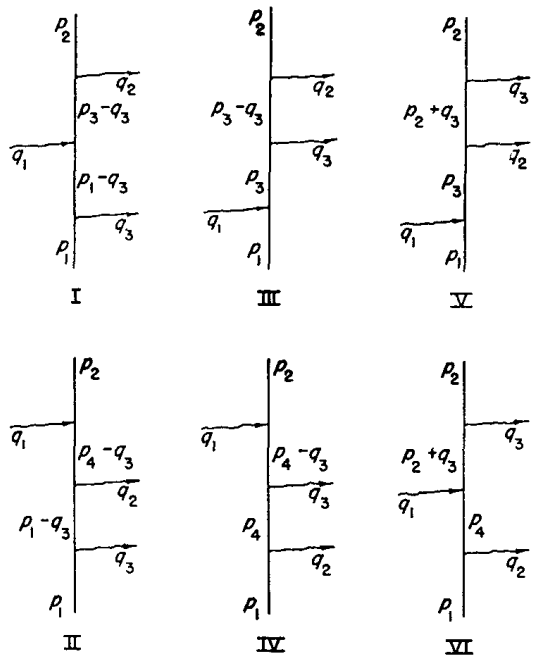


FIG. 4. Diagrams for the double Compton effect. Here $p_3=p_1+q_1$, $p_4=p_1-q_2$; the momentum condition is $p_1+q_1=p_2+q_2+q_3$.

the single scattering cross section, but we must obtain at least the zero order term in $k_{\max}$ (for $k_{\max}\ll m$) to obtain a useful result. The completely differential cross section for the double scattering has been computed by Eliezer,[10] but since we wish to make approximations and carry out an integration, it is simpler for us to obtain the desired cross section from the beginning.

Figure 4 gives the diagrams necessary for computing the cross section for double Compton scattering. The photon momentum $q_3=(\omega_3, \mathbf{q}_3)$ is assumed to be small in the following ($\omega_3\ll m$). This, of course, implies a definite coordinate system. To obtain a finite result we assume the photon has a small rest mass λ, so that $q^2=\lambda^2$. Keeping terms only to order ω_3^{-1}, we neglect q_3 occurring in the numerator of the (rationalized) matrix element terms, and terms of order ω_3 compared to $p_3^2-m^2$ and $p_4^2-m^2$ in the denominators.

We find that terms III and IV are not of the desired order. In view of the fact that we are to make matrix elements between the free electron states u_1, and u_2, a factor $(\not{p}_1+m)\not{e}_3$ (with e_3 the polarization vector of q_3) operating on the left of u_1, is equivalent to $2p_1\cdot e_3$ and a factor $\not{e}_3(\not{p}_2+m)$ operating on the right of u_2 is equivalent to $2p_2\cdot e_3$. Thus, with q_3 small, we get

$$\langle \mathrm{I}\rangle=-Rp_1\cdot e_3/p_1\cdot q_3,\quad \langle \mathrm{II}\rangle=-Sp_1\cdot e_3/p_1\cdot q_3,$$
$$\langle \mathrm{V}\rangle=Rp_2\cdot e_3/p_2\cdot q_3,\quad \langle \mathrm{VI}\rangle=Sp_2\cdot e_3/p_2\cdot q_3. \tag{36}$$

Adding these we find the matrix for the double Compton process:

$$(R+S)\left(\frac{p_2\cdot e_3}{p_2\cdot q_3}-\frac{p_1\cdot e_3}{p_1\cdot q_3}\right),\quad |\mathbf{q}_3|\ll m. \tag{37}$$

Taking the absolute square of (37) and averaging over polarizations and spins in the usual manner, it is clear that we obtain the Klein-Nishina cross section $d\sigma_{\mathrm{K.N.}}$ (Eq. (11)) multiplied by the following factors: (a) $d^3\mathbf{q}_3/(2\pi)^3$, the density of states for $\boldsymbol{q}_3$ (neglecting its effect on the momentum balance, and therefore assuming it is emitted independently of $\boldsymbol{q}_2$), (b) e^2, from the additional interaction vertex, (c) $2\pi/\omega_3$, the normalization factor for the photon $\boldsymbol{q}_3$, (d)

$$\sum_{\mathrm{pol}}\left(\frac{p_2\cdot e_3}{p_2\cdot q_3}-\frac{p_1\cdot e_3}{p_1\cdot q_3}\right)^2=-\left(\frac{p_2}{p_2\cdot q_3}-\frac{p_1}{p_1\cdot q_3}\right)^2.$$

We collect these factors and integrate the photon momentum over all angles and from $\mathbf{q}_3=0$ to the sphere $|\mathbf{q}_3|=k_{\max}$, where $k_{\max}\ll m$. Thus,[11]

$$d\sigma_D=-\frac{e^2}{(2\pi)^2}d\sigma_{\mathrm{K.N.}}\times\int_{|\mathbf{q}_3|=0}^{|\mathbf{q}_3|=k_{\max}}\left(\frac{p_2}{p_2\cdot q_3}-\frac{p_1}{p_1\cdot q_3}\right)^2\frac{d^3\mathbf{q}_3}{(\mathbf{q}_3^2+\lambda^2)^{\frac{1}{2}}}. \tag{38}$$

Observe that if we replace $d\sigma_{\mathrm{K.N.}}$ by the cross section $d\sigma_0$ for an arbitrary process of one electron, the result (37) is valid for that process with one additional photon of small energy in the final state. For if the "small" photon $\boldsymbol{q}_3$ is emitted from an electron line of momentum $\boldsymbol{p}$ where $\boldsymbol{p}^2\neq m^2$, its effect will be negligible. The only diagrams contributing to the process with $\boldsymbol{q}$ emitted will be those of the original process, modified by the emission of $\boldsymbol{q}$ either before or after the original process. Thus the factor of $(R+S)$ in (37) will always be the factor modifying the original matrix element (for small $\boldsymbol{q}$) and the result (38) is the general result for an arbitrary process of one electron, $d\sigma_0$ replacing $d\sigma_{\mathrm{K.N.}}$.

It is most convenient to impose our restriction $k_{\max}\ll m$ in the laboratory system.[12] In this case, the result of integrating (38) (expressed in terms of invariants) is

$$d\sigma_D=-(e^2/\pi)d\sigma_{\mathrm{K.N.}}\{2(1-2y\ \mathrm{ctnh}2y)[\ln(2k_{\max}/\lambda)-\tfrac{1}{2}]+4y\ \mathrm{ctnh}2y[h(2y)-1]\}, \tag{39}$$

with y and $h(y)$ defined in (30).

[10] C. J. Eliezer, Proc. Roy. Soc. (London) **A187**, 210 (1946). When the class of photons inaccessible to the experiment does not consist simply of those below a given very small energy $k_{\max}$ (but consists, for example, of those in a given solid angle, or with a limited momentum component, or having energies too large to permit the approximations we have made) the contribution which these events make to the measured cross section can be obtained from Eliezer's formula. Explicitly, one must add to our result (39) the cross section for the double process given by Eliezer, integrated over all the photons in the class inaccessible to the experiment but which also exceed some arbitrary very small energy $k_{\max}$. The sum, of course, will not depend on $k_{\max}$.

[11] This expression (with $\lambda=0$) has been obtained previously. See, for example, R. Jost, Phys. Rev. **72**, 815 (1947), and F. Bloch and A. Nordsieck, Phys. Rev. **52**, 54 (1937).

[12] J. Schwinger, Phys. Rev. **76**, 790 (1949) has integrated (38) for the case of the scattering of an electron in a potential.

When (39) is added to (35) the effect is to replace the quantity

$$\{2(1-2y\,\mathrm{ctnh}2y)\,\ln\lambda-4y\,\mathrm{ctnh}2y[2h(y)-h(2y)]\}U$$

in $U^{(1)}$ by

$$\{2(1-2y\,\mathrm{ctnh}2y)[\ln(2k_{\max})-\tfrac{1}{2}]+8y\,\mathrm{ctnh}2y[h(2y)-h(y)-\tfrac{1}{2}]\}U. \quad (40)$$

We have now arrived at a physically understandable result as the quantity $k_{\max}$ which replaces λ in $U^{(1)}$ is the sum of the experimental uncertainties in the measurement of the final energies of the Compton scattered photon and electron. Our result can be compared with an experiment providing the energy resolution[10] is known and $k_{\max}$ is sufficiently small.

Under the limits of validity of our formula, the term in (40) containing $k_{\max}$ is positive, and thus makes a negative contribution to the cross section. As the energy resolution of an experiment improves, it is thus found that the measured Compton cross section gets smaller. This is reasonable since we are eliminating from our observations more double Compton events.

The expression (35a) for δ with $U^{(1)}$ given by (29) and U given by (10) is valid also for the correction to the two-quantum pair annihilation and the two-quantum pair production processes, provided that for the former problem we replace κ by $-\kappa$ and for the latter τ by $-\tau$. This occurs because in writing down the matrix element we represent the emission of a photon by $-q$ and its absorption by $+q$, and because a matrix p representing an electron also represents a positron of four-momentum $-p$. However, the infrared divergences in these problems are not compensated by the corresponding three-quantum processes[13] (which are not divergent) but by the effect of Coulomb interaction. This will not be discussed further in this paper.

V. EXTREME RELATIVISTIC LIMIT

The Compton formula (12) can be written in the laboratory system, with scattering angle φ, as

$$(\kappa+\tau)/\tau=\tfrac{1}{2}\kappa(1-\cos\varphi). \quad (41)$$

We assume $|\kappa|\gg1$ and consider the three cases listed in Table I. This table also lists the approximations made in obtaining the formulas for U and $U^{(1)}$ for the three cases and the corresponding conditions on the laboratory and center-of-mass scattering angles. The energies (in units of mc^2) of the incoming and outgoing quanta are represented in the laboratory system by ω_1 and ω_2, respectively, and in the c.m. system by ν. The results are as follows:

Case I.

$$U=2$$

$$U^{(1)}=4(1-2y\,\mathrm{ctnh}2y)\,\ln\lambda-8y\,\mathrm{ctnh}2y[2h(y)-h(2y)]+4yh(y)\,\mathrm{ctnh}y+\ln|\kappa|\,(4y\tanh y-1)-2y^2-4y\tanh y+3-(\ln|\kappa|)^2-\pi^2/6. \quad (42)$$

[13] Note that the two quantum pair processes are symmetric with respect to interchange of p_1 and p_2 so that (37) vanishes.

Case II.

$$U=-[(\kappa/\tau)+(\tau/\kappa)]$$

$$U^{(1)}=U\{(1-2y)(\tfrac{3}{2}+2\ln\lambda)+2y^2-\pi^2/6\}+\left(1+\frac{\tau}{2\kappa}+\frac{\kappa}{\tau}\right)\left\{\left[\ln\left(1+\frac{\tau}{\kappa}\right)\right]^2-\ln\left(1+\frac{\tau}{\kappa}\right)+2\ln\frac{\tau}{|\kappa|}\right\}+\left(1+\frac{\tau}{\kappa}+\frac{\kappa}{2\tau}\right)\left\{\left(\ln\left|1+\frac{\kappa}{\tau}\right|\right)^2-\ln\left(1+\frac{\tau}{\kappa}\right)-\ln\left|\frac{\tau}{\kappa}\right|+\pi^2\right\}. \quad (43)$$

Case III.

$$U=-\kappa/\tau$$

$$U^{(1)}=U\left\{2(1-2y)\,\ln\lambda+\ln\tau\left[2y-\frac{3}{2}\frac{\tau+1}{\tau}+\frac{1}{2\tau(\tau-1)}\right]-\frac{\pi^2}{3}-G_0(\tau)\left(\frac{1}{\tau}+\frac{\tau}{2}\right)+\frac{3}{2}+\frac{2}{\tau}\right\}. \quad (44)$$

The corrected cross sections, in the relativistic limit are given in the laboratory system by

$$d\sigma=(r_0^2/2)d\Omega(\tau^2/\kappa^2)[U-(e^2/\pi)U^{(1)}] \quad (45)$$

and in the c.m. system by

$$d\sigma=(r_0^2/8\nu^2)d\Omega[U-(e^2/\pi)U^{(1)}] \quad (46)$$

with $r_0=e^2/mc^2$.

As we have explained in the previous section, for actual comparison with experiment one must add to (45) the cross section for double Compton scattering (35) which is valid, of course, only in the laboratory system with $k_{\max}\ll m$. If we write (39) as

$$d\sigma_D=(-e^2/\pi)(r_0^2/2)d\Omega(\tau^2/\kappa^2)U_D \quad (47)$$

then we must replace $U^{(1)}$ in (45) and (46) by $U^{(1)}+U_D$. For our three cases, we get:

TABLE I. Approximation made in obtaining the extreme relativistic limit of (30), expressed in invariants, laboratory system quantities, and c.m. system quantities.

Case	Defined by	Leads to	Lab system conditions	c.m. conditions
I	$\lvert(\kappa+\tau)/\tau\rvert\ll1$	$\lvert\kappa\rvert\approx\tau\gg1$	$\omega_2\approx\omega_1$, $1-\cos\varphi\ll1/\omega_1$ (φ near 0)	$\tan^2(\theta/2)\ll1$
II	$\lvert(\kappa+\tau)/\tau\rvert\sim1$	$\tau\gg1$, $\lvert\kappa+\tau\rvert\gg1$	ω_2 near $\omega_1/2$, $1-\cos\varphi\sim1/\omega_1$ (φ near $(2/\omega_1)^{\frac{1}{2}}$)	$\nu^2\cos^2(\theta/2)\gg1$, $\nu^2\sin^2(\theta/2)\gg1$
III	$\lvert(\kappa+\tau)/\tau\rvert\gg1$	$\lvert\kappa+\tau\rvert\gg1$, $\lvert\kappa\rvert\gg\tau$	ω_2 near 1, $1-\cos\varphi\gg1/\omega_1$ ($\varphi\gg(2/\omega_1)^{\frac{1}{2}}$)	$\tan^2(\theta/2)\gg1$

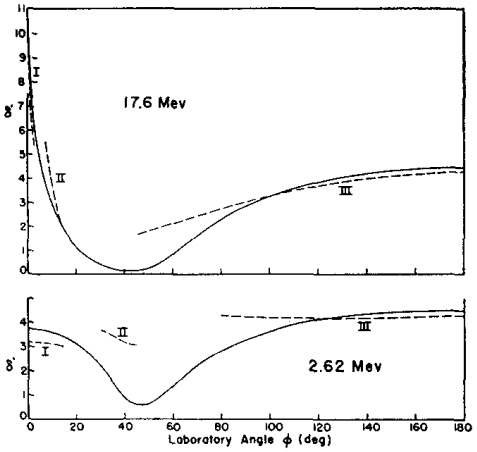


FIG. 5. Plot of $\delta' = -U^{(1)}/U - 2(1-2y\,\mathrm{ctnh}2y)\ln\lambda$ for 2.62 Mev and 17.6 Mev as calculated from the exact expressions (29) and (30) (solid curve). The dotted curves are calculated from the extreme relativistic formulas (42), (43), and (44), and are numbered accordingly.

Case I.

Same as (39).

Cases II and III.

$$U_D = U\{2(1-2y)[\ln(2k_{\max}/\lambda) - \tfrac{1}{2}] + 4y[y + (\pi^2/24y) - 1]\}. \quad (48)$$

In Fig. 5 we have plotted a comparison of the exact expression for $-U^{(1)}/U$ (leaving out the term proportional to $\ln\lambda$) with the limiting cases expressed by (42), (43), and (44). That is, if we write $U^{(1)} = a\ln\lambda + b$, we have plotted $-b/U$. Especially simple formulas result for b in the extreme relativistic limit for the cases $\varphi = 0$ and $\varphi = 180°$. At zero angle (where incidentally, the double Compton effect and the $\ln\lambda$ term in $U^{(1)}$ vanish):

$$b(0°) = -\ln\tau(1+\ln\tau) + 1.355. \quad (49)$$

At 180 degrees,

$$b(180°) = -4.225U. \quad (50)$$

Another simple case results from the condition $\kappa = -2\tau$, in the extreme relativistic limit. This corresponds to 90° scattering in the c.m. system. Here we get $U = 5/2$ and

$$b(90°, \text{c.m.}) = (2y^2 - 3y - 2.29)U. \quad (51)$$

TABLE II. Percent correction to the Compton cross section for unpolarized light arising from Eq. (30), excluding the term proportional to $\ln\lambda$, at zero degrees and at ninety degrees in the c.m. system; computed as a function of the laboratory energy of the incident photon from the special equations (49) and (51).

Laboratory energy Mev	$-(e^2/\pi)b(0°)/U$ percent	$-(e^2/\pi)b(90°, \text{c.m.})/U$ percent
50	3.80	−0.32
150	5.26	−0.87
300	6.41	−2.13
1000	8.80	−4.35

A few results for high energies computed from (49) and (51) are given in Table II. At 180°, the quantity $-(e^2/\pi)b/U$ is +0.98 percent.

VI. NONRELATIVISTIC LIMIT

In the nonrelativistic or Thompson limit our results will be equally valid in the laboratory and c.m. systems, since we will keep only the first nonvanishing terms. In the c.m. system, we let the scattering angle be φ and $|\mathbf{q}_1| = |\mathbf{q}_2| = \omega$. Then

$$\kappa = -2p_1\cdot q_1 = -2[(1+\omega^2)^{\frac{1}{2}}\omega + \omega^2]$$
$$\tau = 2p_1\cdot q_2 = 2[(1+\omega^2)^{\frac{1}{2}}\omega + \omega^2\cos\varphi], \quad (52)$$

so that for $\omega \ll 1$:

$$\kappa = -2\omega(1+\omega+\omega^2/2+\cdots)$$
$$\tau = 2\omega(1+\omega\cos\varphi+\omega^2/2+\cdots). \quad (53)$$

Since κ depends only on ω and $d\tau = \omega^2 d\Omega/\pi$, Eq. (11) becomes

$$d\sigma = (e^4/2m^2)d\Omega(1-2\omega+2\omega^2+\cdots)[U - (e^2/\pi)U^{(1)}]. \quad (54)$$

In Table III we give the nonrelativistic limits of functions occurring in $U^{(1)}$ (Eq. 30). A straightforward calculation then yields:

$$U = 1+\cos^2\varphi+0(\omega) \quad (55)$$

$$U^{(1)} = -(4/3)\omega^2(1-\cos\varphi)U\ln\lambda + (1+\cos\varphi+\cos^2\varphi-\tfrac{1}{3}\cos^3\varphi)4\omega^2\ln\omega + 0(\omega^2). \quad (56)$$

The angular dependence of the $4\omega^2\ln\omega$ term is plotted as $f(\varphi)$ in Fig. 6. Expression (56) disagrees with the result of Schafroth.[2]

Since $(1/\kappa)+(1/\tau) = \frac{1}{2}(1-\cos\varphi) - (\omega/2)\sin^2\varphi$ in the c.m. system and the same invariant is $\frac{1}{2}(1-\cos\theta)$ in the laboratory system, with θ the laboratory scattering angle, it is clear that $\cos\theta = \cos\varphi + 0(\omega)$ for given κ, τ. Also, since $\kappa+\tau = -2\omega_1\omega_2(1-\cos\theta) = -2\omega^2(1-\cos\varphi)$ and $\omega_1 = \omega_2 + 0(\omega_1^2)$, we get that $\omega_1 = \omega + 0(\omega^2) \approx \omega_2$ and therefore (56) is valid with φ interpreted as the laboratory scattering angle and ω the incident photon energy. Equation (55) is valid, with this interpretation, to order ω^2.

The double Compton effect (39) gives in this limit (with U_D defined as in (47))

$$U_D = -(4/3)\omega^2(1-\cos\varphi)U\ln(2k_{\max}/\lambda) + 0(\omega^2). \quad (57)$$

It will be observed that all the corrections vanish in the zero energy limit.

APPENDIX W. ON THE TRANSCENDENTAL FUNCTIONS $G_0(\kappa)$ AND $h(y)$

The complete expression for the radiative correction (29) is expressed in terms of the relatively unfamiliar transcendental integrals $G_0(\kappa)$ and $h(y)$. These can both be expressed, however, in terms of one of the

so-called Spence functions,[14] namely

$$L(x)=\int_0^x \ln(1-u)du/u \tag{A1}$$

which we shall consider briefly.

It is well known that $L(-1)=\pi^2/12$, $L(1)=-\pi^2/6$. If $x<1$,

$$L(x)=-\int_0^x (u+\tfrac{1}{2}u^2+\cdots)du/u$$
$$=-\left(x+\frac{x^2}{2^2}+\frac{x^3}{3^2}+\cdots\right). \tag{A2}$$

If $x>1$,

$$L(x)=L(1)+\int_1^x \ln(1-u)du/u$$
$$=L(1)+\int_1^x [\ln|1-u|\pm i\pi]du/u$$
$$=\bar{L}(x)\pm i\pi \ln x \tag{A3}$$

where

$$\bar{L}(x)=\int_0^x \ln|1-u|\,du/u.$$

For computational convenience, we can also write for $x>1$:

$$\bar{L}(x)=L(1)+\int_{1/x}^1 \ln\left(\frac{1-v}{v}\right)\frac{dv}{v}$$
$$=-\tfrac{1}{3}\pi^2-L(1/x)+\tfrac{1}{2}(\ln x)^2, \tag{A4}$$

and in a similar manner

$$L(-x)=(\pi^2/6)-L(-1/x)+\tfrac{1}{2}(\ln x)^2. \tag{A5}$$

Since $\kappa=(m^2-p_3^2)$ is always negative, $G_0(\kappa)=2\kappa^{-1}[L(1-\kappa)-L(1)]$ has an imaginary part whose sign is not determined by (A3). To fix the sign we must recall that according to the scheme of reference 1(b), all photons and electrons are considered to have a small additional negative imaginary mass. Thus κ has a small positive imaginary part, i.e., $\kappa=-|\kappa|+i\delta$ with δ vanishingly small. Therefore

$$G_0(\kappa)=2\kappa^{-1}[\bar{L}(1-\kappa)-L(1)]+i\pi 2\kappa^{-1}\ln(1-\kappa). \tag{A6}$$

(Similarly, the term $\ln\kappa$ in (30) is equal to $\ln|\kappa|+i\pi$.) We might also note that since τ is always positive, $G_0(\tau)$ has no imaginary part.

Thus if $-\kappa\gg 1$, $\tau\gg 1$ (from A4, A5):

$$G_0(\kappa)\simeq 2\kappa^{-1}[\tfrac{1}{2}(\ln|\kappa|)^2-(\pi^2/6)+i\pi\ln|\kappa|] \tag{A7}$$

$$G_0(\tau)\simeq 2\tau^{-1}[\tfrac{1}{2}(\ln\tau)^2+(\pi^2/3)]. \tag{A8}$$

[14] For references see Fletcher, Miller, and Rosenhead, *An Index of Mathematical Tables* (Scientific Computing Service, Ltd., London, 1946), p. 343.

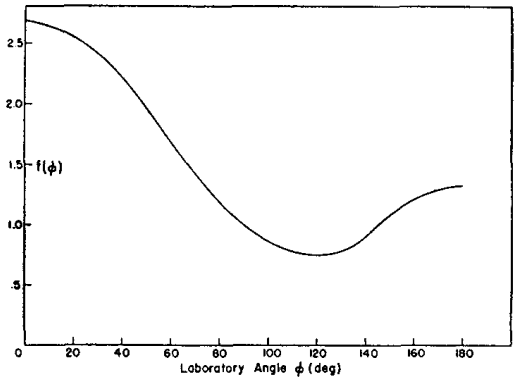


FIG. 6. Angular dependence of nonrelativistic limit of $U^{(1)}$ (without $\ln\lambda$ term).

Our other transcendental function

$$h(y)=y^{-1}\int_0^y u\,du\,\mathrm{ctnh}u$$

can also be expressed in terms of $L(x)$. Integration by parts gives

$$h(y)=\ln(\sinh y)-y^{-1}\int_0^y \ln(\sinh u)du. \tag{A9}$$

Letting $t=e^{-2u}$ yields $\ln(\sinh u)=\ln(1-t)-\ln 2t$ and the integral is obtained directly:

$$h(y)=\ln(2\sinh y)-y/2+(2y)^{-1}[\pi^2/6+L(e^{-2y})]. \tag{A10}$$

If y is sufficiently large that we can neglect e^{-2y} compared to unity, we get

$$h(y)\simeq(y/2)+(\pi^2/12y). \tag{A11}$$

APPENDIX X. TABLE OF INTEGRALS

In this appendix we will simply list the integrals which enter this problem, reserving for Appendix Y a discussion of the methods used. To simplify the presentation of the integrals (which occur also in other problems) we introduce the following definitions:

For factors of the denominator we write $k^2-2p_1\cdot k=(1)$, $k^2-2p_2\cdot k=(2)$, $k^2-2p_3\cdot k-\kappa=(\kappa)$, $k^2=(0)$. We write for frequently occurring vectors

$$p_3=p_1+q_1, \qquad p_4=p_1-q_2,$$
$$2p_0=p_1+p_2, \qquad 2q_0=q_1+q_2,$$
$$2Q=p_1-p_2=q_2-q_1.$$

TABLE III. Nonrelativistic limits of functions occurring in Eq. (30). These expressions are valid in either the c.m. or laboratory system with φ the scattering angle and ω the incident photon energy.

$$y^2=-\tfrac{1}{4}(\kappa+\tau)-(1/48)(\kappa+\tau)^2+0(\omega^6)$$
$$h(y)=1+y^2/9+\cdots$$
$$G_0(\kappa)=2(1+\kappa/2+\kappa^2/9+\cdots)-2\ln|\kappa|(1+\kappa/2+\kappa^2/3+\cdots)$$
$$\ln|\kappa|=\ln 2\omega+\omega+0(\omega^3)$$
$$\ln\tau=\ln 2\omega+\omega\cos\varphi+\tfrac{1}{2}\omega^2(1-\cos^2\varphi)+0(\omega^3)$$

The integrals are defined by

$$A=8i\int d^4k/(1)(2), \qquad B^{(1)}=8i\int d^4k/(1)(\kappa)$$

$$C^{(1)}=8i\int d^4k/(1)(0), \qquad D=8i\int d^4k/(0)(\kappa)$$

$$F=8i\int d^4k/(1)(2)(\kappa), \quad G^{(1)}=8i\int d^4k/(1)(\kappa)(0)$$

$$H=8i\int d^4k/(1)(2)(0), \qquad J=8i\int d^4k/(1)(2)(\kappa)(0).$$

Integrals $B^{(2)}$, $C^{(2)}$, $G^{(2)}$ are defined as $B^{(1)}$, $C^{(1)}$, $G^{(1)}$ but with (2) replacing (1). Their values are obtained from those of $B^{(1)}$, $C^{(1)}$, $G^{(1)}$ by replacing p_1 by p_2, q_1 by q_2, and Q by $-Q$. We use the notation, as in reference 1(b), that

$$F_{(0;\sigma;\sigma\tau)}=8i\int(1;k_\sigma;k_\sigma k_\tau)d^4k/(1)(2)(\kappa), \text{ etc.}$$

Let:

$$\mu=Q^2=-\sinh^2 y=\tfrac{1}{4}(\kappa+\tau), \quad (y \text{ is real as } Q^2<0)$$

$$a=-\ln\Lambda^2$$

$$b=1-y\coth y$$

$$c=(\kappa/\kappa-1)\ln\kappa$$

$$d=2y\,\mathrm{csch}2y=(1-b)/(1-\mu)$$

$$L(x)=\int_0^x\frac{du}{u}\ln(1-u)$$

$$h(y)=(1/y)\int_0^y u\coth u\,du$$

$$v=\kappa^2+4\mu(1-\kappa).$$

Integrals

$$A_0=2a-4b+2$$

$$B_0^{(1)}=2a+2=B_0$$

$$C_0^{(1)}=2a-2=C_0$$

$$D_0=2a-2+2c$$

$$F_0=y^2\,\mathrm{csch}^2 y$$

$$G_0^{(1)}=(2/\kappa)[L(1-\kappa)-L(1)]=G_0$$

$$H_0=2d[-\ln\lambda+h(2y)-h(y)]$$

$$J_0=(2d/\kappa)[2h(y)-h(2y)-\ln(\kappa/\lambda)]$$

$$A_\sigma=(2a-4b+3)p_{0\sigma}$$

$$B_\sigma^{(1)}=(a+\tfrac{3}{2})(p_{1\sigma}+p_{3\sigma})$$

$$C_\sigma^{(1)}=(a+\tfrac{1}{2})p_{1\sigma}$$

$$D_\sigma=[(\kappa/\kappa-1)(c-1)+a+\tfrac{1}{2}]p_{3\sigma}$$

$$F_\sigma=F_0p_{0\sigma}+[F_0-(2b/\mu)]q_{0\sigma}$$

$$G_\sigma^{(1)}=[G_0-(2c/\kappa)]p_{1\sigma}+(2/\kappa)[G_0-(2-\kappa/\kappa)c-2]q_{1\sigma}$$

$$H_\sigma=2dp_{0\sigma}$$

$$vJ_\sigma=[2\mu F_0-(2\mu-\kappa)Z-\kappa G_0]p_{0\sigma}+[(2\mu-\kappa)F_0+2(1-\mu)Z-(2-\kappa)G_0]q_{0\sigma}$$

with

$$Z=H_0+\kappa J_0=2d[-\ln\kappa+h(y)]$$

$$F_{\sigma\tau}=F_\sigma p_{0\tau}+[(F_0-2b/\mu)p_{0\sigma}+(F_0+Y_0-2b/\mu)q_{0\sigma}]q_{0\tau}-Y_0Q_\sigma Q_\tau+[\mu Y_0+\tfrac{1}{2}a+\tfrac{3}{4}]\delta_{\sigma\tau}$$

with

$$Y_0=(F_0-2b-1)/2\mu$$

$$G_{\sigma\tau}^{(1)}=p_{1\sigma}p_{1\tau}\left[G_0-\frac{3\kappa-2}{\kappa-1}\frac{c}{\kappa}+\frac{1}{\kappa-1}\right]+(p_{1\sigma}q_{1\tau}+q_{1\sigma}p_{1\tau})\left[\frac{3}{\kappa}G_0+\frac{2\kappa^2-9\kappa+6}{\kappa(\kappa-1)}\frac{c}{\kappa}+\frac{6-5\kappa}{\kappa-1}\frac{1}{\kappa}\right]+q_{1\sigma}q_{1\tau}\left[\frac{6}{\kappa^2}G_0+\frac{\kappa^3+4\kappa^2-18\kappa+12}{\kappa^2(\kappa-1)}\frac{c}{\kappa}-\frac{2\kappa^2+9\kappa-12}{\kappa(\kappa-1)}\frac{1}{\kappa}\right]+\frac{1}{4}\delta_{\sigma\tau}\left[2G_0-3+2a+2\frac{\kappa-2}{\kappa}c\right]$$

$$H_{\sigma\tau}=dp_{0\sigma}p_{0\tau}+(b/\mu)Q_\sigma Q_\tau+\tfrac{1}{2}\delta_{\sigma\tau}(a-2b+\tfrac{1}{2})$$

$$J_{\sigma\tau}=\alpha_\sigma Q_\tau+\beta_\sigma p_{0\tau}+\gamma_\sigma p_{3\tau}+\epsilon s_{\sigma\tau}$$

where:

$$\epsilon=F_0-\alpha_\sigma Q_\sigma-\beta_\sigma p_{0\sigma}-\gamma_\sigma P_{3\sigma}$$

$$4\mu\alpha_\sigma=G_\sigma^{(1)}-G_\sigma^{(2)}$$

$$v\beta_\sigma=\kappa F_\sigma-(2-\kappa)(H_\sigma+\kappa J_\sigma)+(1-\kappa)(G_\sigma^{(1)}+G_\sigma^{(2)})$$

$$v\gamma_\sigma=(2\mu-\kappa)F_\sigma+2(1-\mu)(H_\sigma+\kappa J_\sigma)-\tfrac{1}{2}(2-\kappa)(G_\sigma^{(1)}+G_\sigma^{(2)})$$

$$s_{\sigma\tau}=\delta_{\sigma\tau}-(Q_\sigma Q_\tau/\mu)+4(1-\kappa)(p_{0\sigma}p_{0\tau}/v)-2(2-\kappa)[(p_{0\sigma}p_{3\tau}+p_{3\sigma}p_{0\tau})/v]+4(1-\mu)(p_{3\sigma}p_{3\tau}/v)$$

$$[s_{\sigma\sigma}=1]$$

$$J_{\sigma\tau\rho}=\alpha_{\sigma\tau}Q_\rho+\beta_{\sigma\tau}p_{0\rho}+\gamma_{\sigma\tau}p_{3\rho}+\epsilon_\sigma s_{\tau\rho}+\epsilon_\tau s_{\sigma\rho}$$

where

$$\epsilon_\sigma=F_\sigma-\alpha_{\sigma\tau}Q_\tau-\beta_{\sigma\tau}p_{0\tau}-\gamma_{\sigma\tau}p_{3\tau}$$

$$4\mu\alpha_{\sigma\tau}=G_{\sigma\tau}^{(1)}-G_{\sigma\tau}^{(2)}$$

$$v\beta_{\sigma\tau}=\kappa F_{\sigma\tau}-(2-\kappa)(H_{\sigma\tau}+\kappa J_{\sigma\tau})+(1-\kappa)(G_{\sigma\tau}^{(1)}+G_{\sigma\tau}^{(2)})$$

$$v\gamma_{\sigma\tau}=(2\mu-\kappa)F_{\sigma\tau}-2(1-\mu)(H_{\sigma\tau}+\kappa J_{\sigma\tau})-\tfrac{1}{2}(2-\kappa)(G_{\sigma\tau}^{(1)}+G_{\sigma\tau}^{(2)}).$$

APPENDIX Y. ON THE INTEGRALS IN THE CORRECTED CROSS SECTION

In this section we will describe the methods by which the integrations occurring in the e^4-order matrix element have been performed and will give some examples of the calculations.

Those integrals which are scalars (those with no k in the numerator) have been done by the parametric method discussed in the appendix to reference 1(b). Those which are tensors (having one or more k's in the numerator) can be done either by the parametric method, or can be derived by an algebraic procedure described below from those of lower tensor order.

We now do several examples of the integrals to indicate the methods employed.

(a) The Two-Denominator Integrals

These all have the form $\chi(\Lambda^2)$ (where we need only the case of Λ much greater than any of the momenta involved in the problem), with

$$\chi=8i\int(1;k_\sigma)(k^2-2p_1\cdot k-\Delta_1)^{-1}$$
$$\times(k^2-2p_2\cdot k-\Delta_2)^{-1}(-\Lambda^2)(k^2-\Lambda^2)^{-1}d^4k. \quad \text{(A12)}$$

To reduce this integral and those considered below we use the methods of reference 1(b).

In the first place we combine the denominators by making use of the relation

$$1/ab=\int_0^1 dy[ay+b(1-y)]^{-2} \quad \text{(A13)}$$

and similar expressions for $1/ab^2$, $1/ab^3$, etc., obtainable from (A13) by differentiation. Thus χ becomes

$$\chi=8i\int_0^1\int_0^1\int(1;k_\sigma)dy2zdz(-\Lambda^2)d^4k$$
$$\times[k^2-2zp_y\cdot k-z\Delta_y-(1-z)\Lambda^2]^{-3} \quad \text{(A14)}$$

with $p_y=p_1y+p_2(1-y)$; $\Delta_y=\Delta_1y+\Delta_2(1-y)$. Using (12a) of reference 1(b), we get

$$\chi=\int_0^1\int_0^1 2zdzdy(-\Lambda^2)(1;zp_{y\sigma})$$
$$\times[z^2p_y^2+z\Delta_y+(1-z)\Lambda^2]^{-1}. \quad \text{(A15)}$$

To facilitate the work we observe that in the limit of very large Λ^2 (with $\delta\ll1$)

$$\int_0^1\frac{(1-z)^ndz(-\Lambda^2)}{P(z)+\Lambda^2(1-z)}$$
$$\cong-\int_0^{1-\delta}(1-z)^{n-1}dz+\int_{1-\delta}^1\frac{(1-z)^ndz(-\Lambda^2)}{P(1)+\Lambda^2(1-z)}$$
$$\cong-1/n, \quad n>0$$
$$\cong\ln[P(1)/\Lambda^2], \quad n=0. \quad \text{(A16)}$$

Writing

$$(1;zp_{y\sigma})=(1-(1-z);[(1-z)^2-2(1-z)+1]p_{y\sigma}),$$

(A15) becomes

$$\chi=\int_0^1 2dy\left\{(1;p_{y\sigma})\ln\frac{p_y^2+\Delta_y}{\Lambda^2}+(1;\tfrac{3}{2}p_{y\sigma})\right\}. \quad \text{(A17)}$$

The second term gives $(2;3(p_{1\sigma}+p_{2\sigma}))$. In the first term

$$\chi_a=\int_0^1 2dy(1;p_{y\sigma})\ln[(p_y^2+\Delta_y)/\Lambda^2], \quad \text{(A18)}$$

notice that with $2Q=(p_1-p_2)$,

$$p_y^2=p_2^2+Q^2y^2+2p_2\cdot(p_1-p_2)y,$$
$$\Delta_1=m^2-p_1^2, \quad \Delta_2=m^2-p_2^2$$

so that

$$p_y^2+\Delta_y=m^2+4Q^2(y^2-y)=m^2+Q^2[(2y-1)^2-1].$$

Now let $m=1$, $Q^2=\sin^2\theta$, and $2y-1=\tan\alpha/\tan\theta$, so that $dy(\sec^2\alpha/2\tan\theta)d\alpha$. We get

$$p_y^2+\Delta_y=\cos^2\theta\sec^2\alpha$$

so that

$$\chi_a=\int_{-\theta}^{\theta}(\sec^2\alpha/\tan\theta)d\alpha(1;p_{y\sigma}(\alpha))$$
$$\times\ln(\cos^2\theta\sec^2\alpha/\Lambda^2). \quad \text{(A19)}$$

This integral can be done easily. For example,

$$\int_{-\theta}^{\theta}\sec^2\alpha d\alpha\ln(\sec^2\alpha)=\int_{-\tan\theta}^{\tan\theta}dy\ln(1+y^2)$$
$$=4\tan\theta[\ln(\sec\theta)-1]+4\theta,$$

the last integral being performed by parts. In this manner we obtain integrals A, B, C, D.

(b) The Intregal $G_0^{(1)}$

As an example of the three denominator integrals we integrate $G_0^{(1)}$. The parameterization method gives

$$G_0^{(1)}=8i\int d^4k(k^2-2p_1\cdot k)^{-1}(k^2-2p_3\cdot k-\kappa)^{-1}(k^2)^{-1}$$
$$=8i\int_0^1\int_0^1\int d^4k2dyxdx(k^2-2p_z\cdot k-\Delta_z)^{-3} \quad \text{(A20)}$$

with

$$p_y=(1-y)p_1+yp_3=p_1+yq_1, \quad p_z=xp_y, \quad \Delta_z=xy\kappa.$$

Using (12a) of reference 1(b) again,

$$G_0^{(1)}=\int_0^1\int_0^1 2dxdy(xp_y^2+y\kappa)^{-1}. \quad \text{(A21)}$$

Since

$$\int_0^1 dx(ax+b)^{-1}=a^{-1}\ln(b+a)/b,$$

and since $p_y^2=1-\kappa y$

$$G_0^{(1)}=-\int_0^1 2dy(1-\kappa y)^{-2}\ln\kappa y$$

$$=-\frac{2}{\kappa}\int_{1-\kappa}^1\frac{dv}{v}\ln(1-v). \quad \text{(A22)}$$

In the last step we have let $v=1-\kappa y$. We obtain finally:

$$G_0^{(1)}=G_0^{(2)}=G_0=(2/\kappa)[L(1-\kappa)-L(1)]. \quad \text{(A23)}$$

(c) The Integrals H

These are the same as those done in the radiationless scattering problem. They are given in reference 1(b), appendix. Equations (23a), (24a), and (25a) should have the signs of their left-hand sides changed.

(d) The Integral J_0

$$J_0=8i\int d^4k(k^2-2p_1\cdot k)^{-1}(k^2-2p_3\cdot k-\kappa)^{-1}$$

$$\times(k^2-2p_2\cdot k)^{-1}(k^2-\lambda^2)^{-1}. \quad \text{(A24)}$$

We have given the photon k a small mass λ as an infrared cutoff. When this integral is parameterized it becomes

$$J_0=8i\int_0^1\int_0^1\int_0^1\int 6dxdydz(1-x)z^2d_4k$$

$$\times[k^2-2zp_x\cdot k-xz\kappa-\lambda^2(1-z)]^{-4}. \quad \text{(A25)}$$

By (13a) of reference 1(b),

$$8i\int d^4k(k^2-2p\cdot k-\Delta)^{-4}=-\tfrac{1}{3}(p^2+\Delta)^{-2}$$

so that

$$J_0=-\int_0^1\int_0^1\int_0^1 2dxdydz(1-x)z^2$$

$$\times[z^2p_x^2+xz\kappa+\lambda^2(1-z)]^{-2} \quad \text{(A26)}$$

with $p_x=(1-x)p_y+xp_3$, $p_y=yp_1+(1-y)p_2$.

We break the x integration into two regions ($\epsilon\ll1$):

$$\int_0^1 dx\int_0^1 dy\int_0^1 dz=\int_0^1 dy\int_\epsilon^1 dx\int_0^1 dz+\int_0^1 dy\int_0^\epsilon dx\int_0^1 dz.$$

(I) (II)

In (I) we let $\lambda\to0$, getting

$$\text{(I)}=\int_0^1 dy\int_\epsilon^1 dx\int_0^1 2dz(1-x)(zp_x^2+x\kappa)^{-2}$$

$$=2\int_0^1\int_\epsilon^1\frac{(1-x)dxdy}{\kappa x(p_x^2+\kappa x)} \quad \text{(A27)}$$

In region (II), since x is small, we neglect x^2 compared to x:

$$\text{(II)}=\int_0^1 dy\int_0^1 dz\int_0^\epsilon 2z^2dx[z^2p_y^2+2xz^2a+z\kappa x+\lambda^2(1-z)]^{-2}$$

$$=2\int_0^1 dy\int_0^1 dz\epsilon z^2[z^2p_y^2+\lambda^2(1-z)]^{-1}$$

$$\times[z^2p_y^2+2\epsilon z^2a+z\kappa\epsilon+\lambda^2(1-z)]^{-1} \quad \text{(A28)}$$

with $a=p_y\cdot(p_3-p_y)$.

In (II) we now break the z integration into two regions; $0\le z\le z_c$ and $z_c\le z\le1$ such that $\lambda^2\ll z_c^2p_y^2\ll z_c\kappa\epsilon$. Thus for $z<z_c$ we neglect z relative to unity and $z^2p_y^2$ relative to $z\kappa\epsilon$. For $z>z_c$ we neglect λ. There results:

$$\int_{z_c}^1\frac{2\epsilon dzdy}{p_y^2z[z(p_y^2+2\epsilon a)+\kappa\epsilon]}=\frac{2}{\kappa p_y^2}\ln\frac{\kappa\epsilon}{z_cp_y^2},$$

$$\int_0^{z_c}\frac{2\epsilon z^2dzdy}{(z^2p_y^2+\lambda^2)(z\kappa\epsilon+\lambda^2)}$$

$$=-\int_0^{z_c}\frac{2\epsilon dz}{(\lambda^2p_y^2+\kappa^2\epsilon^2)}\left(\frac{\lambda^2-\kappa\epsilon z}{z^2p_y^2+\lambda^2}-\frac{\lambda^2}{\kappa\epsilon z+\lambda^2}\right)$$

$$=\frac{1}{\kappa p_y^2}\ln\frac{z_c^2p_y^2}{\lambda^2}\text{(neglecting terms of order }\lambda).$$

Adding these together we get

$$\text{(II)}=\int_0^1\frac{dy}{\kappa p_y^2}\ln\frac{\kappa^2\epsilon^2}{\lambda^2p_y^2}. \quad \text{(A29)}$$

We can now use the same substitution for y as that leading to (A8). Notice that in this case $\Delta_y=0$. We get $dy/\kappa p_y^2=\mp d\alpha/\sin2\theta$, $p_y^2=\cos^2\theta/\cos^2\alpha$ so that (II) becomes:

$$\text{(II)}=\int_{-\theta}^{\theta}\frac{2d\alpha}{\sin2\theta}\left[\ln\frac{\kappa\epsilon}{\lambda\cos\theta}+\ln(\cos\alpha)\right]$$

$$=\frac{4\theta}{\sin2\theta}\ln\frac{\kappa\epsilon}{\lambda\cos\theta}+\int_0^\theta\frac{4d\alpha}{\sin2\theta}\ln(\cos\alpha). \quad \text{(A30)}$$

We have still to finish evaluating (I), Eq. (A27). First investigate the denominator $p_x^2+\kappa x$. Since $p_3^2=1-\kappa$ and $2p_3\cdot p_y=2-\kappa$,

$$p_x^2=(1-x)^2p_y^2+x^2(1-\kappa)+x(1-x)(2-\kappa)$$
$$p_x^2+\kappa x=(1-x)^2(p_y^2-1)+1.$$

Letting $1-x=\sin\phi/\sin\theta$, $2y-1=\tan\alpha/\tan\phi$ with $Q^2=\sin^2\theta$, we get

$$p_y^2=\sin^2\theta[(\tan^2\alpha/\tan^2\phi)-1],$$
$$p_x^2+\kappa x=\sin^2\phi\ \mathrm{ctn}^2\phi(\sec^2\alpha-\sec^2\phi)+1=\cos^2\phi\ \sec^2\phi,$$
$$-dx=\cos\phi d\phi/\sin\phi,\quad dy=\sec^2\alpha d\alpha/2\tan\phi.$$

The integrand of (I) becomes

$$\frac{1}{\kappa}\left(\frac{\sin\theta}{\sin\theta-\sin\phi}\right)\left(\frac{\sin\phi}{\sin\theta}\right)\left(\frac{\cos^2\alpha}{\cos^2\phi}\right)\left(\frac{\cos\phi d\phi}{\sin\theta}\right)\left(\frac{\sec^2\alpha d\alpha}{\tan\phi}\right)$$
$$=\frac{1}{\kappa\sin\theta}\frac{d\phi d\alpha}{\sin\theta-\sin\phi}$$

and

$$(\mathrm{I})=\int_0^{\theta-\epsilon\tan\theta}\frac{d\phi}{\kappa}\int_{-\phi}^{\phi}\frac{d\alpha}{\sin\theta(\sin\theta-\sin\phi)}$$
$$=\int_0^{\theta-\epsilon\tan\theta}\frac{2\phi d\phi}{\kappa\sin\theta(\sin\theta-\sin\phi)}. \quad \text{(A31)}$$

Integrate (A31) by parts, using No. 436, Dwight's *Tables of Integrals*, to get

$$(\mathrm{I})=-\frac{4\theta}{\kappa\sin2\theta}\ln\left(\frac{\epsilon\tan\theta}{2\cos\theta}\right)$$
$$+\int_0^{\theta}\frac{4d\phi}{\sin2\theta}\ln\left(\frac{\sin\frac{1}{2}(\theta-\phi)}{\cos\frac{1}{2}(\theta+\phi)}\right) \quad \text{(A32)}$$

which can be put in the form

$$(\mathrm{I})=-\frac{4\theta}{\kappa\sin2\theta}\ln\frac{\epsilon\sin\theta}{\cos^2\theta}-\int_0^{\theta}\frac{4d\alpha}{\sin2\theta}\ln(\cos\alpha)$$
$$+\int_0^{\theta}\frac{4d\alpha}{\sin2\theta}\ln(\tan\alpha)$$

and can now be added to (A19).

This gives the result,

$$J_0=\frac{-4\theta}{\kappa\sin2\theta}\left[\ln\frac{\kappa}{\lambda\tan\theta}+\theta^{-1}\int_0^{\theta}dv\ln(\tan v)\right]. \quad \text{(A33)}$$

Another integration by parts and the substitution $\theta=iy$ gives the result in the table:

$$J_0=(2d/\kappa)[2h(y)-h(2y)-\ln(\kappa/\lambda)]. \quad \text{(A34)}$$

(e) The Integrals J_σ, $J_{\sigma\tau}$, $J_{\sigma\tau\nu}$

From the preceding work on J_0 it may be supposed that to attempt these more complicated integrals by the parametric method would involve great labor. Fortunately there is a way to reduce these integrals to a combination of integrals of a lower tensor order, and those of a smaller number of denominators. We will do J_σ as an example but it should be clear from this how $J_{\sigma\tau}$ and $J_{\sigma\tau\nu}$ are done. This method can of course, be applied also to G_σ, $G_{\sigma\tau}$, etc. We shall be able to express J_σ in terms of J_0 and the integrals F, G having only three denominators.

Using the notation indicated in Appendix X for the denominators, $(1)=k^2-2p_1\cdot k$, etc., we write

$$J_\sigma=8i\int\frac{k_\sigma d^4k}{(1)(2)(\kappa)(0)}=\alpha p_{1\sigma}+\beta p_{2\sigma}+\gamma p_{3\sigma}, \quad \text{(A35)}$$

α, β, γ being scalar functions of p_1, p_2, and p_3. The vectors p_1, p_2, p_3 will in general define a three-space. It is clear that the vector J_σ cannot have a component in the direction P which is perpendicular to this three-space, since for k_σ in the P direction, the integrand is an odd function and therefore J_σ must vanish.

If we now take the scalar products of J_σ with p_1, p_2, p_3, since $2p_1\cdot k=(0)-(1)$, etc., we get

$$2p_1\cdot J=8i\int\frac{2p_1\cdot kd^4k}{(1)(2)(\kappa)(0)}$$
$$=8i\int\frac{d^4k}{(1)(2)(\kappa)}-8i\int\frac{d^4k}{(\kappa)(2)(0)}=F_0-G_0$$
$$2p_2\cdot J=8i\int\frac{2p_2\cdot kd^4k}{(1)(2)(\kappa)(0)}$$
$$=8i\int\frac{d^4k}{(1)(2)(\kappa)}-8i\int\frac{d^4k}{(\kappa)(2)(0)}=F_0-G_0 \quad \text{(A36)}$$
$$2p_3\cdot J=8i\int\frac{2p_3\cdot kd^4k}{(1)(2)(\kappa)(0)}$$
$$=8i\int\frac{d^4k}{(1)(2)(\kappa)}-8i\int\frac{d^4k}{(1)(2)(0)}$$
$$-8i\int\frac{\kappa d^4k}{(1)(2)(\kappa)(0)}$$
$$=F_0-H_0-\kappa J_0.$$

Taking also the scalar products with the right-hand side of (A35), we get the set of linear equations:

$$F_0-G_0=2\alpha+\beta(1-\tfrac{1}{2}\kappa-\tfrac{1}{2}\tau)+\gamma(2-\kappa)$$
$$F_0-G_0=\alpha(1-\tfrac{1}{2}\kappa-\tfrac{1}{2}\tau)+2\beta+\gamma(2-\kappa) \quad \text{(A37)}$$
$$F_0-H_0-\kappa J_0=(\alpha+\beta)(2-\kappa)+2\gamma(1-\kappa).$$

J_σ can now be readily obtained by solving these equations for α, β, and γ.

To obtain $J_{\sigma\tau}$ we write

$$J_{\sigma\tau}=8i\int\frac{k_\sigma k_\tau d^4k}{(1)(2)(\kappa)(0)}$$
$$=\alpha_\sigma p_{1\tau}+\beta_\sigma p_{2\tau}+\gamma_\sigma p_{3\tau}+\epsilon\delta_{\sigma\tau}, \quad \text{(A38)}$$

α_σ, β_σ, γ_σ being vector functions of p_1, p_2, p_3, and ϵ a scalar function of the same variables. The tensor $\epsilon\delta_{\sigma\tau}$ now occurs on the right-hand side as it is possible for $J_{\sigma\tau}$ to have nonzero components depending on P. (If $k_\sigma=k_\tau=k_P$, $J_{\sigma\tau}$ need not vanish.) If we take inner products with p_1, p_2, and p_3 now we get

$$F_\sigma-G_\sigma^{(1)}=2\alpha_\sigma+2\beta_\sigma(1-\tfrac{1}{2}\kappa-\tfrac{1}{2}\tau)+\gamma_\sigma(2-\kappa)+2\epsilon p_{1\sigma}$$

for $2p_{1\tau}J_{\sigma\tau}$ and similar equations for p_2 and p_3. This gives us three equations for the four quantities α_σ, β_σ, γ_σ, ϵ. However, there is the additional independent result obtained by summing $J_{\sigma\sigma}$ over σ:

$$J_{\sigma\sigma}=8i\int\frac{k_\sigma k_\sigma d^4k}{(1)(2)(\kappa)(0)}$$
$$=F_0=\alpha_\sigma p_{1\sigma}+\beta_\sigma p_{2\sigma}+\gamma_\sigma p_{3\sigma}+4\epsilon. \quad \text{(A39)}$$

Solving these four equations we obtain $J_{\sigma\tau}$ algebraically in terms of simpler integrals. In a similar manner $J_{\sigma\tau\nu}$ can be expressed in the form given in Appendix X.

APPENDIX Z. EXAMPLES OF THE CALCULATIONS

We shall here illustrate by two examples the method of evaluation of the transition amplitude.

The matrix T (18) can be written, if we rationalize the denominator, as ($p^2=m^2$, $q^2=0$)

$$T=\int(k^2-2p\cdot k-2q\cdot k-\kappa)^{-1}(k^2-2p\cdot k)^{-1}$$
$$\times d^4k k^{-2}C(k^2)\boldsymbol{T} \quad \text{(A40)}$$

with

$$\boldsymbol{T}=\gamma_\mu(\boldsymbol{p}+\boldsymbol{q}-\boldsymbol{k}+m)\boldsymbol{e}(\boldsymbol{p}-\boldsymbol{k}+m)\gamma_\mu. \quad \text{(A41)}$$

T can now be split into terms involving no k, one k, and two k's, and the results of the integrations over k-space inserted from Appendix Y. Thus,

$$\boldsymbol{T}=\gamma_\mu(\boldsymbol{p}+\boldsymbol{q}+m)\boldsymbol{e}(\boldsymbol{p}+m)\gamma_\mu-\gamma_\mu\boldsymbol{k}\boldsymbol{e}(\boldsymbol{p}+m)\gamma_\mu$$
$$-\gamma_\mu(\boldsymbol{p}+\boldsymbol{q}+m)\boldsymbol{e}\boldsymbol{k}\gamma_\mu+\gamma_\mu\boldsymbol{k}\boldsymbol{e}\boldsymbol{k}\gamma_\mu$$
$$=2\boldsymbol{p}(\boldsymbol{p}+\boldsymbol{q}+m)\boldsymbol{e}-2\boldsymbol{p}\boldsymbol{k}\boldsymbol{e}+2\boldsymbol{k}\boldsymbol{e}(\boldsymbol{p}+\boldsymbol{q})$$
$$-4m(e\cdot k)-2\boldsymbol{k}\boldsymbol{e}\boldsymbol{k},$$

where we have used Eq. (4a) of reference 1(b) and the fact that the matrix T operates from the left on a state u such that $\boldsymbol{p}u=mu$. Inserting the integrals and grouping terms we now have,

$$8i\boldsymbol{T}=2\boldsymbol{p}(\boldsymbol{p}+\boldsymbol{q}+m)\boldsymbol{e}G_0-2(\boldsymbol{p}\gamma_\sigma\boldsymbol{e}-\gamma_\sigma\boldsymbol{e}(\boldsymbol{p}+\boldsymbol{q})+2me_\sigma)G_\sigma$$
$$-2\gamma_\sigma\boldsymbol{e}\gamma_\tau G_{\sigma\tau}. \quad \text{(A42)}$$

This expression can now be further simplified. For example, the term in G_σ is (using Appendix X)

$$-2[G_0-(2c/\kappa)][\boldsymbol{p}^2\boldsymbol{e}-\boldsymbol{p}\boldsymbol{e}(\boldsymbol{p}+\boldsymbol{q})+2m(e\cdot p)]$$
$$-(4/\kappa)[G_0-(2-\kappa/\kappa)c-2][\boldsymbol{p}\boldsymbol{q}\boldsymbol{e}-\boldsymbol{q}\boldsymbol{e}(\boldsymbol{p}+\boldsymbol{q})]$$

with $\kappa=-2p\cdot q$. Since $p^2=m^2$, $q^2=0$, $e\cdot q=0$, and quite generally $\boldsymbol{a}\boldsymbol{b}+\boldsymbol{b}\boldsymbol{a}=2a\cdot b$, this term becomes finally

$$-2[G_0-(2c/\kappa)](2m^2\boldsymbol{e}-\boldsymbol{p}\boldsymbol{e}\boldsymbol{q})$$
$$+(4/\kappa)[G_0-(2-\kappa/\kappa)c-2]\cdot[2(e\cdot p)\boldsymbol{q}+\kappa\boldsymbol{e}].$$

Combining this with the terms in G_0 and $G_{\sigma\tau}$ in (A42) (expanded in the same manner) we obtain the expression for T given in the text (19).

To illustrate the simplification that occurs upon taking the spur, for unpolarized light, consider the term J (21). This may be decomposed, as was T above, into a sum of terms involving various numbers of k's in the numerator. For example, the term involving three k's is

$$-\int\gamma_\mu\boldsymbol{k}\boldsymbol{e}_2\boldsymbol{k}\boldsymbol{e}_1\boldsymbol{k}\gamma_\mu d^4k/(1)(2)(\kappa)(0) \quad \text{(A43)}$$

using the notation of Appendix X. If we replace $\boldsymbol{e}_1$ by γ_α and $\boldsymbol{e}_2$ by γ_β, this gives a contribution to the spur $P(\kappa,\tau)$, Eq. (28), of a numerical factor times

$$\int g(k)d^4k/(1)(2)(\kappa)(0) \quad \text{(A44)}$$

with

$$g(k)=\mathrm{Sp}[(\boldsymbol{p}_2+m)\gamma_\mu\boldsymbol{k}\gamma_\beta\boldsymbol{k}\gamma_\alpha\boldsymbol{k}\gamma_\mu(\boldsymbol{p}_1+m)\bar{\boldsymbol{W}}]$$
$$=\kappa^{-1}\{4(p_1\cdot k)(p_2\cdot k)(p_3\cdot k)$$
$$+k^2[\tfrac{1}{2}m^2\kappa(p_1+p_2)\cdot k-(p_1\cdot p_2+2m^2)(p_3\cdot k)]\}$$
$$+\tau^{-1}\{4(p_1\cdot k)(p_2\cdot k)(p_4\cdot k)$$
$$-m^2k^2(p_1+p_2+p_3)\cdot k\}. \quad \text{(A45)}$$

It will now be seen that the factor k^2 will cancel the factor (0) in the denominator of (A44) leading to the integral F_σ (Appendix X). Also, we can write

$$2p_1\cdot k=-(k^2-2p_1\cdot k)+k^2=-(1)+(0)$$

leading to integrals $G_{\sigma\tau}^{(2)}$ and $F_{\sigma\tau}$. Thus it is not necessary, for unpolarized light, to use integral $J_{\sigma\tau\rho}$. (A44) can now be written

$$2(\kappa^{-1}p_{2\sigma}p_{3\tau}+\tau^{-1}p_{2\sigma}p_{4\tau})(F_{\sigma\tau}-G_{\sigma\tau}^{(2)})$$
$$+[\tfrac{1}{2}m^2(p_{1\sigma}+p_{2\sigma})-\kappa^{-1}(p_1\cdot p_2+2m^2)p_{3\sigma}$$
$$-m^2\tau^{-1}(p_{1\sigma}+p_{2\sigma}+p_{3\sigma})]F_\sigma. \quad \text{(A46)}$$

THE PRESENT STATUS OF QUANTUM ELECTRODYNAMICS

by **Richard P. FEYNMAN**

California Institute of Technology, Pasadena, California.

Fifty years ago at this Conference one of the problems most energetically discussed was the apparent quantum nature of the interaction of light and matter. It is a privilege to be able, after half a century, to give a report on the progress that has been made in its solution. No problem can be solved without it dragging in its wake new problems to be solved. But the incompleteness of our present view of quantum electrodynamics, although presenting us with the most interesting challenges, should not blind us to the enormous progress that has been made. With the exception of gravitation and radioactivity, all of the phenomena known to physicists and chemists in 1911 have their ultimate explanation in the laws of quantum electrodynamics.

Stricktly speaking, " quantum electrodynamics " might be expected to deal only with the quantum theory of the electromagnetic field, and not with the theory of the motion of the matter which generates it or reacts to it. But conventionally, the motion of that matter whose motion is understood, namely electrons and possibly muons, is included, while the motion of baryons and mesons is not. I will use the term in the conventional sense here. If I wish to refer to the narrower field I will call it simply the " quantum theory of the electromagnetic field ".

Lorentz [1] showed in his 1911 report at this conference that beside an instantaneous coulomb interaction the electromagnetic field could be represented as a set of harmonic oscillators, each

driven by the transverse component of the current produced by matter in the corresponding mode. That the quantum theory of electromagnetic interaction results directly from the simple assumption that these oscillators are quantum oscillators obeying a Schrodinger equation was noted by Dirac [2]. Since that time a bewildering variety of mathematically equivalent formulations of that idea have been made. These are published [3] in many articles and text books and I will assume that you are familiar with some of them, and will not discuss them further. Here I shall simply report first on the comparison of quantum electrodynamic calculations with experiment, and second on some of the unanswered theoretical questions in this field.

COMPARISON TO EXPERIMENT

General remarks.

Considerable evidence for the general validity of Q.E.D. is, of course, provided by the enormous variety of ordinary phenomena which, under rough calculation, are seen to be consistent with it. The superfluidity of helium and the superconductivity of metals having recently been explained, there are to my knowledge no phenomena occuring under known conditions, where quantum electrodynamics should provide an explanation, and where at least a qualitative explanation in these terms has not been found. The search for discrepancies has turned from looking for gross deviations in complex situations to looking either for large discrepancies at very high energies, or by looking for tiny deviations from the theory in very simple, but very accurately measured situations.

High energy experiments.

The experiments at high energy which are most significant for us here are those of the elastic scattering of energetic electrons (up to 1 Gev) by protons at appreciable angles [4]. The scattering is very different from what it would be for an unstructured proton. The proton should have some structure, however, as a result of the unknown strong interactions between mesons and baryons. One usually interprets all the deviation as due to this structure. On the other hand some of it may be a failure of quantum electrodynamics.

According to quantum electrodynamics the scattering amplitude should be

$$(\bar{u}_2 \gamma_\mu u_1) J\mu \frac{1}{q^2} \tag{1}$$

where q is the momentum transferred by the virtual photon, u_1 and u_2 are electron spinors in and out and J_μ is the matrix element of electric current between the nucleon states of four momentum P_1 and P_2. From relativistic invariance arguments J_μ must have the form

$$\gamma_\mu F_1(q^2) + \sigma_{\mu\nu} q_\nu F_2(q) \tag{2}$$

taken between proton spinors, where F_1 and F_2 are unknown functions of q^2. For a point particle $F_2 = 0$, $F_1 = 1$.

It is difficult to say what would happen if electrodynamics failed, as long as the exact manner of a supposed failure is not specified. A conventional way to assume the failure is to suppose that the propagator is altered from $1/q^2$ to

$$1/q^2 (1 - q^2/\Lambda^2) \tag{3}$$

and to tell how large Λ would have to be for such a modification to remain undetected in a given experiment.

In the proton scattering the effective $F_1(q^2)$ is found to fall, for smaller q^2, as $1 + q^2/(560 \text{ Mev})^2$ (q^2 is negative). If F_1 did not fall at all, but the altered propagator were responsible Λ would be 560 Mev. If Λ is much less than this the proton would look much softer and extended than it does. We can probably safely conclude from this experiment that Λ exceeds 500 Mev.

According to (1) the scattering for different angle and energies, *providing they correspond to the same q* should all be related via just two unknown numbers F_1, F_2. This will fail to be exactly so because of corrections, probably not large, due to the exchange of two or more photons. These corrections may be computed, although with some uncertainty due to proton structure. If a large deviation still persists it would mean that Q.E.D. fails in a very peculiar way — for example, that another virtual object of higher spin is exchanged, or that there is a new coupling of proton and electron so that they may combine to form a neutral heavy particle which disintegrates back again to proton and electron, etc.

So far most physicists believe that the general behavior of the functions F_1 and F_2 is understandable as a proton structure effect and therefore that Q.E.D. can be trusted to perhaps at least as high as 1 Gev.

Clearly these uncertainties of proton structure would not arise if electrons or positrons were scattered from electrons. For example, there are measurements to 5% accuracy of the cross section for annihilation of positrons in flight [5] up to laboratory energies of nearly 10 Gev. In the center of gravity system, however, the positron momentum is only 50 Mev and the virtual state momenta of importance are much lower still. The experiments agree with theory but this does not put a very great lower limit on Λ. The same comment applies to μ^-, e^- collisions measured [6] up to 8 Gev, also in agreement with theory.

On the other hand, these experiments must not be treated too lightly. They are only uninteresting if they agree with theory. Although they do not yet involve nearly as high virtual energies as the proton scattering experiment, they test different things, such as the electron propagator, or the muon structure.

The types of experiments at high energy which would more effectively test the predictions of Q.E.D. are discussed by J. Bjorken and S. Drell, *P. R.*, **114**, 1368 (1959).

Energy levels in hydrogen.

Turning now to precision low energy experiments, the classic experiment [7] is the direct measure of the $2s_{1/2}$ and $2p_{1/2}$ energy separation in hydrogen, deuterium and ionized helium. It was the analysis of this experiment by Weisskopf and by Bethe which led them to discover a way to circumvent the divergent self-energy which, up to then, had bedeviled any attempt to compute higher order effects from Q.E.D. The need to put their ideas into a relativistically invariant form led to the formulations of Schwinger and of the author. The Lamb effect still remains one of the most delicate tests of Q.E.D. A comparison [8] of theory and experiment is given in Table I (after Peterman [9]).

Contributions to the Lamb shift arise from several sources :

(*a*) *Virtual emission and readsorption of one virtual photon.*

Here in the initial state an electron is in a definite state ($2s_{1/2}$ or $2p_{1/2}$) in the nuclear Coulomb potential. In the intermediate state it is in some other exact state of the Coulomb potential. The wave function for these states should be found by the Dirac equation. The labor in doing this has been too great, so far, even for computing machines, because for each intermediate state n, matrix elements of the current times exp. $i\vec{K}\cdot\vec{X}$ must be found for every K and a double integral on $\vec{K}$ and n is involved. However for low energy photons the dipole approximation and Schrodinger wave functions can be used and the sums performed. This determines [10] the constants $K_o(2,0)$ and $K_o(2,1)$. For higher intermediate energy it is usual to make an expansion of the intermediate wave functions as plane waves, perturbed by the Coulomb potential (hence an expansion in orders of $Z\alpha$). Combining the first term with the low energy contribution gives the largest part of the Lamb shift, items 1 plus 2. Item 2 is separated here, for it is easy to understand as the correction to fine structure due to the apparent anomalous moment of the electron.

The correction to include two potential scatterings is included [11] in item 4. To include three scatterings is very difficult, but Layzer [12] has shown that it is a quadratic form in $ln(Z\alpha)$ and has computed everything but the constant term. This makes an uncertainty in the total effect from one virtual photon. It is unlikely that the part marked " ? " exceeds $\pm$ 10, so we may take this as a kind of limit of error. Terms of higher order in the potential are probably too small to be significant.

(b) *Emission and reabsorption of two virtual photons.*

This is of order one higher in α, and no large logarithm from low energy photons arises, so the effect is very small indeed. Although the magnetic moment part, item 7, has been worked out exactly [13], the potential spreading effect, item 6, has only been partially evaluated, in such a way that limits of error can be given [14]. This work should be completed because the uncertainty here is the largest contributor to the theoretical uncertainty in the Lamb shift.

TABLE I. — (After Petermann)[9].

Lamb shift : $2S_{1/2}$-$2P_{1/2}$ $(1/\alpha = 137.0389)$

Order		Formula, units $Z^4L(\mu/m)^3$	H	D	He^+
1. $\alpha(Z\alpha)^4$	Rad	$-2lnZ\alpha + \frac{m}{M} + \frac{11}{24} - ln\frac{K_o(2,0)}{K_o(2,1)}$	1009.85	1010.64	13168.4
2. $\alpha(Z\alpha)^4$	Mag Mom	$\frac{1}{2}(1 + \frac{1}{4}\frac{m}{M})$	67.71	67.77	1084.7
3. $\alpha(Z\alpha)^4$	Vac Pol	$-\frac{1}{5}$	—27.08	—27.11	—433.9
4. $\alpha(\alpha Z)^5$	2nd order V	$3\pi Z\alpha\,(1 + \frac{11}{128} - \frac{1}{2}ln2 + \frac{5}{192})$	7.14	7.13	228.4
5. $\alpha(\alpha Z)^6$	3rd order V	$(\alpha Z)^2[3ln^2Z\alpha + (4ln2 + 1 + \frac{7}{48})lnZ\alpha + ?]$	—.25 ± .07	—.25 ± .07	—9.5 ± 4.
6. $\alpha^2(\alpha Z)^4$	4th order Rad	$\frac{3\alpha}{2\pi}(0.52 \pm .3)$	.24 ± .13	.24 ± .13	3.9 ± 2.
7. $\alpha^2(\alpha Z)^4$	Mag Mom	$-.328\,\alpha/\pi$	—.11	—.11	—1.7
8. $\alpha^2(\alpha Z)^4$	Vac Pol	$-\frac{41}{54}\frac{\alpha}{\pi}$	—.24	—.24	—3.8
9. $\alpha(\alpha Z)^4\frac{Zm}{M}$	Finite Mass	$\frac{Zm}{M}[4.862 - \frac{1}{2}ln\,Z]$	.36	.18	2.5
10.	Finite size	$\alpha(m/\alpha)^2 \langle R^2\rangle_{nuc.}$	.12 ± .02	.73 ± .02	7.1 ± 1.
		Total, Mc	1057.74 ± .22	1058.98 ± .22	14046.1 ± 7.
		Exp. Mc	1057.77 ± .10	1059.00 ± .10	14040.2 ± 4.5

$L = \alpha^3\,Ryd_\infty\,c/3\pi = 135.6353Mc$, $\mu/m = (1 + m/M)^{-1}$

m = mass of electron, M = mass of nucleus.

(c) *Vacuum polarization.*

The effective potential from the nucleus is altered by the existence of virtual pairs of electrons and positrons created by the potential. This has the main effect given in item 3. A correction of relative order α has also been calculated [15], item 8. Corrections of order $Z\alpha$ to the vacuum polarization are included in item 4 (the term 5/192). In fact the vacuum polarization has been calculated [16] for arbitrarily strong fields (arbitrary $Z\alpha$), but additional corrections to the term first order in $Z\alpha$ are small even for Pb.

(d) *Finite nuclear mass.*

The biggest correction resulting from the finite nuclear mass is the correction to the probability of finding the electron at the nucleus due to use of the reduced mass μ rather than the electron mass m in the Schrodinger equation. Beside this the mass appears in the logarithms for item (1) and in the fine structure correction item (2) in a way that is readily evaluated. There are, however, additional corrections of order Zm/M from two photon exchanges between the electron and the recoiling nucleus. They are given by [17] [18]

$$Z\frac{m}{M}L\left(2ln\frac{2K_o(2,1)}{Z\alpha K_o(2,0)}+\frac{83}{24}+\frac{3}{2}\left[ln\,Z\alpha+\frac{7}{4}+\frac{4}{3}(1-ln2)\right]\right)$$

and are included in item 9. (My figures seem to differ slightly from those of Peterman.)

(e) *Nuclear structure.*

If the nucleus is not a point charge the potential near the nucleus is slightly altered, perturbing the s state but not the p state in first approximation. This is an effect calculated by elementary perturbation theory if the mean square radius of the nuclear charge density $<R^2>_{nuc.}$ is known. This can be got directly from scattering experiments and the effect evaluated. The error is just a reflection of experimental error.

(f) *Higher order terms.*

Terms in next order in α should probably contribute at most a few hundredths of a Mc to H and D and perhaps up to 1 Mc to He^+.

The last three columns give the shift calculated in megacycles from each of these terms for H, D, and He^+. They are calculated using $1/\alpha = 137.0389$. If $1/\alpha$ is larger than this by ε the correction to the theoretical value for H and D is -22ε Mc which is almost certainly less than $\pm .02$ Mc (the present uncertainty [8] in $1/\alpha$ is $\varepsilon = \pm .0006$).

The agreement between theory and experiment exhibited in the last line is excellent. The errors quoted should be considered more as limits of error than probable errors. The error in the theoretical estimate could probably be reduced by a factor nearly 10 by a more detailed calculation of item 6, and, what might prove even harder, an estimate within ± 1 unit of the constant term " ? " in the expression for item 5. [It might also be possible to compute the total for one virtual photon to all orders in $Z\alpha$ exactly, as described in part (a) above.] It may be very hard to reduce the experimental error, however, for the position of a line is being measured to about one-thousandth part of its width.

What is the significance of this agreement, let us say, to ± 0.1 Mc in the hydrogen Lamb shift ? Again, an evaluation depends on how you expect Q.E.D. to fail. If it is expected that the failure appears only at high momentum transfer, say in the photon propagator, very little is checked here that is not already involved in the electron-proton scattering experiments. In the hydrogen atom the electron is successively scattered by the proton and if this, at very short distances, is not the ideal point charge scattering it can be corrected by using the directly measured scattering, without regard for the reason for the difference from the ideal scattering. For example, in the very unlikely event that Q.E.D. and proton structure effects are compensating each other in the proton scattering experiments, they will compensate here too; the correct net effect being still given in item 9.

On the other hand, one might contemplate a failure involving a modification of the propagator extending out to very large distances (compared to 10^{-13} cm), but having a very small coefficient. For example, suppose it is suggested that Coulomb's law is altered so that the potential from a charge can be approximated by the form $1/r^{(1+\varepsilon)}$ in the range of r of order 10^{-8} cm. We can conclude from the Lamb experiment that ε is less than 10^{-10}. This is because

the close coincidence of the $2s_{1/2}$ and $2p_{1/2}$ levels, each of which has an energy of the order of 10^9 Mc, is a kind of accident involving the perfection of the Coulomb law. The Lamb experiment tells us that any modification that perturbs the $2s$ energy level (in a substantially different manner than it disturbs the $2p$ level) must disturb it by less than one part in 10^{10}.

But perhaps the most satisfying aspect of the agreement of theory and experiment here is that it checks the general theoretical viewpoint. There cannot be much argument that the effects which we ascribe to virtual photons or virtual pairs acting in various orders do exist (although the philosophical ideas used to describe them may someday be altered drastically, of course).

The analogous Lamb shift in other states, such as $3s_{1/2} - 3p_{1/2}$, etc. have also been measured [7] and agree in a satisfying way with theory, although the test here is not quite as stringent as for the $2s_{1/2} - 2p_{1/2}$ because of the somewhat larger experimental error.

The $2s_{1/2} - 2p_{3/2}$ separation is also measured so, as a byproduct, we have a measurement [20] of the fine structure separation $2p_{3/2} - 2p_{1/2}$ in deuterium. The formula for this separation is [20] [18] [12]

$$\Delta = \frac{1}{16}\alpha^2 \,\mathrm{Ryd}_\infty \, c \left(\frac{\mu}{m}\right)^3 \left[(g_{el}\frac{m}{\mu} - 1) + \frac{5}{8}\alpha^2 - 2\frac{\alpha^3}{\pi}\,(ln(1/\alpha) + ?)\right] .$$

This formula can be derived by the usual complete Q.E.D. analysis, but its terms are easy to understand. The first is the energy of interaction of the electron's magnetic moment with the nucleus revolving about it, considering the electron at rest. It involves the anomalous magnetic moment of the electron, calculated [13] to be

$$g_{el} = 2\,[1 + \alpha/2\pi - 0.328\,(\alpha/\pi)^2] = 2\,(1.00115961) \qquad (4)$$

and a factor $m/\mu = 1 + m/\mathrm{M}$, where M is the deuteron mass, to correctly represent the velocity, relative to the electron, of the nucleus generating the magnetic field. The factor in front of the [ ] is obtained by averaging the interaction over the Schrodinger wave function for the p state. The next term, -1, is the Thomas precession correction. The relativistic correction to these two terms combined is to order α^2 just $\frac{5}{8}\alpha^2$, in accordance with the fine structure formula of the Dirac theory of hydrogen. Because the electron can emit and absorb virtual photons its effective location is smeared

over a range (the main source of the Lamb shift of the $2s$ state), so the fine structure interaction is corrected in order α^3. The biggest part of this is a term in $ln\ \alpha$ whose coefficient is easily understood, but a more complete evaluation of the α^3 term, up to the constant " ? ", has not been carried out. Since the $\alpha^3\ ln\ \alpha$ term only amounts to one ppm (part per million) and the experimental result

$$\Delta = 10971.58 \pm .20\ \mathrm{M}c$$

is available only to 20 ppm (limit of error) there is no great need to evaluate the α^3 term more completely than at present. Aside from the small terms $\alpha^3\ ln\ \alpha$, if an experimental value for g_{el} is used, this formula contains no subtle virtual state effects of Q.E.D. that cannot be understood from the Dirac equation and semi-classical arguments. It has been used to obtain a value for α, or as one of the equations in a general evaluation of the fundamental constants.

Anomalous magnetic moment of electron and muon.

From the emission of virtual photons the predicted value 2 for the gyromagnetic ratio of the electron is altered to the expression (4) valid to order α^2. The first term is from one virtual photon. The second term contains two effects : (*a*) the effect of two virtual photons, and (*b*) the vacuum polarization correction to the propagator of the first virtual photon. They have been calculated by Peterman. For the muon, supposing it to satisfy the Dirac equation but with a different mass, all the terms are, of course, the same except term (*b*). In the vacuum polarization for the muon there are terms for virtual electron pairs as well as muon pairs, and the predicted [21] g value for the muon is

$$g = 2\,[1 + \alpha/2\pi + 0.75\,(\alpha/\pi)^2] = 2\,(1.001165) \qquad (5)$$

A value for g_{el} has been obtained by Hardy and Purcell by combining a measurement of Gardner and Purcell (see reference [24]) of the cyclotron frequency of the electron to a magnetic moment measurement of Beringer and Heald [22] to get

$$g_{el} = 2\,(1.0011552 \pm 8)\,.$$

However, an independent measurement of the cyclotron frequency by Franken and Liebes [23] gave somewhat different results, and leads to a g value of

$$g_{el} = 2\,(1.001168 \pm 7)\,.$$

A direct measurement of Schupp, Pidd and Crane [24] gives

$$g_{el} = 2\,(1.0011609 \pm 24)\,.$$

These results are in fair agreement with the theoretical result (4). The result of Schupp *et al.* implies that the coefficient of the α^2/π^2 is -0.1 ± 0.4.

The theoretical result should probably be accurate to one part in about 10^{-8} guessing that the next term in the series is roughly $\pm (\alpha/\pi)^3$. If the photon propagator is modified as (3), the correction to $g/2$ is $-\frac{\alpha}{3\pi}\left(\frac{m}{\Lambda}\right)^2$ so an agreement to 1 ppm means only that Λ exceeds 15 Mev. More information comes from the measurement of $g_\mu - 2$ for the muon. The experimental result [25] for g_μ is $2\,(1.001145 \pm 22)$ agreeing with the theory within its error of 22 ppm. In this case a propagator like (3) would correct $g/2$ by $-\frac{\alpha}{3\pi}\left(\frac{m\mu}{\Lambda}\right)^2$ where m_μ is the meson mass of 105 Mev. If this is not to exceed 22 ppm, Λ must exceed 630 Mev. This is therefore at least as good a test as is provided by the proton-scattering experiments. It is remarkable in that it tests at the same time that the heretofore unfamiliar particle, muon, satisfies the Dirac equation with no appreciable structure comparable to its own Compton wave length.

Hyperfine interaction.

The hyperfine splitting in the ground state of hydrogen resulting from interaction of the nuclear moment and the electron has been measured [26] very accurately for the three isotopes of H, and for the He^3 ion. The theoretical formula [27] for this splitting is

$$\Delta\nu = \frac{16\alpha^2 c}{3}\,\mathrm{Ryd}_\infty\left(\frac{\mu_p\,\mu_{el}}{\mu_o^2}\right)\left(\frac{\mu}{m}\right)^3\left[1 + \frac{3}{2}(Z\alpha)^2 - (\frac{5}{2} - ln2)Z\alpha^2 - X\frac{\alpha m}{M}\right] \tag{6}$$

where μ_p, μ_{el} are the magnetic moments of proton and electron and μ_o is the Bohr magneton, $\mu/m = \left(1 + \frac{m}{M}\right)^{-1}$ where M is the mass of the nucleus. The terms in front of the bracket gives the value expect from a non-relativistic analysis, given by Fermi. In the bracket the $\frac{3}{2}(Z\alpha)^2$ is a Breit correction resulting from the use of the Dirac equation instead of the Schrodinger equation. The next term is a correction from virtual photons in Q.E.D.

The last term is a correction for recoil and finite size of the

nucleus. A direct calculation assuming a point charge and dipole, leads to a result which diverges logarithmically [28]. It is sensitive to the electromagnetic structure of the nucleus. It arises from the exchange of *two* virtual photons with the proton (in H). If the amplitude for this exchange at high energy were known, the term could be evaluated. Data could come from the forward spin flip Compton scattering from a proton. It cannot come directly from proton-electron scattering experiments for these only give the form factor for a one photon interaction with the proton. Assuming that each of the two interactions has the same form factor as does a single photon, C. Iddings and P. Platzmann, *P. R.*, **113**, 192 (1959) evaluated X as 8.7, corresponding to a correction of — 35 ppm. However, I think uncertainties in the assumption relating one and two photon structure factors, as well as uncertainties of the one photon structure factor itself for high energy may make the error in X as high as — 2 or + 10 ppm. [There is a relative insensitivity to these assumptions because a major part of X involves the large $ln(M/m)$.]

Unfortunately, therefore, we cannot use these high precision measurements directly to test Q.E.D. independent of our uncertainties in the electrical properties of the nucleus *. Nevertheless, there may be a discrepancy here for the measurement of Lambe and Dicke [19] corrected by 27.5 ppm for a diamagnetic correction gives, with $1/\alpha = 137.0389 \pm .0006$, a result for the X term of $+ 0.7 \pm 8.8$ ppm instead of — 35 ppm, but terms in (6) of order α^3 may be unusually large **.

---

* The hyperfine splitting for deuterium is still more dependent upon the nuclear structure, this time of the deuteron. The ratio υ_D/υ_H divided by the ratio of the magnetic moment of D and H and by the reduced mass factors cubed is one minus 1.703×10^{-4} experimentally. An attempt to estimate this from nuclear theory by Low and Salpeter, *P. R.*, **83,** 478 (1951) gave $1.98 \pm .20 \times 10^{-4}$, but this calculation could probably be improved today.

** NOTE ADDED IN PROOF :

D.E. Zwanziger has just informed me of calculations of corrections of order $\alpha^3(ln\alpha)^2$ and $\alpha^3 ln\alpha$ to the hyperfine separation in hydrogen made by him and A.J. Layzer independently. They find a contribution of — 9 ppm so the total predicted term is — 44 ppm. Thus, a real discrepency to the " measured " + 0.7 ± 8.8 ppm seems to be developing here. However, the trouble might lie instead in the measurement of the fine structure separation in deuterium, for E. Richard Cohen has kindly informed me that if this fine-structure measurement is omitted from a least-square reduction of the fundamental constants the value of $1/\alpha$ is $137.0417 \pm .0025$. The " measured " value of the hyperfine separation term would then be $— 40 \pm 36$ ppm (instead of + 0.7) which would be consonant with the theoretical value.

Zwanziger has noted that these uncertainties disappear if one compares, in the same atom, the hyperfine structure in the $2s$ state to that in the $1s$ state. Measurements of the hyperfine shift in the $2s$ state have been made by Heberle, Reich and Kusch, *P. R.*, **101**, 612 (1956), **104**, 1585 (1956). According to the Fermi formula the ratio should, in first approximation, be as the probability of finding the electron at the nucleus, (1/8), but the actual measurements lead to a result

$$8\,\nu_{2s}/\nu_{1s} = 1 + d \text{ where } d = 34.6 \pm 0.3 \text{ ppm for H} \\ = 34.2 \pm 0.6 \text{ ppm for D} \qquad (7)$$

The formula (6) is not adequate to calculate d to this accuracy, for the expression in brackets has not been carried to a high enough order. It might be expected that if divergences arise already in this order they should be still worse for higher order, but Zwanziger shows that this is not true if the ratio ν_{2s}/ν_{1s} is calculated. The term d has several contributions. The electron magnetic moment is spread by vacuum polarization, and by the form factor in the theory of this moment. In addtion the interaction of the electron with the nucleus is altered because the wave function of the electron is altered. This is because, just as in the usual Lamb shift, the electron sees a modified potential since it emits and readsorbs photons [giving a term in $ln(m/\text{Ryd})$] and further, the potential is actually modified by the vacuum polarization. This calculation is similar to the first order Lamb calculation. Terms of order $\alpha^2 m/M$ have been calculated by Schwartz (see reference [27]). The theoretical result for d is

$$d = 34.5 \pm 0.2 \text{ ppm}$$

which agrees excellently with the experiments (7). The Breit term alone gives $\frac{5}{8}\alpha^2$ or 33.3 ppm so that we do not here have any sharp test of Q.E.D. at short distances. But it does confirm our general ideas and checks again, but less accurately, that there are no small deviations at larger distances.

There is a measurement of the hyperfine structure of the metastable triplet state of He^3 by White *et al.*, *P.R.L.*, **3**, 428. If this is compared to the He^3 ion hyperfine separation, the dependence on nuclear structure cancels out. For the ratio of frequencies they get 6.2211384 ± 12. To calculate this it is necessary to know the wave function for the triplet state to find the probability (times 6)

that one electron is at the nucleus. The non-relativistic theory with the best known wave function gives 6.222030 for this. Relativistic corrections reduce this to the theoretical prediction 6.221199 $\pm$ 6 for the frequency ratio. The remaining deviation is likely to be a slight inadequacy (only 10 ppm) of the variational wave functions. There is no evidence that Q.E.D. is any less adequate in handling systems with two electrons than it is for one.

Positronium.

The " atom " formed from an electron and a positron presents none of the uncertainties of nucleon structure that the hydrogen atom does. It is therefore an interesting object for calculation, although the experiments are much more difficult because of its transient nature. A still better object would be muonium but the experiments here have not yet been performed.

Just after the newer methods of Q.E.D. were developed, since they were so readily applied to perturbation theory of free systems, it was supposed by many that bound state problems presented some special difficulty. That this was not so was noted by W.E. Lamb, *P. R.*, **85**, 259 (1952) and E.E. Salpeter, **87**, 328 (1952). Evidently one cannot analyze the bound state by starting a perturbation series from non-interacting particles. But one very effective way is to use, as a starting point, the system held together by instantaneous Coulomb potentials. This system can be analyzed by an ordinary differential equation in time, like the Schrodinger equation, because of the instantaneous nature of the interaction. The perturbation then consists of adding the effect of virtual transverse photons in various orders. Of course, any other unperturbed system, held together by some approximation to the true interaction, will serve as well; the perturbation being the difference between the true interaction of Q.E.D. and the approximate interaction assumed. The instantaneous Coulomb potential is a good starting point because its initial approximation is so good.

The most complete analysis of the hydrogen-like atom with arbitrary mass ratio of the two charges has been given by T. Fulton and P. Martin [18], where references to earlier work will be found. They have used their equations to compute the energies up to the first order Lamb effect (i.e., to order α^3 Ryd) of many states in

positronium. The most delicate test is the separation between the singlet and triplet $1s$ states of positronium measured by Weinstein, Deutsch and Brown, *P. R.*, **98**, 223 (1955) to be $2.0338 \pm 4 \times 10^5$ Mc.

The theoretical value [29] for this shift is

$$\alpha^2 \,\mathrm{Ryd}_\infty \, c \left[\frac{2}{3} + \frac{1}{2} - \left(\frac{16}{9} + ln2\right)\frac{\alpha}{\pi}\right] = 2.0337 \times 10^5 \,\mathrm{M}c$$

the first term represents the first order interaction of the spins and the second the amplitude for virtual annihilation into a photon with re-creation of the positron again. It is clear experimentally that such a term exists, once again confirming our general view of what virtual processes go on. The last term is the first order Lamb correction for this system; it amounts to 0.0100 Mc.

General conclusions.

There are many other calculations and experiments in which some aspect of Q.E.D. is involved (such as vacuum polarization effects in mu-mesic atoms, or relativistic corrections to the computed helium ground state energy, etc.). We shall not go on to describe them for, although confirming Q.E.D., they do not provide sharper tests than the examples already given.

All this may be summarized by saying that no error in the predictions of quantum electrodynamics has yet been found. The contributions expected from the various virtual processes envisaged have been found again and again, and there is very little doubt that in the low energy region, at least, our methods of calculation seem adequate today. The region of energy (of virtual states) that has not yet been explored even for gross errors exceeds 600 Mev (the Compton wave length corresponding to this is 2π times 3×10^{-14} cm). There are no experimental indications that the laws of Q.E.D. cannot be exact. Are there any theoretical reasons to expect a failure ? I will discuss such questions in the next few sections.

Before we do this I should like to make a remark on the character of these calculations. It seems that very little physical intuition has yet been developed in this subject. In nearly every case we are reduced to computing exactly the coefficient of some specific term. We have no way to get a general idea of the result to be expected.

To make my view clearer, consider, for example, the anomalous electron moment given in (4). We have no physical picture by which we can easily see that the correction is roughly $\alpha/2\pi$, in fact, we do not even know why the sign is positive (other than by computing it). In another field we would not be content with the calculation of the second order term to three significant figures without enough understanding to get a rational estimate of the order of magnitude of the third. We have been computing terms like a blind man exploring a new room, but soon we must develop some concept of this room as a whole, and to have some general idea of what is contained in it. As a specific challenge, is there any method of computing the anamalous moment of the electron which, on first rough approximation, gives a fair approximation to the α term and a crude one to α^2; and when improved, increases the accuracy of the α^2 term, yielding a rough estimate to α^3 and beyond ?

THEORETICAL QUESTIONS

Self-energy.

The first difficulty which arises if one assumes that each of the transverse electromagnetic modes in a box is a quantized oscillator is that each should have a zero point energy $\omega/2$. Since there are an infinite number of modes this zero point energy is infinite. It is easy to get around this difficulty, however, by supposing that absolute energy cannot be measured (leaving a question for the theory of gravitation) so that all the zero point energy is subtracted. But now suppose we put atoms or other objects in the box at a small density N per unit volume, so that the index of refraction is changed from 1 to $1 + 2\pi \mathrm{N} f_{\mathrm{K}}^{\gamma}/\omega^2$ where f_{K}^{γ} is the real part of the forward scattering amplitude of the object for light of mode K. The wave lengths which fit into the box are still the same, but the frequency of the modes is changed by $2\pi \mathrm{N} f_{\mathrm{K}}^{\gamma}/\omega$ so the total zero point energy is changed, per object, by

$$\Delta \mathrm{E} = \frac{1}{2} \Sigma_{\mathrm{K}} \frac{4\pi f_{\mathrm{K}}^{\gamma}}{2\omega_{\mathrm{K}}} .$$

This shift in energy we would associate with the object and would call it the self-energy of the object. There are higher terms from

the effect of scattering two photons at once, but we shall just go to the first order in α. In addition, we have the positron-electron Dirac field and will have to add a contribution for the shift in energy of the electrons in the negative energy sea. This will involve f_{N}^{pos} the amplitude for the object to scatter forward a positron in state N. It is better to deal with electrons and positrons symmetrically and we find the following formula for the self-energy of an object :

$$(4\pi)^{-1}\Delta \mathrm{E} = \frac{1}{2}\Sigma_{\mathrm{K}} \frac{f_{\mathrm{K}}^{\gamma}}{2\omega_{\mathrm{K}}} + \frac{1}{2}\Sigma_n \frac{f_n^{el}}{2\,\mathrm{E}_n} - \frac{1}{2}\Sigma^{\mathrm{N}} \frac{f_{\mathrm{N}}^{pos}}{2\,\mathrm{E}_{\mathrm{N}}} \qquad (8)$$

where f_n^{el} is the real part of the amplitude for the object to scatter forward an electron in state n, f_{N}^{pos} that for scattering a positron in state N, and f_{K}^{γ} that for scattering a photon in state K (K specifying momentum and polarization), everything to first order in α *.

Applied to a free electron, however, (8) still gives a divergent result. It might be thought that this could also be subtracted away, and only differences taken for the electron in different states i, but these differences are also infinite. This is most easily seen if one compares the energy of two photons with the energy of the pair one expects to create from them. No completely satisfactory way has been found out of these difficulties.

Electron mass.

In making actual calculations, one way to handle the difficulty is this : temporarily stop the divergent integrals at some high energy and note that the self-energy effect, at least insofar as it depends (logarithmically) on the cut-off is equivalent to changing the mass of the electron from m_o to $m_o + \Delta m$ in every process (of energy well below the cut-off). If we write $m = m_o + \Delta m$, interpret it as the experimentally observed mass, and write all results in

* To be more explicit if the object is an H atom with an electron in state i this formula gives the correct level shift to order α if n and N are states in the nuclear potential Z, but the interaction of the electron in i and the electron n is calculated only to first order in α. That this is true can be easily demonstrated by writing out each amplitude by diagrams, adding the results, and comparing with the unusual diagram for the virtual photon level shift. Actually, only the exchange scattering in f^{el} and the annihilation scattering in f^{pos} need be taken. The rest cancels out. If the object is charged, like a free electron, there is no true forward scattering f^{el} as it goes, by the Rutherford law, as θ^{-2} but f^{pos} does likewise and, in (8), they cancel out to a finite limit as $\theta \to 0$.

terms of m, then the limit can be taken as the cut-off energy goes to infinity. Expressing everything in terms of m is called renormalizing the mass, and a theory for which results are then independent of the cut-off as it goes to infinity are called renormalizable. Q.E.D. is renormalizable if the electron mass, and the charge, are both renormalized (it is, I think, not necessary to renormalize the mass ratio of muon and electron).

What is the meaning of this ability to renormalize the mass ? From one point of view it is no problem at all. After all, only m is observable, not m_o so the m_o is a construct which should be got rid of anyhow. Only the renormalized theory should have been written in the first place. Two questions arise, however.

The first question is whether the renormalized theory is, in fact, a logically consistent theory. With any finite cut-off the theory is not consistent, slight deviations from unitarity (the principle that the sum of probabilities for all alternatives should be unity) occur. These get smaller as the cut-off energy Λ goes to infinity. But the mass m_o that may be needed to get a finite m for very large Λ may be negative. That is, the theory may contain hidden difficulties if we computed processes for energies E such that $\alpha\, ln(E/m) \gtrsim 1$. This is such an extreme energy that such matters are of no apparent concern to calculation of lower energy phenomena. However, from a strictly theoretical view it would be nice to know whether renormalized Q.E.D. is a consistent theory, or whether difficulties may not arise of relative magnitude e^{-137}. The great difficulty in answering such questions is our limited mathematical ability to deal with situations where some kind of perturbation theory does not suffice. I do not know if it has even ever been proved that Δm still diverges if all orders of perturbation theory are included. [The perturbation result for Δm is $m_o\,(3\alpha/2\pi) ln\, \Lambda/m_o$].

The second question concerning the renormalization idea is that renormalization of a quantity Λ gives up any possibility of calculating that quantity. Now it may be that the " electromagnetic part of the electron mass " is unobservable, but this is not true of other particles. The difference in mass of proton and neutron, or of π^+ and π^o, or of K^+ and K^o, etc. are almost certainly electromagnetic in origin. They cannot be computed with a renormalized theory, for in such a theory any constant can be added to the masses of

each particle. It should be objected that these baryons and mesons are complicated systems, in virtue of the strong interactions, and we therefore do not know the correct laws of coupling to calculate this mass difference. If we knew them, perhaps the mass differences would converge without any modification of Q.E.D. itself. This is indeed possible, in principle, but so far it has not been demonstrated to be true in fact. Any complete field theory yet written for strongly interacting particles and electromagnetic interactions, has always appeared to be unrenormalizable unless these mass differences are renormalized as well. It seems odd to try to put the complete burden of the divergences of Q.E.D. on special properties of the strong couplings, but on the other hand, that is where they may indeed lie. At any rate, close study of these mass differences would probably teach something; either of the breakdown of Q.E.D. or of the electromagnetic characteristics of the particles.

If it is assumed that the nucleon and meson structure is not solely responsible for the convergence of the mass differences, and that it is a failure of Q.E.D. instead, then the numerical values of these differences suggest that the failure of Q.E.D. should begin to show up strongly at virtual energies around 1 Gev.

Finally, we should remark on the possibility that all of the mass of the electron is electromagnetic in origin. First, of course, the correction seems to be too small to do that (yet how can something that is infinite appear to be too small ?). But disregarding that, there is the argument that, if the electron mass is zero, the change of the electron field operator from ψ to $\gamma_5\psi$ will not alter things. Put otherwise, a state of right helicity will never be converted to one of left helicity, no matter how often it interacts with real or virtual photons. But an electron with mass does not have this property, so it has been believed that mass cannot come from no mass. But recently several physicists (Heisenberg, Nambu, Schwinger, for example) have argued that this is, in fact, not true. In a ferromagnet the original system has the symmetry that all space directions are equal; but in fact the interaction can produce a polarized background state in which any excited state has an energy depending on its alignment to an axis. That is, if we go beyond perturbation theory such symmetry arguments may fail.

On the other hand, pure Q.E.D. with only zero mass electrons

and photons interacting with no other particles and having no cut-off energy probably cannot produce a finite electron mass. This is because the system is also invariant to a change of scale, there is no parameter to determine a length. Yet an electron with a mass involves such a length. I am not certain, but it appears to me impossible to generate a specific length from no scale whatsoever.

However, in fact, there are two places in Nature from which such a length could come. One is the theory of gravitation, the gravitational constant involves length dimensions and light interacts with gravitons. Further, Machs' principle in quantum mechanics is equivalent to the statement that the surrounding nebulae determine the atomic scale of length in a local vicinity. However, these theories are not developed far enough for us to compute the electron mass starting only with zero mass electron, photons and gravitons in interaction.

A more practical point to notice is that, in fact, photons do interact with nucleons and mesons, and to allow that there is, in that system, an independent scale of length, say the nucleon mass. This would serve as the small length-determining perturbation, which works its way back to determine the mass of the electron. It is always possible that the equation determining the electron mass has more than one solution, and that a second solution is the muon, but we are engaged in pure speculation here.

One way to find out about this is to study more seriously Q.E.D. with electron mass exactly zero. On the one hand, it may be useful for getting a better understanding of Q.E.D. when electron energies are high, but most particularly, it would be interesting to see if such a theory is consistent at all. For example, a charge entering a magnetic field presents problems; it seems to radiate at an infinite rate. Perhaps a careful study of this problem, and the effect of a small length-determining perturbation on the result, would lead us to an understanding of the ratio of electron mass to nucleon mass.

Electric charge.

Q.E.D. contains two constants which must be determined by experiment. One is the electron mass, which we have just discussed. The second is the electric charge, or the dimensionless combination $a = \hbar c/e^2 = 137.039$. For both of these the renormalization process must be applied, so we have foregone computing a also.

It is interesting that all the „ fundamental „ particles which are charged have the same charge, but we have no remarks to make on that point.

The way that the charge becomes renormalized is via the polarization of the vacuum. A virtual pair produced by a photon (also virtual) annihilates again to re-create the photon. In first approximation the correction to a from virtual electrons is $\Delta a = (2/3\pi)\, ln(\lambda/m)$. That is to say, it is divergent, so we have this time to cut-off the electron propagator at some energy λ (this is in addition and different than the photon propagator cut-off discussed above — but they may ultimately have the same origin). Virtual pairs of other particles such as muons simply add their contribution to Δa in first order.

At first it may be argued that here, at least, the philosophy of renormalization is unassailable. The " free charge " must be unobservable (although Gell-Mann has suggested that this may not be so, but high energy interactions may determine it). Only the total corrected a can be measured. Unlike the case of mass where we have charged and uncharged particles, like neutron and proton, to compare, here we have nothing but the single measured a.

This is true, but are you willing to give up, forever, the possibility of computing this remarkable constant a ? If in some ultimate future theory a is to be computed, will we find ourselves correcting some value a_o for virtual pairs ? One can hardly begin to speculate from our present position, but the question of how a can be computed at all has always been intriguing, and I should like to make a few remarks about it.

The quantum theory of electromagnetic interaction can be formulated roughly as follows. Let S_o be the S-matrix operator for the system of particles not interacting with the electromagnetic field, let j_μ be the current density operator of this matter omitting the charge factor, and A_μ be the electromagnetic potential operator times the charge. Then the S-matrix including the field is something like (I am merely outlining here, the precise definitions are assumed to be familiar)

$$S = S_o\, e^{i \int j_\mu A_\mu d\tau}\, e^{i L_o(A)} \tag{9}$$

where

$$L_o(A) = \frac{a_o}{8\pi} \int F_{\mu\nu} F_{\mu\nu}\, d\tau \tag{10}$$

$a_o = \hbar c/e_o^2$ is the unrenormalized a and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Note that we are using somewhat unfamiliar notation since neither j_μ nor A_μ require knowledge of the charge in their definition, it appears only in L_o.

Now if we take expectation values between states of the vacuum for matter (to get at the vacuum polarization and other effects of virtual pairs) we shall have to calculate the expectation

$$<S_o\, e^{i\int j_\mu A_\mu d\tau}> = e^{i\,L(A)} \tag{11}$$

it being sufficient to consider A_μ as a c-number function and L(A) defined here is simply a functional of A_μ. The completion of a problem requires that we integrate

$$\int e^{i\,(L(A) + L_o(A))}\, DA \tag{12}$$

over all potential functions A_μ satisfying the boundary conditions of the problems.

In a way of speaking, then, if we were not aware of the pairs we would say that electrodynamics has the effective Lagrangian $L(A) + L_o(A)$.

The L(A) defined in (11) can be expanded in powers of A_μ and in powers of its derivatives assuming in some situation that A is smoothly varying and small. A constant A_μ has no effect, assuming $\partial_\mu j_\mu = 0$, the conservation of charge, or guage invariance, so the expansion begins :

$$L(A) = c\int F_{\mu\nu}\,F_{\mu\nu}\,d\tau + c'\int (F_{\mu\nu,\,\sigma})^2\,d\tau + c''\int (F_{\mu\nu})^4\,d\tau + \ldots \tag{13}$$

where, except for the first term, only the type of term is meant to be indicated, two F's and two derivatives in the second term, four F's in the third, etc. The c, c', c'' ... are constants. The first term can be combined with L_o in (12) to form $\frac{a_1}{8\pi}\int F_{\mu\nu}\,F_{\mu\nu}\,d\tau$ with $a_1 = a_o + 8\pi c$ and this is the origin of part [34] of the charge renormalization. For electron pairs $c = (1/12\pi^2)\,ln(\lambda/m)$ (as we have said), which is very small (although infinite !) compared to a.

The other terms in the effective Lagrangian $L_o + L$ generate modifications of Coulomb's law, the scattering of light by light, etc. Had these phenomena been discovered before Q.E.D., they would have been representable in classical theory as just such

modifications of Maxwell's Lagrangian, so that the complete Lagrangian would be considered to be a complicated, non-linear and partly unknown affair. With the advent of Q.E.D. explaining the origin of the non-linearities, etc. the natural reaction would be *to try to explain the complete Lagrangian* in this way. The fact that c is infinite, although discouragingly small, would be exciting. It is, in fact, possible that c is not so small; perhaps the cut-off is controlled by gravitation, or, perhaps, some other modification of the laws of matter replaces the logarithm with a less convergent expression, or the cut-off may not be necessary if perturbation theory could be avoided. Finally, every charged particle makes a nearly equal contribution (although not to c' or c'' which vary inversely as the mass squared, or fourth power). So we must add contributions from muons, nucleons and mesons. We do not know how many nor how much each contributes and, although difficult, it is not impossible that a complete future calculation would give a value near 137 to $8\pi c$ so that no a_0 is required at all !

If that is the case, quantum electrodynamics takes a very simple form. In (9) $L_0(A)$ must be omitted. Then the functional integral on A_μ, which must be taken, gives, since $\int \exp.(i\int j_\mu A_\mu d\tau) DA_\mu = \delta[j_\mu]$, a functional delta function of j_μ. It says therefore simply that all amplitude and *expectations* of S_0 *must be taken subject to the condition that the total current density is everywhere and always zero.*

What is electromagnetic interaction ? If a real charge exists it must generate its own four-current (charge density, if at rest). Since total current vanishes this must be compensated by an opposite current in the vacuum sea of charged particles. Because of the dynamics of these particles the compensation cannot occur just locally, but another counter current is generated nearby, compensated by vacuum current again, etc. until this effect is propagated out to infinity. The energy associated with these compensating currents is the electromagnetic self-energy of the charge. Another charge placed in the vicinity adds its system of compensating currents so there is an energy of interaction between these charges. At any rate something like that is implied by these unsupported speculations.

Interaction of other particles.

Although Q.E.D. is very accurate, there is of course one way in which it *must* be wrong; it is incomplete. There are not only

photons, electrons, and muons in the world, but charged baryons and mesons as well. It will not do to say that Q.E.D. is exactly right as it stands in a limited situation where only electrons and muons are present, because virtual states of charged baryons must have an influence. Two sufficiently energetic photons, colliding, will not do just what Q.E.D. in that limited sense supposes; they also produce pions. Or, more subtly, a sufficiently accurate analysis of the energy levels of positronium would fail, for the vacuum polarization from mesons and nucleons would have been omitted.

On the other hand, we use that theory as best we can to discuss the Coulomb potential from the nucleus, the nuclear emission of γ-rays, the Bremstrahlung expected from pions, the chance that γ-rays will be found in the disintegration of the K^+, etc. How do we do it ?

We expect that the principles of Q.E.D. will extend to these particles too, and replace our incomplete knowledge of them by a set of constants (charge, magnetic moment, etc.) which will suffice for low energy analysis.

Yet we use the tool of electrodynamics for much more than that. We make some hypothesis about how the photons are " ultimately " coupled to the new particles. We suppose for example, that the anamolous magnetic moment of the proton has its ultimate origin, not in an extra Pauli term $\frac{\mu}{4M} \sigma_{\mu\nu} F_{\mu\nu}$ in the " original Lagrangian " but in the currents of virtual mesons which surround the nucleon. That is, we assume that the coupling is in some sense ultimately as simple as possible, and all apparent anamolous effects have their origin in complexities of the strongly interacting particles themselves. This hypothesis is universally used, permitting us to use electromagnetic interaction to learn something about the strange particles. Yet it has never been formulated in a completely precise manner. Its importance was first emphasized by Gell-Mann who called it the principle of minimal electromagnetic interaction, but following a suggestion of Telegdi, I shall call it Amperes-hypothesis (the assumption that all magnetism comes from currents).

There is one exception already known to this principle; photons interact with the gravitational field without a charged intermediary. But this interaction can be viewed the other way about, that gravity

interacts with all energy, photon energy included. We shall therefore disregard this counter-example for the present.

The simplest suggestion for defining Amperes-hypothesis would be to say that in the " fundamental Lagrangian " (the exact form of which, at present, unknown) all gradient operators ∂_μ on charged fields are to be replaced by $\partial_\mu - A_\mu$ and no other coupling to A_μ is to be assumed. There are two objections to this formulation. First, we do not know the form of the future theory; no Lagrangian may exist. Is there not some formulation closer to the observed properties of the particles ? The second objection is possibly academic; if the Lagrangian were before us, we would probably know exactly what to do. Yet a term like $\sigma_{\mu\nu}\partial_\mu\partial_\nu$ which is evidently zero could be added, but when ∂_μ is changed to $\partial_\mu - A_\mu$ it is no longer zero but is a Pauli term $\sigma_{\mu\nu} F_{\mu\nu}$ instead. Such an ambiguity would arise, say, if the Lagrangian contained second derivatives, and it was not clear whether to write $(\gamma_\mu\partial_\mu)^2$ or $\partial_\mu\partial_\mu$ at some point.

One of the effects of terms like a Pauli moment is that certain processes in Q.E.D. become uncalculable. For example, if the muon carried a true Pauli moment the hyperfine split of muonium could not be computed, as the term " X " in (6) is divergent in that case. In other words, the theory would not be renormalizable. Perhaps then Amperes-hypothesis is equivalent to the assumption that the theory is renormalizable. On the other hand, the proton-neutron mass difference may not be one of those computable quantities. We must understand renormalizability, then, as the hypothesis that only a finite number of unknown quantities must be attached, before everything else can be computed.

Another effect of a term like a Pauli moment is to drastically alter the behavior of cross-sections at high energy. From the dispersion point of view, discussed below, constants such as the charge, anamalous moments, etc. appear in the form of subtraction constants required to ensure the convergence of the dispersion integrals involved. More constants are needed if the high energy cross-sections remain large, or increase, with rising energy. Thus Amperes-hypothesis in this viewpoint would take the form of a statement that a certain minimum number of subtraction constants are required.

What amounts to the same thing, but is more readily available to experiment is to try to replace Amperes-hypothesis by a statement

that the size of high energy photon cross-sections are limited in some specific way.

As a last remark, if the current j_μ comes from the Lagrangian so easily, might the structure of this current not tell us something about that Lagrangian (or whatever more fundamentally replaces it) ? This possibility is discussed at this conference by Gell-Mann.

Dispersion theory [31].

In the lowest order in which any process occurs, there are no integrals over virtual states (closed loops), and no divergences will arise. The Q.E.D. difficulties arise on integrating over invisible virtual processes. This generates a feeling that such virtual state integrations are unreal, or at least are not handled quite correctly, and that all of the formulas should be put in terms of directly measurable quantities. This can be done because of the analytic character of the functions involved. Their real parts can be expressed in terms of their imaginary parts. The imaginary parts can be expressed as the rate for real processes of lower order. Thus a diagram involving one virtual momentum integral can be expressed as a dispersion integral over a function determined from processes without a virtual integral at all. Put in this way, it sounds trivial, one integral is replaced by another, almost identical, and in fact from a practical calculational point of view there is often little to gain. But it is hoped that this viewpoint gives a clearer insight into the virtual state integrals and a closer relation to experiment. (The greatest utility of this method results, of course, in the analysis of strong coupling where the fundamental equations are unknown, for here one experimental result can be related to another.)

We can illustrate by the simplest example, the second order vacuum polarization effect of electrons. The amplitude that a virtual photon of momentum q^2 makes a pair and annihilates again is written $q^2 f(q^2)$, so that the entire dependence of L(A) on A expanded to second order is $\int f(q^2)\, F_{\mu\nu}(q)\, F_{\mu\nu}(q)\, d^4q$ written in momentum space.

Now the imaginary part of $q^2 f$ is the rate that a (virtual) photon makes real pairs. It only exists, writing $q^2 = 4m^2x$ for $x > 1$. Choose the time axis in the direction of q and the photon polariz-

ation in the z direction and we require the probability that a vector potential of frequency $2m\sqrt{x}$ constant in space, produces a pair of electrons each of energy $E = m\sqrt{x}$, of momentum $p = m\sqrt{x-1}$. This is one of the simplest elementary problems in Q.E.D. The amplitude for this is $(\bar{u}_2\gamma_z u_1)$, squared and summed on polarizations it is $-sp\,[(\bar{p}_2+m)\,\gamma_z(\bar{p}_1+m)\gamma_z] = -(p_2\cdot p_1 - m^2) - 2p_{2z}p_{1z}$ or averaging over directions, $E^2 + \frac{1}{3}p^2 + m^2 = \frac{2m^2}{3}(1+2x)$. The phase space factor is $pE/(2E)^2$ so, dividing by q^2, we find for the imaginary part of f (times $(4\pi\alpha)^{-1}$)

$$f_i = \frac{1+2x}{3x}\left(\frac{x-1}{x}\right)^{1/2}.$$

The real part is now given by a dispersion integral from Cauchy's theorem as

$$f_R(q^2) = \pi^{-1}\int f_i(q_1^2)\,dq_1^2/(q^2 - q_1^2)$$

or

$$f_R(x) = \frac{1}{\pi}\int_1^\infty \frac{1+2y}{3y}\left(\frac{y-1}{y}\right)^{1/2}\frac{dy}{x-y} \tag{14}$$

The integral, however, is divergent (so the Cauchy theorem is not strictly true, there is a contribution from the contour at infinity). This can be handled in the following way. We can assume f_R for some x is known experimentally or by definition. In this case $f_R(0)$ is the renormalization of the charge, so in the spirit of renormalization we can take it to be zero. Then we can use the Cauchy relation for the more convergent expression $[f(x) - f(0)]/x$. What it leads to is the same as if we subtract from (14) the same equation with $x = 0$:

$$f_R(x) = f_R(0) + x\int_1^\infty \frac{1+2y}{3y}\left(\frac{y-1}{y}\right)^{1/2}\frac{dy}{y(x-y)} \tag{15}$$

an integral which is now convergent and whose integrated value $(2/3x)\,[(2x+1)\,(1 - \beta\,ctn\,\beta) - x/3]$, where $\sin^2\beta = x$, is exactly the finite part of the vacuum polarization effect obtained by the usual integral over a closed loop [30].

It is clear that Q.E.D. in its renormalized form may have its simplest expression in this mathematical scheme. We need merely say that $f_R(0)$ vanishes, for we wish to work with the constants

already renormalized. Several authors have discussed Q.E.D. from this point of view [32].

If the integral (15) were still divergent we could perform another subtraction, but generate another uncalculable constant (analogous, in a different problem, to a Pauli anamalous moment). Amperes-hypothesis is that this is not necessary.

It must be admitted, however, that no fundamental change in position on the renormalization question is really involved. If the integral (14) *did* converge we certainly *would* compute it, and call its value at $x = 0$ the change Δa in a_o induced by pairs. In the usual loop integral method the integral is completed by subtracting the effect that the pairs would have if the mass of the electron were changed from m to λ. The same method here makes (14) convergent and gives the same value for $f_R(0) = \Delta a = (2/3\pi)\, ln(\lambda/m)$.

It might be hoped, therefore, that such dispersion relations involving the entire system of strange particles and Q.E.D. may be a satisfactory way of representing nature. It has much to recommend it; its close relation to experiment, the possibility of interdetermining coupling constants, the avoidance of the possibly meaningless question of which particles are fundamental and which compound, etc. These points have been emphasized by Chew in a remarkable speech at the conference in La Jolla, California this year. One serious question appears, however. Integrals over all energies are still required and at high energies the real processes involve all kinds of particles in considerable numbers. Thus the set of interconnected equations becomes enormously elaborate just when it becomes interesting. It is not clear how to get started grappling with this complexity.

On the other hand, the experimentally observed extreme energy phenomena suggest that they may have certain regularities. If this is so, a central theoretical problem is to formulate these regularities [33]. Only then may it be possible to close in an intelligent way the wide-open hierachy of dispersion relations.

It is in the spirit that all quantities should be reexpressed in terms of others, in principle, observable that the formula (8) for the Lamb shift self-energy was developed. Dispersion theory will probably also permit its being simplified still further. The real

part of the forward photon scattering cross-section f_K^{γ} can, of course, be immediately expressed as an integral on the imaginary part, and a similar reduction can likely be made for the other terms. If this can be done the calculation of the Lamb shift for H (to first order in α, all orders in $Z\alpha$) will only involve true rates of real photon absorption or of pair annihilation. Thus, for each state n a definite K value is associated (so $\omega_K = E_n - E_i$), so that the computation may be possible on machines. In fact, many of the matrix elements, useful in calculating internal conversion coefficients, etc. have been already calculated.

The formula (8) is not valid directly for calculating such things as the proton-neutron mass difference, because other virtual fields, like the meson field, must be included. The necessary generalizations of (8) can be written, but we shall have to see how useful they are and if the necessary experimental quantities are available.

Conclusion.

In writing this report on the present state of quantum electrodynamics, I have been converted from a long-held strong prejudice that it *must* fail significantly (other than by simply being incomplete) at around 1 Gev virtual energy. The origin of this feeling was the belief that the mass of the electron (relative to the nucleon, say) and its charge, must be ultimately computable and that Q.E.D. must play some part in this future analysis. I still hold this belief, and do not subscribe to the philosophy of renormalization. But I now realize that there is much to be said for considering theoretically the possibility that Q.E.D. is exact, although incomplete. This assumption may be wrong, but it is precise and definite, and suggests many things to study theoretically, while the other negative assumption, (that it fails somehow) is not enough to suggest definite theoretical research. This is Wheeler's principle of " radical conservatism ".

Things are, of course, quite the other way for experimental research. One should look very hard for an " expected " failure. I have probably been converted from my prejudice, that it must fail, just in time to be caught off base by an experiment next month showing that indeed it does.

REFERENCES

(1) Report of the Solvay Conference (1911).

(2) P. Dirac, *Proceedings of the Royal Society* (1927);
E. Fermi, *Rev. Mod. Phys.*, **4**, 87 (1932).

(3) See, for example, J.S. Schwinger, " Quantum Electrodynamics ", Dover Pub. Co. (1958); J.M. Jauch and F. Rohrlich, " The Theory of Photons and Electrons ", Addison-Wesley Pub. Co. (1955); S.S. Schweber, " An Introduction to Relativistic Quantum Field Theory ", Row Peterson and Co. (1961).

(4) Hofstadter and Herman, *P.R.L.*, **6**, 293 (1961);
Olsen, Schopper and Wilson, *P.R.L.*, **6**, 286 (1961).

(5) F. Fabianini *et al.*, Conference at Aix-en-Provence (1961).

(6) G. Backenstoss *et al.*, Conference at Aix-en-Provence (1961).

(7) See W.E. Lamb, T.M. Saunders, *P. R.*, **119**, 1901 (1960) and L. Wilcox, W.E. Lamb, *P. R.*, **119**, 1915 (1960) for a complete set of references.

(8) All calculations in this paper are made on the basis of the latest (1961), least squares adjustment of the fundamental constants, by J. DuMond and R. Cohen. We are indebted to them for a private communication of the results of this adjustment.

(9) A. Petermann, *Fortschr. Physik*, **6**, 505 (1958).

(10) J. Harriman, *P. R.*, **101**, 594 (1956);
H. Bethe, L. Brown, and J. Stehn, *P. R.*, **77**, 370 (1950).

(11) M. Baranger, H. Bethe, and R. Feynman, *P. R.*, **92**, 482 (1953);
R. Karplus, A. Klein, and J. Schwinger, *P. R.*, **86**, 288 (1952).

(12) A.J. Layzer, *P.R.L.*, **4**, 580 (1960).

(13) A. Petermann, *Helv. Phys. Acta*, **30**, 407 (1957);
C.M. Sommerfield, *P. R.*, **107**, 328 (1957).

(14) J. Weneser, R. Bersohn, and N. Kroll, *P. R.*, **91**, 1257 (1953).

(15) M. Baranger, F. Dyson, and E. Salpeter, *P. R.*, **88**, 680 (1952);
G. Källen and A. Sabry, *K. Dansk. Videk Medd.*, **29**, No. 17 (1955).

(16) E. Wichmann and N.M. Kroll, *P. R.*, **96**, 232 (1954).

(17) E. Salpeter, *P. R.*, **87**, 328 (1953).

(18) T. Fulton and P. Martin, *P. R.*, **95**, 811 (1954).

(19) B.D. Lambe and R.H. Dicke, Lambe thesis, Princeton (1959).

(20) S. Triebwasser, E. Dayhoff, and W. Lamb, *P. R.*, **89**, 98 (1953); **89**, 106 (1953).

(21) H. Suura and K. Wichmann, *P. R.*, **105**, 1930 (1957);
A. Petermann, *P. R.*, **105**, 1931 (1957).

(22) Beringer and Heald, *P. R.*, **95**, 1474 (1954).

(23) Franken and Liebes, *P. R.*, **104**, 1197 (1956).

(24) Schupp, Pidd and Crane, *P. R.*, **121**, 1 (1961).

(25) Charpak *et al.*, *P.R.L.*, **6**, 128 (1961).

(26) Anderson *et al.*, *P.R.L.*, **4**, 69 (1960);
Kush, *P. R.*, **100**, 1188 (1955);
Witte and Dicke, *P. R.*, **96**, 530 (1954);
Kusch, *P. R.*, **106**, 89 (1957).

(27) A very good discussion of the theory giving a complete set of references may be found in Zwanziger, *P. R.*, **121**, 1128 (1961).

(28) Arnowitz, *P. R.*, **92,** 1002 (1953);
Newcomb and Salpeter, *P. R.*, **97,** 1146 (1955).

(29) R. Karplus and A. Klein, *P. R.*, **86,** 257 (1952).

(30) R. Serber, *P. R.*, **48,** 49 (1935), or, for example, R.P. Feynman, *P. R.*, **76,** 769 (1949).

(31) These matters are discussed in detail in other reports at this conference, but because I am not sure exactly what will be said, I have included this section, in spite of the undoubted duplications of material.

(32) For example, see K. Nishijima, *P. R.*, **122,** 298 (1961).

(33) Some apparent advances in this direction are being made by several authors. See, for example, work by Amati, Fubini *et al.*, and the report by Drell at the Conference at Aix-en-Provence (1961).

(34) The quantity a_1 will be further renormalized by the effects of terms such as the c'' term before it can be compared to a.

Discussion des rapports de Heitler et Feynman

W. Heisenberg. — From Heitler's lecture we have learned that Lorentz-invariance and local interaction seem not to be compatible in the conventional quantum-theory. On the other hand Feynman explained to us that quantum electrodynamics gives very accurate results on a very wide range of phenomena. These facts suggest, as has often been discussed, that in Qu. E. Dyn. we might meet a situation very similar to the Lee model with local interaction. Let me specify this assumption somewhat further : it would mean that by the process of infinite renormalization we have unconsciously introduced " ghost-states " of very high energy, i.e. an indefinite metric in Hilbert space. If this was true it would easily explain why for all low energy phenomena Qu. E. Dyn. gives excellent results and is a perfectly " closed theory ". At higher energies in realistic physics some modifications of Qu. E. Dyn. will occur since there will be the possibility of creating pairs of nucleons, π-mesons, etc. It may in fact be that the ghost states could be identified to some extend with the baryons, since the norm of baryon-states may have opposite sign to the norm of the electron-states. (That would not in itself interfere with the unitarity of the S-matrix, since we have baryon and lepton-conservation.) At the same time the indefinite metric would explain why it has not been possible to formulate Qu. E. Dyn. without these divergences and limiting processes. Because if, like in the Lee-model, the renormalized operators commute or anti-commute everywhere for a given time, then these operators at a given time are not sufficient to define the complete Hilbert space and we would need the operators in some arbitrarily small but finite time interval in order to define it. This would be equivalent to an infinite renormalization. Quite generally, and independently of Qu. E. Dyn. it seems natural to assume that a local interaction will have the tendency to eliminate the δ-functions on the light cone in the commutators and replace them by a minor singularity — which would be equivalent to an indefinite metric in Hilbert space. Certainly we have no general proof, that we can in

such a formalism always avoid the well known difficulties with the probability interpretation. On the other hand several cases are known in which an indefinite metric in Hilbert space is compatible with a unitary S-matrix. Therefore we should investigate this possibility for reconciling local interaction and Lorentz-invariance, unless some better solution can be suggested.

P.A.M. Dirac. — I would like to give briefly my point of view with regard to field theory. The foundation of atomic physics is the superposition principle, which says that states are of such a character that they can be added together to give other quantities of the same nature. The states must be pictured as embedded in space-time; so that if one is given a state, one can apply to it various operations of rotation and translation to get other states. These operations form a group, the inhomogeneous Lorentz group. It follows that the states provide a representation of the inhomogeneous Lorentz group. The problem of setting up a quantum theory thus becomes the problem of finding a certain representation of the inhomogeneous Lorentz group.

One could attack the problem by looking for all the representations of the inhomogeneous Lorentz group. This method was followed by Wigner in 1939. He expressed the representations in terms of the irreducible representations. The irreducible representations correspond to particles by themselves. Particles in interaction also correspond to representations, but reducible ones. All the work that has been done on quantum field theory may be looked upon as attempts to set up a suitable representation of the inhomogeneous Lorentz group corresponding to physical reality. The attempts fail because of the infinities and produce nothing of mathematical significance.

I think it would be worth while to work on the problem from a more general mathematical basis, in which one does not necessarily build up the representation in terms of field quantities suggested by existing physical theories. The task of primary importance is to get the mathematical relations right. One can then afterwards look for the physical interpretation of the various quantities that enter into the mathematical scheme.

W. Heitler. — I would have no objection to the use of indefinite metric provided it can be done without inconsistencies. When

Heisenberg suggests that quantum-electrodynamics is a " closed theory ", this surely can apply to electrons only and implies that the self mass remains infinite. This would not permit a treatment of the electrodynamics of bosons and would not permit a calculation of the mass differences.

W. Heisenberg. — For π-mesons and nucleons it is more reasonable to start not from Qu. E. Dyn., but from an entirely different scheme.

A.S. Wightman. — I should like to comment on Heitler's non-invariant theory. Such theories are of interest from a point of view quite different from that which Heitler considered. They can be studied for the light which they may throw on the theory without cut-off. For this purpose one must examine the dependence on the cut-off of the various quantities occuring in theory. Normally, one considers two cut-offs in this connection, an ultra violet cut-off and a box. In Heitler's theory there is no box and the theory appears Euclidean invariant. This gives rise to phenomena which, I believe, could strongly affect Heitler's conclusions, whatever the purpose for which the theory is studied. What I have in mind is the Haag theorem which says that in a Euclidean invariant theory in which there are canonical variables, the no particle state is necessarily Euclidean invariant. Now the only Euclidean invariant state which admits a reasonable physical interpretation is the physical vacuum. Since in any theory where there is non-trivial pair creation (as in Heitler's) the no particle state is not stationary, there is no reasonable vacuum state. The only way out of this difficulty is to use one of the so-called strange representations of the commutation relations, but in that case the evaluation of the physical quantities of the theory will certainly require some better technique than perturbation theory.

G. Källén. — Perhaps I may be allowed to formulate what Wightman has just said in a slightly different way. In any ordinary field theory you have the interacting field $A(x)$ and the asymptotic incoming field $A^{(in)}(x)$. Further you have a Hamiltonian

$$H(A) = H_0(A) + H_{int}(A).$$

Especially in the old times one often introduced states, mathematical states, as eigenstates of $H_o(A)$

$$H_o(A) \mid n, \text{math}> = E_n \mid n, \text{math}>.$$

In contradistinction to this we have the physical states which are eigenstates of the total Hamiltonian or, which is practically the same thing, $H_o(A^{(in)}) \neq H_o(A)$

$$H_o(A^{(in)}) \mid n, \text{phys.}> = E_n \mid n, \text{phys.}>.$$

Many times one tries to write expansions of the form

$$\mid n, \text{phys.}> = \sum_{n'} C_{nn'} \mid n', \text{math.}>.$$

The so-called " Haag theorem " says that the transformation $C_{n,n'}$ is indeed, very singular from the mathematical point of view. I agree that that is certainly so also for the theory Heitler discussed yesterday. However, I also believe that this fact is not really very serious. If one computes more physical quantities like scattering amplitudes or even self-masses, they can very well exist even if $C_{nn'}$ is singular. Therefore, I believe that this particular argument against the Heitler model is not very relevant.

A.S. Wightman. — The problem stated by Professor Dirac can be regarded as half-solved. I believe that we know *up to unitary equivalence* the reducible unitary representation of the inhomogeneous Lorentz group which belongs to a physical theory with given stable particles. It may be displayed as the representation of the free field theory of particles of the same masses.

To go further one must specify the physical observables of the theory. For the theory of a scalar field, for example, one has

$$U(a, \Lambda)\, \Phi(x)\, U(a, \Lambda)^{-1} = \Phi(\Lambda x + a),$$

$$[\Phi(x), \Phi(y)] = 0 \qquad (x^2 - y^2) \text{ spacelike.}$$

Here $U(a, \Lambda)$ is the unitary representation of the Lorentz group. I believe that the problem of finding the Φ with these properties is a well posed one mathematically and the solutions would give a

natural expression for the basic ideas of field theory put forward thirty some years ago by Dirac, Heisenberg and Pauli. Unfortunately, it is as yet unsolved. Until we understand its solutions or lack of them I feel there will be no physical paradox in the foundations of field theory, just a muddle.

L. Van Hove. — I would like to ask a question to Heisenberg concerning his remarks at the opening of this discussion. If, as you suggest, the baryons regularize the leptons and vice versa, don't you expect that the vacuum polarization effects of the baryons which become important in high energy electrodynamics may come out to be different, for example in sign, from what conventional theory predicts ?

W. Heisenberg. — If the baryons have opposite norm to that of the leptons, this should certainly have some influence on vacuum polarization effects of the baryons. Whether it would change the sign of these effects could probably be answered only by a careful investigation; at least I don't know the answer.

G. Chew. — If none of the strongly interacting particles is elementary, there should be no divergences in calculating electromagnetic mass splittings of isotopic multiplets, even with existing rules of electrodynamics. (Think, for example, of the Coulomb splittings of He^3 and H^3.) The principle that none of the strongly interacting particles is elementary is a feature of Heisenberg's theory and can be incorporated into the S-matrix theory — even though the notion is awkward in conventional (Lagrangian) field theory. One may hope therefore, not to need cut-offs when such mass calculations are finally carried out.

W. Heisenberg. — With regard to Wightman's remark, I would like to emphasize, that in my opinion wellknown difficulties of divergences, etc. are not primarily mathematical problems. Quantum field theory is in two respects essentially different in its physical content from quantum mechanics :

(1) In field theory the interactions are local while in quantum mechanics they are non local.

(2) In field theory we have three boundary conditions while in quantum mechanics we have only two (at infinitely small and in-

finitely large distances of particles). The third boundary condition in field theory refers to an infinite number of particles.

It is a problem of physics and not only of mathematics to see how such a profound change as the replacement of non local interactions by local ones will affect the mathematical representation. We cannot expect that such a change could be represented without radical changes in the formalism of quantum theory.

With respect to the problem of the third boundary we should try to get some information from the experiments on multiple production of particles.

Concerning the views expressed by Chew, I think I can agree in principle with most of his points. It perhaps should be possible in principle to construct the S-matrix simply by considering the group structure of the system of elementary particles and adding the postulate of unitarity and analyticity (except at points representing physical states) in order to represent causality. All this could probably be done without the use of an indefinite metric. On the other hand I cannot see how from a practical point of view one could deal with the enormous complexity of the analytic behaviour of S-matrix elements, without deriving them from some kind of " local interaction ". One should also keep in mind that by the postulate of analyticity one goes already away from the energy-shell into the more " local " regions; and discussing these regions without indefinite metric may be just as complicated as for instance a discussion of the fundamental laws of algebra without introducing $\sqrt{-1}$. But aside from these practical points I would approve of the views expressed by Chew.

L. Van Hove. — At center of mass energies of the order of 1 Gev and higher the vacuum polarization effects of strongly interacting particles become a more important part of radiative corrections than at low energies. Could Feynman comment on the implications of this fact if quantum electrodynamics would be " exact, although incomplete " ?

R.P. Feynman. — I do not want to give a precise answer to this question. It might be possible to write electrodynamics with e and μ, which would be complete but incorrect.

R. Peierls. — Even if there exists a modified form of electrodynamics in which there are no infinities it may not be easy to discover this from the study of electrodynamics by itself, because the modifications are likely to relate to extremely high energies (small distances) and their discussion becomes very academic. The similar difficulties in the theory of strong interactions are very substantial in the region of practical interest. It therefore seems likely that we may first discover the remedy (if it exists) in the strong interactions, and then recognize how to remove the troubles of electrodynamics by similar means.

A. Salam. — I would like to come back to Heitler's formula connecting energy and momentum for an electron and the violation of Lorentz invariance. If so perhaps we could have a discussion of the experimental situation.

W. Heitler. — Concerning the variation of mass with velocity : the formula given in my report for the self mass $\delta m(p)$ refers to free electrons only (not to bound electrons). The coefficients depend very much on the choice of the form factor and it may even be that a form factor exists for which δm is independent of p. Personnally, of course, I do not believe that any departure of this sort from invariance exists, this was merely meant to show what kind of results arise when one insists on the finiteness of the theory, and to suggest that such fundamental relation as Einstein's mass-velocity relation, should be checked as accurately as possible by experiments.

R.P. Feynman. — If the relativity formula is wrong, there are two masses that can be defined : the rest mass or self energy, and the coefficient of $v^2/2$ in the energy for small velocities or the " kinetic mass ". If gravity acts on energy, or, therefore rest mass, and " kinetic mass " represents inertia then we know from the experiment of Eotvös that they differ by less than one part in 10^{-8}. Therefore the electromagnetic part of the proton mass, being of order 10^{-3} must have the right coefficient of v^2 to order 10^{-5}. I pointed out yesterday that if the electron energy had a velocity dependence of the form $m_0(1 + \frac{1}{2} v^2/c^2 + a_1 v^4/c^4)$ the Lamb experiment shows that $a = 3/8$ to one part in 10^{-6} so that if a fractional part of the mass of order 0.1 % is electromagnetic and varies in a

new way with velocity, then it must have the right coefficient of v^4 correct to one part in 10^{-3}.

L. Van Hove. — Present mass determination of the particles at ultrarelativistic energies are, I think, of an order of accuracy of a few percent. Is this sufficient for the question you raised ?

W. Heitler. — I think that an accuracy of 1% would not be good enough.

H.A. Bethe. — I want to draw your attention to *two experiments*, the first of which may be legitimate, in Heitler's sense, and the second may not be. The first experiment compares directly the rest energy of an electron to the kinetic mass of the electron; the kinetic mass is known with absolute accuracy; the rest energy can be measured by determining the energy necessary to produce an electron and a positron from radioactivity; this is known in some cases to a few parts in ten thousand and therefore the equality of rest mass and kinetic mass is established to a few parts in ten thousand. An even more direct and accurate way to measure the rest energy is by means of the wave length of annihilation radiation which has been measured by Dumond at Cal. Tech., also to an accuracy of a few parts in 10,000. The second experiment, I want to mention, refers to high energy; one can measure the total energy of the particle and measure the difference between the velocity of the particle and the velocity of light. This difference for synchrotrons giving electrons of 1 Gev is something like one part in 10^7, and one can measure this difference very accurately by measuring the total intensity of the light emitted in synchrotron radiation. One can measure this difference to something like 1% by measuring the intensity of the light; this has been done at 300 Mev and there is a project for 1 Gev but I am sure it would give the right result. This would give the velocity of the particle to one part in 10^9. Thus it is established with phenomenal accuracy that the velocity of an electron actually approaches that of light. The rest mass of a fast electron is measured by the difference $1 - \beta$ and is therefore known, for 300 Mev electrons, to an accuracy of about 1%. It is of course equal to the familiar kinetic mass. All the experiments I mentioned refer to free electrons.

III. Path Integrals and Operator Calculus: QED and Other Applications

The Feynman path integral approach to quantum mechanics stands presently on a par, both esthetically and practically, with the original formulations of 1925–26.[1] The path integral formulation of the quantum gauge theories which lie at the heart of the Standard Model of elementary particle interactions turned out to be critical in the Veltman–'t Hooft proof that these theories are renormalizable. However, it has an even wider range of applications than to quantum field theories. Feynman's book of 1965 with Albert R. Hibbs [64] uses path integrals to treat problems other than quantum mechanics and quantum electrodynamics, including statistical mechanics, the variational principle, the polaron problem, Brownian motion, and noise. Other applications have been made to quantum liquids and solids, to macromolecules and polymers, and to problems of propagation in dissipative media. The approach is important to various forms of semiclassical approximations in chemical, atomic, and nuclear problems and basic to the instanton problem (barrier penetration between different vacuum ground states). It can be extended to optics and even to the motion of particles in the strong gravitational fields near a black hole.

As Feynman related in his Nobel Lecture [73], he developed the path integral formulation of quantum mechanics in order to quantize the action-at-a-distance theory of classical electrodynamics and thus avoid problems arising in field theory from the self-interaction of the electron. The methods of quantization associated with the names of Heisenberg, Schrödinger, and Dirac all begin with the Hamiltonian function as the generator of the evolution of the system with time, while the classical action-at-a-distance theory used a classical action principle based on the Lagrangian. This involved the interaction of two currents, each a function of an independent space–time variable. In [73] Feynman described how he discovered (with the help of Herbert Jehle) an infinitesimal time development operator of Dirac that involved the classical Lagrangian. Successive applications of this operator to the initial wave function generated the wave function at any later time, and was equivalent to finding the solution of Schrödinger's equation. To obtain the wave function after a *finite* elapsed time, however, one had to integrate over all possible paths connecting two arbitrary space–time points. This is the path integral approach of Feynman.

Although it is possible to begin with the Hamiltonian and arrive at a "manifestly covariant" relativistic QED, as shown by Julian Schwinger and Sin-itiro Tomonaga, with whom Feynman shared the Nobel Prize, the overall space–time point of view of Feynman lends itself to a relativistic formulation in a more natural way because the action, i.e. the time integral of the classical Lagrangian, is a relativistic invariant, while the Hamiltonian function is not.

Before leaving for Los Alamos in the spring of 1942, Feynman wrote his doctoral dissertation on the path integral method. The essential parts of his thesis were published in 1948 in paper [7]. In it, the electrons are treated nonrelativistically and the electromagnetic field is described by its Fourier transform, i.e. by the so-called field oscillators. Fermi's version of QED represented the electromagnetic field this way, and then quantized the classical oscillators. Feynman, instead, *eliminated* the field oscillators, integrating out their coordinates and

[1] "One cannot fail to observe that Feynman's principle in particular — and this is no hyperbole — expresses the laws of quantum mechanics in an exemplary neat and elegant manner, notwithstanding the fact that it employs somewhat unconventional mathematics." — W. Yourgrau and S. Mandelstam, *Variational Principles in Dynamics and Quantum Theory* (Philadelphia 1968), p. 128.

leaving only the "source current" and the "test current." (We use the quotation marks because the two currents are interchangeable.) *All* the oscillators were eliminated. In previous treatments the electromagnetic longitudinal and timelike degrees of freedom were combined to yield the Coulomb interaction, while the transverse degrees of freedom were retained as field oscillators.

In the following year, 1949, Feynman published papers [12] and [13], which are included in Part II.C, using the rules that he had derived for QED by applying the path integral method. In 1950, he wrote paper [14], which establishes the validity of the rules, i.e. the Feynman diagram methods. This paper starts with Fermi's formulation of the field as a set of oscillators and completely eliminates them, as Feynman had done already in paper [7]. Here, however, the charges are treated in a completely relativistic manner, using either the Dirac electron–positron field or, for spinless particles, the relativistic Klein–Gordon (or Pauli–Weisskopf) field. All virtual photons are eliminated; and Feynman shows how real photons can be either introduced *ab initio* or derived from the general formulae for virtual processes.

In 1951, Feynman completed his work on QED in paper [15], describing and applying a new operator calculus for operators dependent on a continuous parameter (such as the time), which describes the ordering of application of the generally noncommuting operators. Using the ordering parameter, the ordinary methods of calculus (integration and power series expansions, for example) can be carried through without regard to the ordering until the final result, which is then appropriately reordered. In this paper, Feynman applied this method to QED and pointed out that while it did not lead to any new results, it made it easier to relate the path integral formulation of QED to the more conventional approaches of Schwinger and Tomonaga. An advantage of the ordered operator calculus over the path integral method is that it can deal more easily with half-integral spin.[2] Evidently Feynman hoped that his new method for dealing with operators would have a wider application, but that does not seem to have been the case (at least so far!).

Among the applications of the path integral method treated by Feynman and Hibbs, we find statistical mechanics,[3] a new variational principle, Brownian motion, and "other problems in probability." This is only a small subset of the many applications which have been found for the method, especially in condensed matter physics.[4] Papers [28] and [49] develop the new variational principle and apply it to the polaron problem, which Feynman describes in [28] as follows:

> An electron in an ionic crystal polarizes the lattice in its neighborhood. This interaction changes the energy of the electron. Furthermore, when the electron moves the polarization state must move with it. An electron moving with its accompanying distortion of the lattice has sometimes been called a polaron. It has an effective mass higher than that of the electron. We wish to compute the energy and effective mass of such an electron.

Feynman developed the new variational principle and used it to obtain a lower bound for the polaron self-energy that was lower than that obtained in five other papers published

[2]See [64], Feynman and Hibbs, p. 355.

[3]See also [88].

[4]See, e.g., Martin C. Gutzwiller, "Resource letter ICQM-1: the interplay between classical and quantum mechanics," *Am. J. Phys.* **66** (1998): 304–324. Among items dealing explicitly with Feynman path integrals are 71–73 and 158–168.

in the early 1950s. Paper [49] employs a similar method to calculate the polaron's mobility. It is a collaboration between Feynman, Robert W. Hellworth (of the Hughes Research Laboratory, where Feynman was a part-time consultant), and two Caltech graduate students: Carl K. Iddings and Phillip M. Platzman.

Paper [55] is part of the doctoral thesis of Frank L. Vernon, Jr., supervised by Feynman. It was not Feynman's usual practice for his name to appear in publications of his students' dissertations, so we can assume he was closely involved in the subject.[5] It deals with a general quantum-mechanical system (e.g. an atom or molecule) interacting with a linear dissipative system. The latter is represented by a collection of harmonic oscillators. In that sense it resembles QED and, as in that case, the oscillators' coordinates are integrated out so that their total effect is replaced by that of an "influence functional." According to the abstract, "In addition, a fluctuation–dissipation theorem is derived relating temperature and dissipation of the linear system to a fluctuating classical potential acting on the system of interest which reduces to the Nyquist–Johnson relation for noise in the case of electric circuits."

Selected Papers

[7] Space–time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.* **20** (1948): 367–387.

[14] Mathematical formulation of the quantum theory of electromagnetic interaction. *Phys. Rev.* **80** (1950): 440–457.

[15] An operator calculus having application in quantum electrodynamics. *Phys. Rev.* **84** (1951): 108–128.

[28] Slow electrons in a polar crystal. *Phys. Rev.* **97** (1955): 660–665.

[49] With R.W. Hellwarth, C.K. Iddings, and R.M. Platzman. Mobility of slow electrons in a polar crystal. *Phys. Rev.* **127** (1962): 1004–1017.

[55] With F.L. Vernon, Jr.. The theory of a general quantum mechanical system interacting with a linear dissipative system. *Ann. Phys.* **24** (1963): 118–173.

[5] According to Frank Vernon (private communication), "I asked [Feynman] if he would coauthor the paper. He seemed surprised but agreed."

Space-Time Approach to Non-Relativistic Quantum Mechanics

R. P. FEYNMAN

Cornell University, Ithaca, New York

Non-relativistic quantum mechanics is formulated here in a different way. It is, however, mathematically equivalent to the familiar formulation. In quantum mechanics the probability of an event which can happen in several different ways is the absolute square of a sum of complex contributions, one from each alternative way. The probability that a particle will be found to have a path $x(t)$ lying somewhere within a region of space time is the square of a sum of contributions, one from each path in the region. The contribution from a single path is postulated to be an exponential whose (imaginary) phase is the classical action (in units of $\hbar$) for the path in question. The total contribution from all paths reaching x, t from the past is the wave function $\psi(x, t)$. This is shown to satisfy Schroedinger's equation. The relation to matrix and operator algebra is discussed. Applications are indicated, in particular to eliminate the coordinates of the field oscillators from the equations of quantum electrodynamics.

1. INTRODUCTION

IT is a curious historical fact that modern quantum mechanics began with two quite different mathematical formulations: the differential equation of Schroedinger, and the matrix algebra of Heisenberg. The two, apparently dissimilar approaches, were proved to be mathematically equivalent. These two points of view were destined to complement one another and to be ultimately synthesized in Dirac's transformation theory.

This paper will describe what is essentially a third formulation of non-relativistic quantum theory. This formulation was suggested by some of Dirac's[1,2] remarks concerning the relation of classical action[3] to quantum mechanics. A probability amplitude is associated with an entire motion of a particle as a function of time, rather than simply with a position of the particle at a particular time.

The formulation is mathematically equivalent to the more usual formulations. There are, therefore, no fundamentally new results. However, there is a pleasure in recognizing old things from a new point of view. Also, there are problems for which the new point of view offers a distinct advantage. For example, if two systems A and B interact, the coordinates of one of the systems, say B, may be eliminated from the equations describing the motion of A. The inter-

[1] P. A. M. Dirac, *The Principles of Quantum Mechanics* (The Clarendon Press, Oxford, 1935), second edition, Section 33; also, Physik. Zeits. Sowjetunion **3**, 64 (1933).
[2] P. A. M. Dirac, Rev. Mod. Phys. **17**, 195 (1945).

[3] Throughout this paper the term "action" will be used for the time integral of the Lagrangian along a path. When this path is the one actually taken by a particle, moving classically, the integral should more properly be called Hamilton's first principle function.

action with B is represented by a change in the formula for the probability amplitude associated with a motion of A. It is analogous to the classical situation in which the effect of B can be represented by a change in the equations of motion of A (by the introduction of terms representing forces acting on A). In this way the coordinates of the transverse, as well as of the longitudinal field oscillators, may be eliminated from the equations of quantum electrodynamics.

In addition, there is always the hope that the new point of view will inspire an idea for the modification of present theories, a modification necessary to encompass present experiments.

We first discuss the general concept of the superposition of probability amplitudes in quantum mechanics. We then show how this concept can be directly extended to define a probability amplitude for any motion or path (position *vs.* time) in space-time. The ordinary quantum mechanics is shown to result from the postulate that this probability amplitude has a phase proportional to the action, computed classically, for this path. This is true when the action is the time integral of a quadratic function of velocity. The relation to matrix and operator algebra is discussed in a way that stays as close to the language of the new formulation as possible. There is no practical advantage to this, but the formulae are very suggestive if a generalization to a wider class of action functionals is contemplated. Finally, we discuss applications of the formulation. As a particular illustration, we show how the coordinates of a harmonic oscillator may be eliminated from the equations of motion of a system with which it interacts. This can be extended directly for application to quantum electrodynamics. A formal extension which includes the effects of spin and relativity is described.

2. THE SUPERPOSITION OF PROBABILITY AMPLITUDES

The formulation to be presented contains as its essential idea the concept of a probability amplitude associated with a completely specified motion as a function of time. It is, therefore, worthwhile to review in detail the quantum-mechanical concept of the superposition of probability amplitudes. We shall examine the essential changes in physical outlook required by the transition from classical to quantum physics.

For this purpose, consider an imaginary experiment in which we can make three measurements successive in time: first of a quantity A, then of B, and then of C. There is really no need for these to be of different quantities, and it will do just as well if the example of three successive position measurements is kept in mind. Suppose that a is one of a number of possible results which could come from measurement A, b is a result that could arise from B, and c is a result possible from the third measurement C.[4] We shall assume that the measurements A, B, and C are the type of measurements that completely specify a state in the quantum-mechanical case. That is, for example, the state for which B has the value b is not degenerate.

It is well known that quantum mechanics deals with probabilities, but naturally this is not the whole picture. In order to exhibit, even more clearly, the relationship between classical and quantum theory, we could suppose that classically we are also dealing with probabilities but that all probabilities either are zero or one. A better alternative is to imagine in the classical case that the probabilities are in the sense of classical statistical mechanics (where, possibly, internal coordinates are not completely specified).

We define P_{ab} as the probability that if measurement A gave the result a, then measurement B will give the result b. Similarly, P_{bc} is the probability that if measurement B gives the result b, then measurement C gives c. Further, let P_{ac} be the chance that if A gives a, then C gives c. Finally, denote by P_{abc} the probability of all three, i.e., if A gives a, then B gives b, and C gives c. If the events between a and b are independent of those between b and c, then

$$P_{abc}=P_{ab}P_{bc}. \tag{1}$$

This is true according to quantum mechanics when the statement that B is b is a complete specification of the state.

[4] For our discussion it is not important that certain values of a, b, or c might be excluded by quantum mechanics but not by classical mechanics. For simplicity, assume the values are the same for both but that the probability of certain values may be zero.

In any event, we expect the relation

$$P_{ac} = \sum_b P_{abc}. \tag{2}$$

This is because, if initially measurement A gives a and the system is later found to give the result c to measurement C, the quantity B must have had some value at the time intermediate to A and C. The probability that it was b is P_{abc}. We sum, or integrate, over all the mutually exclusive alternatives for b (symbolized by $\sum_b$).

Now, the essential difference between classical and quantum physics lies in Eq. (2). In classical mechanics it is always true. In quantum mechanics it is often false. We shall denote the quantum-mechanical probability that a measurement of C results in c when it follows a measurement of A giving a by $P_{ac}{}^q$. Equation (2) is replaced in quantum mechanics by this remarkable law:[5] There exist complex numbers φ_{ab}, φ_{bc}, φ_{ac} such that

$$P_{ab} = |\varphi_{ab}|^2, \quad P_{bc} = |\varphi_{bc}|^2, \quad \text{and} \quad P_{ac}{}^q = |\varphi_{ac}|^2. \tag{3}$$

The classical law, obtained by combining (1) and (2),

$$P_{ac} = \sum_b P_{ab} P_{bc} \tag{4}$$

is replaced by

$$\varphi_{ac} = \sum_b \varphi_{ab} \varphi_{bc}. \tag{5}$$

If (5) is correct, ordinarily (4) is incorrect. The logical error made in deducing (4) consisted, of course, in assuming that to get from a to c the system had to go through a condition such that B had to have some definite value, b.

If an attempt is made to verify this, i.e., if B is measured between the experiments A and C, then formula (4) is, in fact, correct. More precisely, if the apparatus to measure B is set up and used, but no attempt is made to utilize the results of the B measurement in the sense that only the A to C correlation is recorded and studied, then (4) is correct. This is because the B measuring machine has done its job; if we wish, we could read the meters at any time without disturbing the situation any further. The experiments which gave a and c can, therefore, be separated into groups depending on the value of b.

Looking at probability from a frequency point of view (4) simply results from the statement that in each experiment giving a and c, B had some value. The only way (4) could be wrong is the statement, "B had some value," must sometimes be meaningless. Noting that (5) replaces (4) only under the circumstance that we make no attempt to measure B, we are led to say that the statement, "B had some value," may be meaningless whenever we make no attempt to measure B.[6]

Hence, we have different results for the correlation of a and c, namely, Eq. (4) or Eq. (5), depending upon whether we do or do not attempt to measure B. No matter how subtly one tries, the attempt to measure B must disturb the system, at least enough to change the results from those given by (5) to those of (4).[7] That measurements do, in fact, cause the necessary disturbances, and that, essentially, (4) could be false was first clearly enunciated by Heisenberg in his uncertainty principle. The law (5) is a result of the work of Schroedinger, the statistical interpretation of Born and Jordan, and the transformation theory of Dirac.[8]

Equation (5) is a typical representation of the wave nature of matter. Here, the chance of finding a particle going from a to c through several different routes (values of b) may, if no attempt is made to determine the route, be represented as the square of a sum of several complex quantities—one for each available route.

[5] We have assumed b is a non-degenerate state, and that therefore (1) is true. Presumably, if in some generalization of quantum mechanics (1) were not true, even for pure states b, (2) could be expected to be replaced by: There are complex numbers φ_{abc} such that $P_{abc} = |\varphi_{abc}|^2$. The analog of (5) is then $\varphi_{ac} = \sum_b \varphi_{abc}$.

[6] It does not help to point out that we *could* have measured B had we wished. The fact is that we did not.

[7] How (4) actually results from (5) when measurements disturb the system has been studied particularly by J. von Neumann (*Mathematische Grundlagen der Quantenmechanik* (Dover Publications, New York, 1943)). The effect of perturbation of the measuring equipment is effectively to change the phase of the interfering components, by θ_b, say, so that (5) becomes $\varphi_{ac} = \sum_b e^{i\theta_b} \varphi_{ab} \varphi_{bc}$. However, as von Neumann shows, the phase shifts must remain unknown if B is measured so that the resulting probability P_{ac} is the square of φ_{ac} averaged over all phases, θ_b. This results in (4).

[8] If $\mathbf{A}$ and $\mathbf{B}$ are the operators corresponding to measurements A and B, and if ψ_a and ψ_b are solutions of $\mathbf{A}\psi_a = a\psi_a$ and $\mathbf{B}\chi_b = b\chi_b$, then $\varphi_{ab} = \int \chi_b{}^* \psi_a dx = (\chi_b{}^*, \psi_a)$. Thus, φ_{ab} is an element $(a|b)$ of the transformation matrix for the transformation from a representation in which $\mathbf{A}$ is diagonal to one in which $\mathbf{B}$ is diagonal.

Probability can show the typical phenomena of interference, usually associated with waves, whose intensity is given by the square of the sum of contributions from different sources. The electron acts as a wave, (5), so to speak, as long as no attempt is made to verify that it is a particle; yet one can determine, if one wishes, by what route it travels just as though it were a particle; but when one does that, (4) applies and it does act like a particle.

These things are, of course, well known. They have already been explained many times.[9] However, it seems worth while to emphasize the fact that they are all simply direct consequences of Eq. (5), for it is essentially Eq. (5) that is fundamental in my formulation of quantum mechanics.

The generalization of Eqs. (4) and (5) to a large number of measurements, say A, B, C, D, $\cdots$, K, is, of course, that the probability of the sequence a, b, c, d, $\cdots$, k is

$$P_{abcd\cdots k} = |\varphi_{abcd\cdots k}|^2.$$

The probability of the result a, c, k, for example, if b, d, $\cdots$ are measured, is the classical formula:

$$P_{ack} = \sum_b \sum_d \cdots P_{abcd\cdots k}, \tag{6}$$

while the probability of the same sequence a, c, k if no measurements are made between A and C and between C and K is

$$P_{ack}{}^q = |\sum_b \sum_d \cdots \varphi_{abcd\cdots k}|^2. \tag{7}$$

The quantity $\varphi_{abcd\cdots k}$ we can call the probability amplitude for the condition $A=a$, $B=b$, $C=c$, $D=d$, $\cdots$, $K=k$. (It is, of course, expressible as a product $\varphi_{ab}\varphi_{bc}\varphi_{cd}\cdots\varphi_{jk}$.)

3. THE PROBABILITY AMPLITUDE FOR A SPACE-TIME PATH

The physical ideas of the last section may be readily extended to define a probability amplitude for a particular completely specified space-time path. To explain how this may be done, we shall limit ourselves to a one-dimensional problem, as the generalization to several dimensions is obvious.

[9] See, for example, W. Heisenberg, *The Physical Principles of the Quantum Theory* (University of Chicago Press, Chicago, 1930), particularly Chapter IV.

Assume that we have a particle which can take up various values of a coordinate x. Imagine that we make an enormous number of successive position measurements, let us say separated by a small time interval ϵ. Then a succession of measurements such as A, B, C, $\cdots$ might be the succession of measurements of the coordinate x at successive times t_1, t_2, t_3, $\cdots$, where $t_{i+1}=t_i+\epsilon$. Let the value, which might result from measurement of the coordinate at time t_i, be x_i. Thus, if A is a measurement of x at t_1 then x_1 is what we previously denoted by a. From a classical point of view, the successive values, x_1, x_2, x_3, $\cdots$ of the coordinate practically define a path $x(t)$. Eventually, we expect to go the limit $\epsilon\to 0$.

The probability of such a path is a function of x_1, x_2, $\cdots$, x_i, $\cdots$, say $P(\cdots x_i, x_{i+1}, \cdots)$. The probability that the path lies in a particular region R of space-time is obtained classically by integrating P over that region. Thus, the probability that x_i lies between a_i and b_i, and x_{i+1} lies between a_{i+1} and b_{i+1}, etc., is

$$\cdots \int_{a_i}^{b_i}\int_{a_{i+1}}^{b_{i+1}} \cdots P(\cdots x_i, x_{i+1}, \cdots)\cdots dx_i dx_{i+1}\cdots$$

$$= \int_R P(\cdots x_i, x_{i+1}, \cdots)\cdots dx_i dx_{i+1}\cdots, \tag{8}$$

the symbol $\int_R$ meaning that the integration is to be taken over those ranges of the variables which lie within the region R. This is simply Eq. (6) with a, b, $\cdots$ replaced by x_1, x_2, $\cdots$ and integration replacing summation.

In quantum mechanics this is the correct formula for the case that x_1, x_2, $\cdots$, x_i, $\cdots$ were actually all measured, and then only those paths lying within R were taken. We would expect the result to be different if no such detailed measurements had been performed. Suppose a measurement is made which is capable only of determining that the path lies somewhere within R.

The measurement is to be what we might call an "ideal measurement." We suppose that no further details could be obtained from the same measurement without further disturbance to the system. I have not been able to find a precise definition. We are trying to avoid the extra uncertainties that must be averaged over if, for example, more information were measured but

not utilized. We wish to use Eq. (5) or (7) for all x_i and have no residual part to sum over in the manner of Eq. (4).

We expect that the probability that the particle is found by our "ideal measurement" to be, indeed, in the region R is the square of a complex number $|\varphi(R)|^2$. The number $\varphi(R)$, which we may call the probability amplitude for region R is given by Eq. (7) with a, b, $\cdots$ replaced by x_i, x_{i+1}, $\cdots$ and summation replaced by integration:

$$\varphi(R)=\lim_{\epsilon\to 0}\int_R \times\Phi(\cdots x_i, x_{i+1}\cdots)\cdots dx_i dx_{i+1}\cdots. \quad (9)$$

The complex number $\Phi(\cdots x_i, x_{i+1}\cdots)$ is a function of the variables x_i defining the path. Actually, we imagine that the time spacing ϵ approaches zero so that Φ essentially depends on the entire path $x(t)$ rather than only on just the values of x_i at the particular times t_i, $x_i=x(t_i)$. We might call Φ the probability amplitude functional of paths $x(t)$.

We may summarize these ideas in our first postulate:

I. If an ideal measurement is performed to determine whether a particle has a path lying in a region of space-time, then the probability that the result will be affirmative is the absolute square of a sum of complex contributions, one from each path in the region.

The statement of the postulate is incomplete. The meaning of a sum of terms one for "each" path is ambiguous. The precise meaning given in Eq. (9) is this: A path is first defined only by the positions x_i through which it goes at a sequence of equally spaced times,[10] $t_i=t_{i-1}+\epsilon$. Then all values of the coordinates within R have an equal weight. The actual magnitude of the weight depends upon ϵ and can be so chosen that the probability of an event which is certain shall be normalized to unity. It may not be best to do so, but we have left this weight factor in a proportionality constant in the second postulate. The limit $\epsilon\to 0$ must be taken at the end of a calculation.

When the system has several degrees of freedom the coordinate space x has several dimensions so that the symbol x will represent a set of coordinates $(x^{(1)}, x^{(2)}, \cdots, x^{(k)})$ for a system with k degrees of freedom. A path is a sequence of configurations for successive times and is described by giving the configuration x_i or $(x_i^{(1)}, x_i^{(2)}, \cdots, x_i^{(k)})$, i.e., the value of each of the k coordinates for each time t_i. The symbol dx_i will be understood to mean the volume element in k dimensional configuration space (at time t_i). The statement of the postulates is independent of the coordinate system which is used.

The postulate is limited to defining the results of position measurements. It does not say what must be done to define the result of a momentum measurement, for example. This is not a real limitation, however, because in principle the measurement of momentum of one particle can be performed in terms of position measurements of other particles, e.g., meter indicators. Thus, an analysis of such an experiment will determine what it is about the first particle which determines its momentum.

4. THE CALCULATION OF THE PROBABILITY AMPLITUDE FOR A PATH

The first postulate prescribes the type of mathematical framework required by quantum mechanics for the calculation of probabilities. The second postulate gives a particular content to this framework by prescribing how to compute the important quantity Φ for each path:

II. The paths contribute equally in magnitude, but the phase of their contribution is the classical action (in units of $\hbar$); i.e., the time integral of the Lagrangian taken along the path.

That is to say, the contribution $\Phi[x(t)]$ from a given path $x(t)$ is proportional to $\exp(i/\hbar)S[x(t)]$, where the action $S[x(t)]=\int L(\dot{x}(t), x(t))dt$ is the time integral of the classical Lagrangian $L(\dot{x}, x)$ taken along the path in question. The Lagrangian, which may be an explicit function of the time, is a function of position and velocity. If we suppose it to be a quadratic function of the

[10] There are very interesting mathematical problems involved in the attempt to avoid the subdivision and limiting processes. Some sort of complex measure is being associated with the space of functions $x(t)$. Finite results can be obtained under unexpected circumstances because the measure is not positive everywhere, but the contributions from most of the paths largely cancel out. These curious mathematical problems are sidestepped by the subdivision process. However, one feels as Cavalieri must have felt calculating the volume of a pyramid before the invention of calculus.

velocities, we can show the mathematical equivalence of the postulates here and the more usual formulation of quantum mechanics.

To interpret the first postulate it was necessary to define a path by giving only the succession of points x_i through which the path passes at successive times t_i. To compute $S=\int L(\dot{x}, x)dt$ we need to know the path at all points, not just at x_i. We shall assume that the function $x(t)$ in the interval between t_i and t_{i+1} is the path followed by a classical particle, with the Lagrangian L, which starting from x_i at t_i reaches x_{i+1} at t_{i+1}. This assumption is required to interpret the second postulate for discontinuous paths. The quantity $\Phi(\cdots x_i, x_{i+1}, \cdots)$ can be normalized (for various ϵ) if desired, so that the probability of an event which is certain is normalized to unity as $\epsilon \to 0$.

There is no difficulty in carrying out the action integral because of the sudden changes of velocity encountered at the times t_i as long as L does not depend upon any higher time derivatives of the position than the first. Furthermore, unless L is restricted in this way the end points are not sufficient to define the classical path. Since the classical path is the one which makes the action a minimum, we can write

$$S=\sum_i S(x_{i+1}, x_i), \tag{10}$$

where

$$S(x_{i+1}, x_i)=\text{Min.}\int_{t_i}^{t_{i+1}} L(\dot{x}(t), x(t))dt. \tag{11}$$

Written in this way, the only appeal to classical mechanics is to supply us with a Lagrangian function. Indeed, one could consider postulate two as simply saying, "Φ is the exponential of i times the integral of a real function of $x(t)$ and its first time derivative." Then the classical equations of motion might be derived later as the limit for large dimensions. The function of x and $\dot{x}$ then could be shown to be the classical Lagrangian within a constant factor.

Actually, the sum in (10), even for finite ϵ, is infinite and hence meaningless (because of the infinite extent of time). This reflects a further incompleteness of the postulates. We shall have to restrict ourselves to a finite, but arbitrarily long, time interval.

Combining the two postulates and using Eq. (10), we find

$$\varphi(R)=\operatorname*{Lim}_{\epsilon\to 0}\int_R \times\exp\left[\frac{i}{\hbar}\sum_i S(x_{i+1}, x_i)\right]\cdots\frac{dx_{i+1}\,dx_i}{A\quad A}\cdots, \tag{12}$$

where we have let the normalization factor be split into a factor $1/A$ (whose exact value we shall presently determine) for each instant of time. The integration is just over those values $x_i, x_{i+1}, \cdots$ which lie in the region R. This equation, the definition (11) of $S(x_{i+1}, x_i)$, and the physical interpretation of $|\varphi(R)|^2$ as the probability that the particle will be found in R, complete our formulation of quantum mechanics.

5. DEFINITION OF THE WAVE FUNCTION

We now proceed to show the equivalence of these postulates to the ordinary formulation of quantum mechanics. This we do in two steps. We show in this section how the wave function may be defined from the new point of view. In the next section we shall show that this function satisfies Schroedinger's differential wave equation.

We shall see that it is the possibility, (10), of expressing S as a sum, and hence Φ as a product, of contributions from successive sections of the path, which leads to the possibility of defining a quantity having the properties of a wave function.

To make this clear, let us imagine that we choose a particular time t and divide the region R in Eq. (12) into pieces, future and past relative to t. We imagine that R can be split into: (a) a region R', restricted in any way in space, but lying entirely earlier in time than some t', such that $t'<t$; (b) a region R'' arbitrarily restricted in space but lying entirely later in time than t'', such that $t''>t$; (c) the region between t' and t'' in which all the values of x coordinates are unrestricted, i.e., all of space-time between t' and t''. The region (c) is not absolutely necessary. It can be taken as narrow in time as desired. However, it is convenient in letting us consider varying t a little without having to redefine R' and R''. Then $|\varphi(R', R'')|^2$ is the probability that the

path occupies R' and R''. Because R' is entirely previous to R'', considering the time t as the present, we can express this as the probability that the path had been in region R' and will be in region R''. If we divide by a factor, the probability that the path is in R', to renormalize the probability we find: $|\varphi(R', R'')|^2$ is the (relative) probability that if the system were in region R' it will be found later in R''.

This is, of course, the important quantity in predicting the results of many experiments. We prepare the system in a certain way (e.g., it was in region R') and then measure some other property (e.g., will it be found in region R''?). What does (12) say about computing this quantity, or rather the quantity $\varphi(R', R'')$ of which it is the square?

Let us suppose in Eq. (12) that the time t corresponds to one particular point k of the subdivision of time into steps ϵ, i.e., assume $t=t_k$, the index k, of course, depending upon the subdivision ϵ. Then, the exponential being the exponential of a sum may be split into a product of two factors

$$\exp\left[\frac{i}{\hbar}\sum_{i=k}^{\infty} S(x_{i+1}, x_i)\right] \cdot \exp\left[\frac{i}{\hbar}\sum_{i=-\infty}^{k-1} S(x_{i+1}, x_i)\right]. \quad (13)$$

The first factor contains only coordinates with index k or higher, while the second contains only coordinates with index k or lower. This split is possible because of Eq. (10), which results essentially from the fact that the Lagrangian is a function only of positions and velocities. First, the integration on all variables x_i for $i>k$ can be performed on the first factor resulting in a function of x_k (times the second factor). Next, the integration on all variables x_i for $i<k$ can be performed on the second factor also, giving a function of x_k. Finally, the integration on x_k can be performed. That is, $\varphi(R', R'')$ can be written as the integral over x_k of the product of two factors. We will call these $\chi^*(x_k, t)$ and $\psi(x_k, t)$:

$$\varphi(R', R'') = \int \chi^*(x, t)\psi(x, t)dx, \quad (14)$$

where

$$\psi(x_k, t) = \operatorname*{Lim}_{\epsilon\to 0} \int_{R'} \times\exp\left[\frac{i}{\hbar}\sum_{i=-\infty}^{k-1} S(x_{i+1}, x_i)\right]\frac{dx_{k-1}}{A}\frac{dx_{k-2}}{A}\cdots, \quad (15)$$

and

$$\chi^*(x_k, t) = \operatorname*{Lim}_{\epsilon\to 0} \int_{R''} \exp\left[\frac{i}{\hbar}\sum_{i=k}^{\infty} S(x_{i+1}, x_i)\right] \cdot\frac{1}{A}\frac{dx_{k+1}}{A}\frac{dx_{k+2}}{A}\cdots. \quad (16)$$

The symbol R' is placed on the integral for ψ to indicate that the coordinates are integrated over the region R', and, for t_i between t' and t, over all space. In like manner, the integral for χ^* is over R'' and over all space for those coordinates corresponding to times between t and t''. The asterisk on χ^* denotes complex conjugate, as it will be found more convenient to define (16) as the complex conjugate of some quantity, χ.

The quantity ψ depends only upon the region R' previous to t, and is completely defined if that region is known. It does not depend, in any way, upon what will be done to the system after time t. This latter information is contained in χ. Thus, with ψ and χ we have separated the past history from the future experiences of the system. This permits us to speak of the relation of past and future in the conventional manner. Thus, if a particle has been in a region of space-time R' it may at time t be said to be in a certain condition, or state, determined only by its past and described by the so-called wave function $\psi(x, t)$. This function contains all that is needed to predict future probabilities. For, suppose, in another situation, the region R' were different, say r', and possibly the Lagrangian for times before t were also altered. But, nevertheless, suppose the quantity from Eq. (15) turned out to be the same. Then, according to (14) the probability of ending in any region R'' is the same for R' as for r'. Therefore, future measurements will not distinguish whether the system had occupied R' or r'. Thus, the wave function $\psi(x, t)$ is sufficient to define those attributes which are left from past history which determine future behavior.

Likewise, the function $\chi^*(x, t)$ characterizes the experience, or, let us say, experiment to which the system is to be subjected. If a different region, r'' and different Lagrangian after t, were to give the same $\chi^*(x, t)$ *via* Eq. (16), as does region R'', then no matter what the preparation, ψ, Eq. (14) says that the chance of finding the system in R'' is always the same as finding it in r''. The two "experiments" R'' and r'' are equivalent, as they yield the same results. We shall say loosely that these experiments are to determine with what probability the system is in state χ. Actually, this terminology is poor. The system is really in state ψ. The reason we can associate a state with an experiment is, of course, that for an ideal experiment there turns out to be a unique state (whose wave function is $\chi(x, t)$) for which the experiment succeeds with certainty.

Thus, we can say: the probability that a system in state ψ will be found by an experiment whose characteristic state is χ (or, more loosely, the chance that a system in state ψ will appear to be in χ) is

$$\left|\int \chi^*(x, t)\psi(x, t)dx\right|^2. \tag{17}$$

These results agree, of course, with the principles of ordinary quantum mechanics. They are a consequence of the fact that the Lagrangian is a function of position, velocity, and time only.

6. THE WAVE EQUATION

To complete the proof of the equivalence with the ordinary formulation we shall have to show that the wave function defined in the previous section by Eq. (15) actually satisfies the Schroedinger wave equation. Actually, we shall only succeed in doing this when the Lagrangian L in (11) is a quadratic, but perhaps inhomogeneous, form in the velocities $\dot{x}(t)$. This is not a limitation, however, as it includes all the cases for which the Schroedinger equation has been verified by experiment.

The wave equation describes the development of the wave function with time. We may expect to approach it by noting that, for finite ϵ, Eq. (15) permits a simple recursive relation to be developed. Consider the appearance of Eq. (15) if we were to compute ψ at the next instant of time:

$$\psi(x_{k+1}, t+\epsilon) = \int_{R'} \exp\left[\frac{i}{\hbar}\sum_{i=-\infty}^{k} S(x_{i+1}, x_i)\right] \times \frac{dx_k}{A}\frac{dx_{k-1}}{A}\cdots. \tag{15'}$$

This is similar to (15) except for the integration over the additional variable x_k and the extra term in the sum in the exponent. This term means that the integral of (15') is the same as the integral of (15) except for the factor $(1/A)\exp(i/\hbar)S(x_{k+1}, x_k)$. Since this does not contain any of the variables x_i for i less than k, all of the integrations on dx_i up to dx_{k-1} can be performed with this factor left out. However, the result of these integrations is by (15) simply $\psi(x_k, t)$. Hence, we find from (15') the relation

$$\psi(x_{k+1}, t+\epsilon) = \int \exp\left[\frac{i}{\hbar}S(x_{k+1}, x_k)\right]\psi(x_k, t)dx_k/A. \tag{18}$$

This relation giving the development of ψ with time will be shown, for simple examples, with suitable choice of A, to be equivalent to Schroedinger's equation. Actually, Eq. (18) is not exact, but is only true in the limit $\epsilon \to 0$ and we shall derive the Schroedinger equation by assuming (18) is valid to first order in ϵ. The Eq. (18) *need* only be true for small ϵ to the first order in ϵ. For if we consider the factors in (15) which carry us over a finite interval of time, T, the number of factors is T/ϵ. If an error of order ϵ^2 is made in each, the resulting error will not accumulate beyond the order $\epsilon^2(T/\epsilon)$ or $T\epsilon$, which vanishes in the limit.

We shall illustrate the relation of (18) to Schroedinger's equation by applying it to the simple case of a particle moving in one dimension in a potential $V(x)$. Before we do this, however, we would like to discuss some approximations to the value $S(x_{i+1}, x_i)$ given in (11) which will be sufficient for expression (18).

The expression defined in (11) for $S(x_{i+1}, x_i)$ is difficult to calculate exactly for arbitrary ϵ from classical mechanics. Actually, it is only necessary that an approximate expression for $S(x_{i+1}, x_i)$ be

used in (18), provided the error of the approximation be of an order smaller than the first in ϵ. We limit ourselves to the case that the Lagrangian is a quadratic, but perhaps inhomogeneous, form in the velocities $\dot{x}(t)$. As we shall see later, the paths which are important are those for which $x_{i+1}-x_i$ is of order $\epsilon^{\frac{1}{2}}$. Under these circumstances, it is sufficient to calculate the integral in (11) over the classical path taken by a *free* particle.[11]

In *Cartesian coordinates*[12] the path of a free particle is a straight line so the integral of (11) can be taken along a straight line. Under these circumstances it is sufficiently accurate to replace the integral by the trapezoidal rule

$$S(x_{i+1}, x_i) = \frac{\epsilon}{2} L\left(\frac{x_{i+1}-x_i}{\epsilon}, x_{i+1}\right) + \frac{\epsilon}{2} L\left(\frac{x_{i+1}-x_i}{\epsilon}, x_i\right) \quad (19)$$

or, if it proves more convenient,

$$S(x_{i+1}, x_i) = \epsilon L\left(\frac{x_{i+1}-x_i}{\epsilon}, \frac{x_{i+1}+x_i}{2}\right). \quad (20)$$

These are not valid in a general coordinate system, e.g., spherical. An even simpler approximation may be used if, in addition, there is no vector potential or other terms linear in the velocity (see page 376):

$$S(x_{i+1}, x_i) = \epsilon L\left(\frac{x_{i+1}-x_i}{\epsilon}, x_{i+1}\right). \quad (21)$$

Thus, for the simple example of a particle of mass m moving in one dimension under a potential $V(x)$, we can set

$$S(x_{i+1}, x_i) = \frac{m\epsilon}{2}\left(\frac{x_{i+1}-x_i}{\epsilon}\right)^2 - \epsilon V(x_{i+1}). \quad (22)$$

[11] It is assumed that the "forces" enter through a scalar and vector potential and not in terms involving the square of the velocity. More generaliy, what is meant by a free particle is one for which the Lagrangian is altered by omission of the terms linear in, and those independent of, the velocities.

[12] More generally, coordinates for which the terms quadratic in the velocity in $L(\dot{x}, x)$ appear with constant coefficients.

For this example, then, Eq. (18) becomes

$$\psi(x_{k+1}, t+\epsilon) = \int \exp\left[\frac{i\epsilon}{\hbar}\left\{\frac{m}{2}\left(\frac{x_{k+1}-x_k}{\epsilon}\right)^2 - V(x_{k+1})\right\}\right]\psi(x_k, t)dx_k/A. \quad (23)$$

Let us call $x_{k+1}=x$ and $x_{k+1}-x_k=\xi$ so that $x_k=x-\xi$. Then (23) becomes

$$\psi(x, t+\epsilon) = \int \exp\frac{im\xi^2}{\epsilon\cdot 2\hbar} \cdot\exp\frac{-i\epsilon V(x)}{\hbar}\cdot\psi(x-\xi, t)\frac{d\xi}{A}. \quad (24)$$

The integral on ξ will converge if $\psi(x, t)$ falls off sufficiently for large x (certainly if $\int\psi^*(x)\psi(x)dx=1$). In the integration on ξ, since ϵ is very small, the exponential of $im\xi^2/2\hbar\epsilon$ oscillates extremely rapidly except in the region about $\xi=0$ (ξ of order $(\hbar\epsilon/m)^{\frac{1}{2}}$). Since the function $\psi(x-\xi, t)$ is a relatively smooth function of ξ (since ϵ may be taken as small as desired), the region where the exponential oscillates rapidly will contribute very little because of the almost complete cancelation of positive and negative contributions. Since only small ξ are effective, $\psi(x-\xi, t)$ may be expanded as a Taylor series. Hence,

$$\psi(x, t+\epsilon) = \exp\left(\frac{-i\epsilon V(x)}{\hbar}\right) \times\int \exp\left(\frac{im\xi^2}{2\hbar\epsilon}\right)\left[\psi(x, t) - \xi\frac{\partial\psi(x, t)}{\partial x} + \frac{\xi^2}{2}\frac{\partial^2\psi(x, t)}{\partial x^2} - \cdots\right]d\xi/A. \quad (25)$$

Now

$$\int_{-\infty}^{\infty}\exp(im\xi^2/2\hbar\epsilon)d\xi = (2\pi\hbar\epsilon i/m)^{\frac{1}{2}},$$

$$\int_{-\infty}^{\infty}\exp(im\xi^2/2\hbar\epsilon)\xi d\xi = 0, \quad (26)$$

$$\int_{-\infty}^{\infty}\exp(im\xi^2/2\hbar\epsilon)\xi^2 d\xi = (\hbar\epsilon i/m)(2\pi\hbar\epsilon i/m)^{\frac{1}{2}},$$

while the integral containing ξ^3 is zero, for like

the one with ξ it possesses an odd integrand, and the ones with ξ^4 are of at least the order ϵ smaller than the ones kept here.[13] If we expand the left-hand side to first order in ϵ, (25) becomes

$$\psi(x,t)+\epsilon\frac{\partial\psi(x,t)}{\partial t}$$
$$=\exp\left(\frac{-i\epsilon V(x)}{\hbar}\right)\frac{(2\pi\hbar\epsilon i/m)^{\frac{1}{2}}}{A}$$
$$\times\left[\psi(x,t)+\frac{\hbar\epsilon i}{m}\frac{\partial^2\psi(x,t)}{\partial x^2}+\cdots\right]. \quad (27)$$

In order that both sides may agree to *zero* order in ϵ, we must set

$$A=(2\pi\hbar\epsilon i/m)^{\frac{1}{2}}. \quad (28)$$

Then expanding the exponential containing $V(x)$, we get

$$\psi(x,t)+\epsilon\frac{\partial\psi}{\partial t}=\left(1-\frac{i\epsilon}{\hbar}V(x)\right)$$
$$\times\left(\psi(x,t)+\frac{\hbar\epsilon i}{2m}\frac{\partial^2\psi}{\partial x^2}\right). \quad (29)$$

Canceling $\psi(x,t)$ from both sides, and comparing terms to first order in ϵ and multiplying by $-\hbar/i$ one obtains

$$-\frac{\hbar}{i}\frac{\partial\psi}{\partial t}=\frac{1}{2m}\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2\psi+V(x)\psi, \quad (30)$$

which is Schroedinger's equation for the problem in question.

The equation for χ^* can be developed in the same way, but adding a factor *decreases* the time by one step, i.e., χ^* satisfies an equation like (30) but with the sign of the time reversed. By taking complex conjugates we can conclude that χ satisfies the same equation as ψ, i.e., an experiment can be defined by the particular state χ to which it corresponds.[14]

This example shows that most of the contribution to $\psi(x_{k+1},t+\epsilon)$ comes from values of x_k in $\psi(x_k,t)$ which are quite close to x_{k+1} (distant of order $\epsilon^{\frac{1}{2}}$) so that the integral equation (23) can, in the limit, be replaced by a differential equation. The "velocities," $(x_{k+1}-x_k)/\epsilon$ which are important are very high, being of order $(\hbar/m\epsilon)^{\frac{1}{2}}$ which diverges as $\epsilon\to 0$. The paths involved are, therefore, continuous but possess no derivative. They are of a type familiar from study of Brownian motion.

It is these large velocities which make it so necessary to be careful in approximating $S(x_{k+1},x_k)$ from Eq. (11).[15] To replace $V(x_{k+1})$ by $V(x_k)$ would, of course, change the exponent in (18) by $i\epsilon[V(x_k)-V(x_{k+1})]/\hbar$ which is of order $\epsilon(x_{k+1}-x_k)$, and thus lead to unimportant terms of higher order than ϵ on the right-hand side of (29). It is for this reason that (20) and (21) are equally satisfactory approximations to $S(x_{i+1},x_i)$ when there is no vector potential. A term, linear in velocity, however, arising from a vector potential, as $A\dot{x}dt$ must be handled more carefully. Here a term in $S(x_{k+1},x_k)$ such as $A(x_{k+1})\times(x_{k+1}-x_k)$ differs from $A(x_k)(x_{k+1}-x_k)$ by a term of order $(x_{k+1}-x_k)^2$, and, therefore, of order ϵ. Such a term would lead to a change in the resulting wave equation. For this reason the approximation (21) is not a sufficiently accurate approximation to (11) and one like (20), (or (19) from which (20) differs by terms of order higher than ϵ) must be used. If $\mathbf{A}$ represents the vector potential and $\mathbf{p}=(\hbar/i)\nabla$, the momentum operator, then (20) gives, in the Hamiltonian operator, a term $(1/2m)(\mathbf{p}-(e/c)\mathbf{A})\cdot(\mathbf{p}-(e/c)\mathbf{A})$, while (21) gives $(1/2m)(\mathbf{p}\cdot\mathbf{p}-(2e/c)\mathbf{A}\cdot\mathbf{p}+(e^2/c^2)\mathbf{A}\cdot\mathbf{A})$. These two expressions differ by $(\hbar e/2imc)\nabla\cdot\mathbf{A}$

[13] Really, these integrals are oscillatory and not defined, but they may be defined by using a convergence factor. Such a factor is automatically provided by $\psi(x-\xi,t)$ in (24). If a more formal procedure is desired replace $\hbar$ by $\hbar(1-i\delta)$, for example, where δ is a small positive number, and then let $\delta\to 0$.

[14] Dr. Hartland Snyder has pointed out to me, in private conversation, the very interesting possibility that there may be a generalization of quantum mechanics in which the states measured by experiment cannot be prepared; that is, there would be no state into which a system may be put for which a particular experiment gives certainty for a result. The class of functions χ is not identical to the class of available states ψ. This would result if, for example, χ satisfied a different equation than ψ.

[15] Equation (18) is actually exact when (11) is used for $S(x_{i+1},x_i)$ for arbitrary ϵ for cases in which the potential does not involve x to higher powers than the second (e.g., free particle, harmonic oscillator). It is necessary, however, to use a more accurate value of A. One can define A in this way. Assume classical particles with k degrees of freedom start from the point x_i, t_i with uniform density in momentum space. Write the number of particles having a given component of momentum in range dp as dp/p_0 with p_0 constant. Then $A=(2\pi\hbar i/p_0)^{k/2}\rho^{-\frac{1}{2}}$, where ρ is the density in k dimensional coordinate space x_{i+1} of these particles at time t_{i+1}.

which may not be zero. The question is still more important in the coefficient of terms which are quadratic in the velocities. In these terms (19) and (20) are not sufficiently accurate representations of (11) in general. It is when the coefficients are constant that (19) or (20) can be substituted for (11). If an expression such as (19) is used, say for spherical coordinates, when it is not a valid approximation to (11), one obtains a Schroedinger equation in which the Hamiltonian operator has some of the momentum operators and coordinates in the wrong order. Equation (11) then resolves the ambiguity in the usual rule to replace p and q by the non-commuting quantities $(\hbar/i)(\partial/\partial q)$ and q in the classical Hamiltonian $H(p, q)$.

It is clear that the statement (11) is independent of the coordinate system. Therefore, to find the differential wave equation it gives in any coordinate system, the easiest procedure is first to find the equations in Cartesian coordinates and then to transform the coordinate system to the one desired. It suffices, therefore, to show the relation of the postulates and Schroedinger's equation in rectangular coordinates.

The derivation given here for one dimension can be extended directly to the case of three-dimensional Cartesian coordinates for any number, K, of particles interacting through potentials with one another, and in a magnetic field, described by a vector potential. The terms in the vector potential require completing the square in the exponent in the usual way for Gaussian integrals. The variable x must be replaced by the set $x^{(1)}$ to $x^{(3K)}$ where $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ are the coordinates of the first particle of mass m_1, $x^{(4)}$, $x^{(5)}$, $x^{(6)}$ of the second of mass m_2, etc. The symbol dx is replaced by $dx^{(1)}dx^{(2)}\cdots dx^{(3K)}$, and the integration over dx is replaced by a $3K$-fold integral. The constant A has, in this case, the value $A=(2\pi\hbar\epsilon i/m_1)^{\frac{1}{2}}(2\pi\hbar\epsilon i/m_2)^{\frac{1}{2}}\cdots(2\pi\hbar\epsilon i/m_K)^{\frac{1}{2}}$. The Lagrangian is the classical Lagrangian for the same problem, and the Schroedinger equation resulting will be that which corresponds to the classical Hamiltonian, derived from this Lagrangian. The equations in any other coordinate system may be obtained by transformation. Since this includes all cases for which Schroedinger's equation has been checked with experiment, we may say our postulates are able to describe what can be described by non-relativistic quantum mechanics, neglecting spin.

7. DISCUSSION OF THE WAVE EQUATION

The Classical Limit

This completes the demonstration of the equivalence of the new and old formulations. We should like to include in this section a few remarks about the important equation (18).

This equation gives the development of the wave function during a small time interval. It is easily interpreted physically as the expression of Huygens' principle for matter waves.* In geometrical optics the rays in an inhomogeneous medium satisfy Fermat's principle of least *time*. We may state Huygens' principle in wave optics in this way: If the amplitude of the wave is known on a given surface, the amplitude at a near by point can be considered as a sum of contributions from all points of the surface. Each contribution is delayed in phase by an amount proportional to the *time* it would take the light to get from the surface to the point along the ray of least *time* of geometrical optics. We can consider (22) in an analogous manner starting with Hamilton's first principle of least *action* for classical or "geometrical" mechanics. If the amplitude of the wave ψ is known on a given "surface," in particular the "surface" consisting of all x at time t, its value at a particular nearby point at time $t+\epsilon$, is a sum of contributions from all points of the surface at t. Each contribution is delayed in phase by an amount proportional to the *action* it would require to get from the surface to the point along the path of least *action* of classical mechanics.[16]

Actually Huygens' principle is not correct in optics. It is replaced by Kirchoff's modification which requires that both the amplitude and its derivative must be known on the adjacent surface. This is a consequence of the fact that the wave equation in optics is second order in the time. The wave equation of quantum mechanics is first order in the time; therefore, Huygens' principle *is* correct for matter waves, action replacing time.

[16] See in this connection the very interesting remarks of Schroedinger, Ann. d. Physik **79**, 489 (1926).

The equation can also be compared mathematically to quantities appearing in the usual formulations. In Schroedinger's method the development of the wave function with time is given by

$$\frac{\hbar}{i}\frac{\partial\psi}{\partial t}=\mathbf{H}\psi, \tag{31}$$

which has the solution (for any ϵ if $\mathbf{H}$ is time independent)

$$\psi(x,\, t+\epsilon)=\exp(-i\epsilon\mathbf{H}/\hbar)\psi(x,\, t). \tag{32}$$

Therefore, Eq. (18) expresses the operator $\exp(-i\epsilon\mathbf{H}/\hbar)$ by an approximate integral operator for small ϵ.

From the point of view of Heisenberg one considers the position at time t, for example, as an operator $\mathbf{x}$. The position $\mathbf{x}'$ at a later time $t+\epsilon$ can be expressed in terms of that at time t by the operator equation

$$\mathbf{x}'=\exp(i\epsilon\mathbf{H}/\hbar)\mathbf{x}\exp-(i\epsilon\mathbf{H}/\hbar). \tag{33}$$

The transformation theory of Dirac allows us to consider the wave function at time $t+\epsilon$, $\psi(x', t+\epsilon)$, as representing a state in a representation in which $\mathbf{x}'$ is diagonal, while $\psi(x,\, t)$ represents the same state in a representation in which $\mathbf{x}$ is diagonal. They are, therefore, related through the transformation function $(x'|x)_\epsilon$ which relates these representations:

$$\psi(x',\, t+\epsilon)=\int(x'|x)_\epsilon\psi(x,t)\, dx.$$

Therefore, the content of Eq. (18) is to show that for small ϵ we can set

$$(x'|x)_\epsilon=(1/A)\exp(iS(x',\, x)/\hbar) \tag{34}$$

with $S(x',\, x)$ defined as in (11).

The close analogy between $(x'|x)_\epsilon$ and the quantity $\exp(iS(x',\, x)/\hbar)$ has been pointed out on several occasions by Dirac.[1] In fact, we now see that to sufficient approximations the two quantities may be taken to be proportional to each other. Dirac's remarks were the starting point of the present development. The points he makes concerning the passage to the classical limit $\hbar\to 0$ are very beautiful, and I may perhaps be excused for briefly reviewing them here.

First we note that the wave function at x'' at time t'' can be obtained from that at x' at time t' by

$$\psi(x'',\, t'')=\lim_{\epsilon\to 0}\int\cdots\int \times\exp\left[\frac{i}{\hbar}\sum_{i=0}^{j-1}S(x_{i+1},\, x_i)\right] \times\psi(x',\, t')\frac{dx_0}{A}\frac{dx_1}{A}\cdots\frac{dx_{j-1}}{A}, \tag{35}$$

where we put $x_0\equiv x'$ and $x_j\equiv x''$ where $j\epsilon=t''-t'$ (between the times t' and t'' we assume no restriction is being put on the region of integration). This can be seen either by repeated applications of (18) or directly from Eq. (15). Now we ask, as $\hbar\to 0$ what values of the intermediate coordinates x_i contribute most strongly to the integral? These will be the values most likely to be found by experiment and therefore will determine, in the limit, the classical path. If $\hbar$ is very small, the exponent will be a very rapidly varying function of any of its variables x_i. As x_i varies, the positive and negative contributions of the exponent nearly cancel. The region at which x_i contributes most strongly is that at which the phase of the exponent varies least rapidly with x_i (method of stationary phase). Call the sum in the exponent S;

$$S=\sum_{i=0}^{j-1}S(x_{i+1},\, x_i). \tag{36}$$

Then the classical orbit passes, approximately, through those points x_i at which the rate of change of S with x_i is small, or in the limit of small $\hbar$, zero, i.e., the classical orbit passes through the points at which $\partial S/\partial x_i=0$ for all x_i. Taking the limit $\epsilon\to 0$, (36) becomes in view of (11)

$$S=\int_{t'}^{t''}L(\dot{x}(t),\, x(t))dt. \tag{37}$$

We see then that the classical path is that for which the integral (37) suffers no first-order change on varying the path. This is Hamilton's principle and leads directly to the Lagrangian equations of motion.

8. OPERATOR ALGEBRA

Matrix Elements

Given the wave function and Schroedinger's equation, of course all of the machinery of operator or matrix algebra can be developed. It is, however, rather interesting to express these concepts in a somewhat different language more closely related to that used in stating the postulates. Little will be gained by this in elucidating operator algebra. In fact, the results are simply a translation of simple operator equations into a somewhat more cumbersome notation. On the other hand, the new notation and point of view are very useful in certain applications described in the introduction. Furthermore, the form of the equations permits natural extension to a wider class of operators than is usually considered (e.g., ones involving quantities referring to two or more different times). If any generalization to a wider class of action functionals is possible, the formulae to be developed will play an important role.

We discuss these points in the next three sections. This section is concerned mainly with definitions. We shall define a quantity which we call a transition element between two states. It is essentially a matrix element. But instead of being the matrix element between a state ψ and another χ corresponding to the *same* time, these two states will refer to different times. In the following section a fundamental relation between transition elements will be developed from which the usual commutation rules between coordinate and momentum may be deduced. The same relation also yields Newton's equation of motion in matrix form. Finally, in Section 10 we discuss the relation of the Hamiltonian to the operation of displacement in time.

We begin by defining a transition element in terms of the probability of transition from one state to another. More precisely, suppose we have a situation similar to that described in deriving (17). The region R consists of a region R' previous to t', all space between t' and t'' and the region R'' after t''. We shall study the probability that a system in region R' is later found in region R''. This is given by (17). We shall discuss in this section how it changes with changes in the form of the Lagrangian between t' and t''. In Section 10 we discuss how it changes with changes in the preparation R' or the experiment R''.

The state at time t' is defined completely by the preparation R'. It can be specified by a wave function $\psi(x', t')$ obtained as in (15), but containing only integrals up to the time t'. Likewise, the state characteristic of the experiment (region R'') can be defined by a function $\chi(x'', t'')$ obtained from (16) with integrals only beyond t''. The wave function $\psi(x'', t'')$ at time t'' can, of course, also be gotten by appropriate use of (15). It can also be gotten from $\psi(x', t')$ by (35). According to (17) with t'' used instead of t, the probability of being found in χ if prepared in ψ is the square of what we shall call the transition amplitude $\int \chi^*(x'', t'')\psi(x'', t'')dx''$. We wish to express this in terms of χ at t'' and ψ at t'. This we can do with the aid of (35). Thus, the chance that a system prepared in state $\psi_{t'}$ at time t' will be found after t'' to be in a state $\chi_{t''}$ is the square of the transition amplitude

$$\langle \chi_{t''} | 1 | \psi_{t'} \rangle_S = \operatorname*{Lim}_{\epsilon \to 0} \int \cdots \int \chi^*(x'', t'') \times \exp(iS/\hbar)\psi(x', t') \frac{dx_0}{A} \cdots \frac{dx_{j-1}}{A} dx_j, \quad (38)$$

where we have used the abbreviation (36).

In the language of ordinary quantum mechanics if the Hamiltonian, $\mathbf{H}$, is constant, $\psi(x, t'') = \exp[-i(t''-t')\mathbf{H}/\hbar]\psi(x, t')$ so that (38) is the matrix element of $\exp[-i(t''-t')\mathbf{H}/\hbar]$ between states $\chi_{t''}$ and $\psi_{t'}$.

If F is any function of the coordinates x_i for $t' < t_i < t''$, we shall define the transition element of F between the states ψ at t' and χ at t'' for the action S as $(x'' \equiv x_j, x' \equiv x_0)$:

$$\langle \chi_{t''} | F | \psi_{t'} \rangle_S = \operatorname*{Lim}_{\epsilon \to 0} \int \cdots \int \times \chi^*(x'', t'') F(x_0, x_1, \cdots x_j) \cdot \exp\left[\frac{i}{\hbar} \sum_{i=0}^{j-1} S(x_{i+1}, x_i)\right] \psi(x', t') \frac{dx_0}{A} \cdots \frac{dx_{j-1}}{A} dx_j. \quad (39)$$

In the limit $\epsilon \to 0$, F is a functional of the path $x(t)$.

We shall see presently why such quantities are important. It will be easier to understand if we

stop for a moment to find out what the quantities correspond to in conventional notation. Suppose F is simply x_k where k corresponds to some time $t=t_k$. Then on the right-hand side of (39) the integrals from x_0 to x_{k-1} may be performed to produce $\psi(x_k, t)$ or $\exp[-i(t-t')\mathbf{H}/\hbar]\psi_{t'}$. In like manner the integrals on x_i for $j \geq i > k$ give $\chi^*(x_k, t)$ or $\{\exp[-i(t''-t)\mathbf{H}/\hbar]\chi_{t''}\}^*$. Thus, the transition element of x_k,

$$\langle \chi_{t''} | F | \psi_{t'} \rangle_S = \int \chi_{t''}^* e^{-(i/\hbar)\mathbf{H}(t''-t)} x e^{-(i/\hbar)\mathbf{H}(t-t')} \psi_{t'} dx = \int \chi^*(x, t) x \psi(x, t) dx \quad (40)$$

is the matrix element of $\mathbf{x}$ at time $t=t_k$ between the state which would develop at time t from $\psi_{t'}$ at t' and the state which will develop from time t to $\chi_{t''}$ at t''. It is, therefore, the matrix element of $\mathbf{x}(t)$ between these states.

Likewise, according to (39) with $F=x_{k+1}$, the transition element of x_{k+1} is the matrix element of $\mathbf{x}(t+\epsilon)$. The transition element of $F=(x_{k+1}-x_k)/\epsilon$ is the matrix element of $(\mathbf{x}(t+\epsilon)-\mathbf{x}(t))/\epsilon$ or of $i(\mathbf{Hx}-\mathbf{xH})/\hbar$, as is easily shown from (40). We can call this the matrix element of velocity $x(t)$.

Suppose we consider a second problem which differs from the first because, for example, the potential is augmented by a small amount $U(, \mathbf{x}t)$. Then in the new problem the quantity replacing S is $S'=S+\sum_i \epsilon U(x_i, t_i)$. Substitution into (38) leads directly to

$$\langle \chi_{t''} | 1 | \psi_{t'} \rangle_{S'} = \left\langle \chi_{t''} \middle| \exp \frac{i\epsilon}{\hbar} \sum_{i=1}^{j} U(x_i, t_i) \middle| \psi_{t'} \right\rangle_S . \quad (41)$$

Thus, transition elements such as (39) are important insofar as F may arise in some way from a change δS in an action expression. We denote, by observable functionals, those functionals F which can be defined, (possibly indirectly) in terms of the changes which are produced by possible changes in the action S. The condition that a functional be observable is somewhat similar to the condition that an operator be Hermitian. The observable functionals are a restricted class because the action must remain a quadratic function of velocities. From one observable functional others may be derived, for example, by

$$\langle \chi_{t''} | F | \psi_{t'} \rangle_{S'} = \left\langle \chi_{t''} \middle| F \exp \frac{i\epsilon}{\hbar} \sum_{i=1}^{j} U(x_i, t_i) \middle| \psi_{t'} \right\rangle_S \quad (42)$$

which is obtained from (39).

Incidentally, (41) leads directly to an important perturbation formula. If the effect of U is small the exponential can be expanded to first order in U and we find

$$\langle \chi_{t''} | 1 | \psi_{t'} \rangle_{S'} = \langle \chi_{t''} | 1 | \psi_{t'} \rangle_S + \frac{i}{\hbar} \langle \chi_{t''} | \sum_i \epsilon U(x_i, t_i) | \psi_{t'} \rangle . \quad (43)$$

Of particular importance is the case that $\chi_{t''}$ is a state in which $\psi_{t'}$ would not be found at all were it not for the disturbance, U (i.e., $\langle \chi_{t''} | 1 | \psi_{t'} \rangle_S = 0$) Then

$$\frac{1}{\hbar^2} | \langle \chi_{t''} | \sum_i \epsilon U(x_i, t_i) | \psi_{t'} \rangle_S |^2 \quad (44)$$

is the probability of transition as induced to first order by the perturbation. In ordinary notation,

$$\langle \chi_{t''} | \sum_i \epsilon U(x_i, t_i) | \psi_{t'} \rangle_S = \int \left\{ \int \chi_{t''}^* e^{-(i/\hbar)\mathbf{H}(t''-t)} \mathbf{U} e^{-(i/\hbar)\mathbf{H}(t-t')} \psi_{t'} dx \right\} dt$$

so that (44) reduces to the usual expression[17] for time dependent perturbations.

9. NEWTON'S EQUATIONS

The Commutation Relation

In this section we find that different functionals may give identical results when taken between any two states. This equivalence between functionals is the statement of operator equations in the new language.

If F depends on the various coordinates, we can, of course, define a new functional $\partial F/\partial x_k$

[17] P. A. M. Dirac, *The Principles of Quantum Mechanics* (The Clarendon Press, Oxford, 1935), second edition, Section 47, Eq. (20).

by differentiating it with respect to one of its variables, say $x_k (0<k<j)$. If we calculate $\langle \chi_{t''} | \partial F/\partial x_k | \psi_{t'} \rangle_S$ by (39) the integral on the right-hand side will contain $\partial F/\partial x_k$. The only other place that the variable x_k appears is in S. Thus, the integration on x_k can be performed by parts. The integrated part vanishes (assuming wave functions vanish at infinity) and we are left with the quantity $-F(\partial/\partial x_k)\exp(iS/\hbar)$ in the integral. However, $(\partial/\partial x_k)\exp(iS/\hbar) = (i/\hbar)(\partial S/\partial x_k)\exp(iS/\hbar)$, so the right side represents the transition element of $-(i/\hbar)F(\partial S/\partial x_k)$, i.e.,

$$\left\langle \chi_{t''} \left| \frac{\partial F}{\partial x_k} \right| \psi_{t'} \right\rangle_S = -\frac{i}{\hbar} \left\langle \chi_{t''} \left| F \frac{\partial S}{\partial x_k} \right| \psi_{t'} \right\rangle_S . \quad (45)$$

This very important relation shows that two different functionals may give the same result for the transition element between any two states. We say they are equivalent and symbolize the relation by

$$-\frac{\hbar}{i}\frac{\partial F}{\partial x_k} \underset{S}{\leftrightarrow} F\frac{\partial S}{\partial x_k}, \quad (46)$$

the symbol $\underset{S}{\leftrightarrow}$ emphasizing the fact that functionals equivalent under one action may not be equivalent under another. The quantities in (46) need not be observable. The equivalence is, nevertheless, true. Making use of (36) one can write

$$-\frac{\hbar}{i}\frac{\partial F}{\partial x_k} \underset{S}{\leftrightarrow} F\left[\frac{\partial S(x_{k+1}, x_k)}{\partial x_k} + \frac{\partial S(x_k, x_{k-1})}{\partial x_k}\right]. \quad (47)$$

This equation is true to zero and first order in ϵ and has as consequences the commutation relations of momentum and coordinate, as well as the Newtonian equations of motion in matrix form.

In the case of our simple one-dimensional problem, $S(x_{i+1}, x_i)$ is given by the expression (15), so that

$$\partial S(x_{k+1}, x_k)/\partial x_k = -m(x_{k+1}-x_k)/\epsilon,$$

and

$$\partial S(x_k, x_{k-1})/\partial x_k = +m(x_k - x_{k-1})/\epsilon - \epsilon V'(x_k);$$

where we write $V'(x)$ for the derivative of the potential, or force. Then (47) becomes

$$-\frac{\hbar}{i}\frac{\partial F}{\partial x_k} \underset{S}{\leftrightarrow} F\left[-m\left(\frac{x_{k+1}-x_k}{\epsilon} - \frac{x_k - x_{k-1}}{\epsilon}\right) - \epsilon V'(x_k)\right]. \quad (48)$$

If F does not depend on the variable x_k, this gives Newton's equations of motion. For example, if F is constant, say unity, (48) just gives (dividing by ϵ)

$$0 \underset{S}{\leftrightarrow} -\frac{m}{\epsilon}\left(\frac{x_{k+1}-x_k}{\epsilon} - \frac{x_k - x_{k-1}}{\epsilon}\right) - V'(x_k).$$

Thus, the transition element of mass times acceleration $[(x_{k+1}-x_k)/\epsilon - (x_k - x_{k-1})/\epsilon]/\epsilon$ between any two states is equal to the transition element of force $-V'(x_k)$ between the same states. This is the matrix expression of Newton's law which holds in quantum mechanics.

What happens if F does depend upon x_k? For example, let $F = x_k$. Then (48) gives, since $\partial F/\partial x_k = 1$,

$$-\frac{\hbar}{i} \underset{S}{\leftrightarrow} x_k\left[-m\left(\frac{x_{k+1}-x_k}{\epsilon} - \frac{x_k - x_{k-1}}{\epsilon}\right) - \epsilon V'(x_k)\right]$$

or, neglecting terms of order ϵ,

$$m\left(\frac{x_{k+1}-x_k}{\epsilon}\right)x_k - m\left(\frac{x_k - x_{k-1}}{\epsilon}\right)x_k \underset{S}{\leftrightarrow} \frac{\hbar}{i}. \quad (49)$$

In order to transfer an equation such as (49) into conventional notation, we shall have to discover what matrix corresponds to a quantity such as $x_k x_{k+1}$. It is clear from a study of (39) that if F is set equal to, say, $f(x_k)g(x_{k+1})$, the corresponding operator in (40) is

$$e^{-(i/\hbar)(t''-t-\epsilon)\mathbf{H}} g(\mathbf{x}) e^{-(i/\hbar)\epsilon\mathbf{H}} f(\mathbf{x}) e^{-(i/\hbar)(t-t')\mathbf{H}},$$

the matrix element being taken between the states $\chi_{t''}$ and $\psi_{t'}$. The operators corresponding to functions of x_{k+1} will appear to the left of the operators corresponding to functions of x_k, i.e., *the order of terms in a matrix operator product corresponds to an order in time of the corresponding factors in a functional.* Thus, if the functional can and is written in such a way that in each term factors corresponding to later times appear to the

left of factors corresponding to earlier terms, the corresponding operator can immediately be written down if the order of the operators is kept the same as in the functional.[18] Obviously, the order of factors in a functional is of no consequence. The ordering just facilitates translation into conventional operator notation. To write Eq. (49) in the way desired for easy translation would require the factors in the second term on the left to be reversed in order. We see, therefore, that it corresponds to

$$\mathbf{px}-\mathbf{xp}=\hbar/i$$

where we have written $\mathbf{p}$ for the operator $m\mathbf{x}$.

The relation between functionals and the corresponding operators is defined above in terms of the order of the factors in time. It should be remarked that this rule must be especially carefully adhered to when quantities involving velocities or higher derivatives are involved. The correct functional to represent the operator $(\dot{x})^2$ is actually $(x_{k+1}-x_k)/\epsilon\cdot(x_k-x_{k-1})/\epsilon$ rather than $[(x_{k+1}-x_k)/\epsilon]^2$. The latter quantity diverges as $1/\epsilon$ as $\epsilon\to 0$. This may be seen by replacing the second term in (49) by its value $x_{k+1}\cdot m(x_{k+1}-x_k)/\epsilon$ calculated an instant ϵ later in time. This does not change the equation to zero order in ϵ. We then obtain (dividing by ϵ)

$$\left(\frac{x_{k+1}-x_k}{\epsilon}\right)^2 \underset{S}{\leftrightarrow} -\frac{\hbar}{im\epsilon}. \tag{50}$$

This gives the result expressed earlier that the root mean square of the "velocity" $(x_{k+1}-x_k)/\epsilon$ between two successive positions of the path is of order $\epsilon^{-\frac{1}{2}}$.

It will not do then to write the functional for kinetic energy, say, simply as

$$\tfrac{1}{2}m[(x_{k+1}-x_k)/\epsilon]^2 \tag{51}$$

for this quantity is infinite as $\epsilon\to 0$. In fact, it is not an observable functional.

One can obtain the kinetic energy as an observable functional by considering the first-order change in transition amplitude occasioned by a change in the mass of the particle. Let m be changed to $m(1+\delta)$ for a short time, say ϵ, around t_k. The change in the action is $\frac{1}{2}\delta\epsilon m[(x_{k+1}-x_k)/\epsilon]^2$ the derivative of which gives an expression like (51). But the change in m changes the normalization constant $1/A$ corresponding to dx_k as well as the action. The constant is changed from $(2\pi\hbar\epsilon i/m)^{-\frac{1}{2}}$ to $(2\pi\hbar\epsilon i/m(1+\delta))^{-\frac{1}{2}}$ or by $\frac{1}{2}\delta(2\pi\hbar\epsilon i/m)^{-\frac{1}{2}}$ to first order in δ. The total effect of the change in mass in Eq. (38) to the first order in δ is

$$\langle\chi_{t''}|\tfrac{1}{2}\delta\epsilon im[(x_{k+1}-x_k)/\epsilon]^2/\hbar+\tfrac{1}{2}\delta|\psi_{t'}\rangle.$$

We expect the change of order δ lasting for a time ϵ to be of order $\delta\epsilon$. Hence, dividing by $\delta\epsilon i/\hbar$, we can define the kinetic energy functional as

$$\text{K.E.}=\tfrac{1}{2}m[(x_{k+1}-x_k)/\epsilon]^2+\hbar/2\epsilon i. \tag{52}$$

This is finite as $\epsilon\to 0$ in view of (50). By making use of an equation which results from substituting $m(x_{k+1}-x_k)/\epsilon$ for F in (48) we can also show that the expression (52) is equal (to order ϵ) to

$$\text{K.E.}=\tfrac{1}{2}m\left(\frac{x_{k+1}-x_k}{\epsilon}\right)\left(\frac{x_k-x_{k-1}}{\epsilon}\right). \tag{53}$$

That is, the easiest way to produce observable functionals involving powers of the velocities is to replace these powers by a product of velocities, each factor of which is taken at a slightly different time.

10. THE HAMILTONIAN

Momentum

The Hamiltonian operator is of central importance in the usual formulation of quantum mechanics. We shall study in this section the functional corresponding to this operator. We could immediately define the Hamiltonian functional by adding the kinetic energy functional (52) or (53) to the potential energy. This method is artificial and does not exhibit the important relationship of the Hamiltonian to time. We shall define the Hamiltonian functional by the changes made in a state when it is displaced in time.

To do this we shall have to digress a moment to point out that the subdivision of time into *equal* intervals is not necessary. Clearly, any subdivision into instants t_i will be satisfactory; the limits are to be taken as the largest spacing, $t_{i+1}-t_i$, approaches zero. The total action S must

[18] Dirac has also studied operators containing quantities referring to different times. See reference 2.

now be represented as a sum

$$S=\sum_i S(x_{i+1}, t_{i+1}; x_i, t_i), \tag{54}$$

where

$$S(x_{i+1}, t_{i+1}; x_i, t_i)=\int_{t_i}^{t_{i+1}} L(\dot{x}(t), x(t))dt, \tag{55}$$

the integral being taken along the classical path between x_i at t_i and x_{i+1} at t_{i+1}. For the simple one-dimensional example this becomes, with sufficient accuracy,

$$S(x_{i+1}, t_{i+1}; x_i, t_i)$$
$$=\left\{\frac{m}{2}\left(\frac{x_{i+1}-x_i}{t_{i+1}-t_i}\right)^2-V(x_{i+1})\right\}(t_{i+1}-t_i); \tag{56}$$

the corresponding normalization constant for integration on dx_i is $A=(2\pi\hbar i(t_{i+1}-t_i)/m)^{-\frac{1}{2}}$.

The relation of H to the change in a state with displacement in time can now be studied. Consider a state $\psi(t)$ defined by a space-time region R'. Now imagine that we consider another state at time t, $\psi_\delta(t)$, defined by another region R_δ'. Suppose the region R_δ' is exactly the same as R' except that it is earlier by a time δ, i.e., displaced bodily toward the past by a time δ. All the apparatus to prepare the system for R_δ' is identical to that for R' but is operated a time δ sooner. If L depends explicitly on time, it, too, is to be displaced, i.e., the state ψ_δ is obtained from the L used for state ψ except that the time t in L_δ is replaced by $t+\delta$. We ask how does the state ψ_δ differ from ψ? In any measurement the chance of finding the system in a fixed region R'' is different for R' and R_δ'. Consider the change in the transition element $\langle\chi|1|\psi_\delta\rangle_{S_\delta}$ produced by the shift δ. We can consider this shift as effected by decreasing all values of t_i by δ for $i\leqslant k$ and leaving all t_i fixed for $i>k$, where the time t lies in the interval between t_{k+1} and t_k.[19] This change will have no effect on $S(x_{i+1}, t_{i+1}; x_i, t_i)$ as defined by (55) as long as both t_{i+1} and t_i are changed by the same amount. On the other hand, $S(x_{k+1}, t_{k+1}; x_k, t_k)$ is changed to $S(x_{k+1}, t_{k+1}; x_k, t_k-\delta)$. The constant $1/A$ for the integration on dx_k is also altered to $(2\pi\hbar i(t_{k+1}-t_k+\delta)/m)^{-\frac{1}{2}}$. The effect of these changes on the transition element is given to the first order in δ by

$$\langle\chi|1|\psi\rangle_S-\langle\chi|1|\psi_\delta\rangle_{S_\delta}=\frac{i\delta}{\hbar}\langle\chi|H_k|\psi\rangle_S, \tag{57}$$

here the Hamiltonian functional H_k is defined by

$$H_k=\frac{\partial S(x_{k+1}, t_{k+1}; x_k, t_k)}{\partial t_k}+\frac{\hbar}{2i(t_{k+1}-t_k)}. \tag{58}$$

The last term is due to the change in $1/A$ and serves to keep H_k finite as $\epsilon\to 0$. For example, for the expression (56) this becomes

$$H_k=\frac{m}{2}\left(\frac{x_{k+1}-x_k}{t_{k+1}-t_k}\right)^2+\frac{\hbar}{2i(t_{k+1}-t_k)}+V(x_{k+1}),$$

which is just the sum of the kinetic energy functional (52) and that of the potential energy $V(x_{k+1})$.

The wave function $\psi_\delta(x, t)$ represents, of course, the same state as $\psi(x, t)$ will be after time δ, i.e., $\psi(x, t+\delta)$. Hence, (57) is intimately related to the operator equation (31).

One could also consider changes occasioned by a time shift in the final state χ. Of course, nothing new results in this way for it is only the relative shift of χ and ψ which counts. One obtains an alternative expression

$$H_k=-\frac{\partial S(x_{k+1}, t_{k+1}; x_k, t_k)}{\partial t_{k+1}}+\frac{\hbar}{2i(t_{k+1}-t_k)}. \tag{59}$$

This differs from (58) only by terms of order ϵ.

The time rate of change of a functional can be computed by considering the effect of shifting both initial and final state together. This has the same effect as calculating the transition element of the functional referring to a later time. What results is the analog of the operator equation

$$\frac{\hbar}{i}\dot{\mathbf{f}}=\mathbf{Hf}-\mathbf{fH}.$$

The momentum functional p_k can be defined in an analagous way by considering the changes

[19] From the point of view of mathematical rigor, if δ is finite, as $\epsilon\to 0$ one gets into difficulty in that, for example, the interval $t_{k+1}-t_k$ is kept finite. This can be straightened out by assuming δ to vary with time and to be turned on smoothly before $t=t_k$ and turned off smoothly after $t=t_k$. Then keeping the time variation of δ fixed, let $\epsilon\to 0$. Then seek the first-order change as $\delta\to 0$. The result is essentially the same as that of the crude procedure used above.

made by displacements of position:

$$\langle\chi|1|\psi\rangle_S-\langle\chi|1|\psi_\Delta\rangle_{S_\Delta}=\frac{i\Delta}{\hbar}\langle\chi|p_k|\psi\rangle_S.$$

The state ψ_Δ is prepared from a region R_Δ' which is identical to region R' except that it is moved a distance Δ in space. (The Lagrangian, if it depends explicitly on x, must be altered to $L_\Delta=L(\dot{x}, x-\Delta)$ for times previous to t.) One finds[20]

$$p_k=\frac{\partial S(x_{k+1}, x_k)}{\partial x_{k+1}}=-\frac{\partial S(x_{k+1}, x_k)}{\partial x_k}. \quad (60)$$

Since $\psi_\Delta(x, t)$ is equal to $\psi(x-\Delta, t)$, the close connection between p_k and the x-derivative of the wave function is established.

Angular momentum operators are related in an analogous way to rotations.

The derivative with respect to t_{i+1} of $S(x_{i+1}, t_{i+1}; x_i, t_i)$ appears in the definition of H_i. The derivative with respect to x_{i+1} defines p_i. But the derivative with respect to t_{i+1} of $S(x_{i+1}, t_{i+1}; x_i, t_i)$ is related to the derivative with respect to x_{i+1}, for the function $S(x_{i+1}, t_{i+1}; x_i, t_i)$ defined by (55) satisfies the Hamilton-Jacobi equation. Thus, the Hamilton-Jacobi equation is an equation expressing H_i in terms of the p_i. In other words, it expresses the fact that time displacements of states are related to space displacements of the same states. This idea leads directly to a derivation of the Schroedinger equation which is far more elegant than the one exhibited in deriving Eq. (30).

11. INADEQUACIES OF THE FORMULATION

The formulation given here suffers from a serious drawback. The mathematical concepts needed are new. At present, it requires an unnatural and cumbersome subdivision of the time interval to make the meaning of the equations clear. Considerable improvement can be made through the use of the notation and concepts of the mathematics of functionals. However, it was thought best to avoid this in a first presentation. One needs, in addition, an appropriate measure for the space of the argument functions $x(t)$ of the functionals.[10]

It is also incomplete from the physical standpoint. One of the most important characteristics of quantum mechanics is its invariance under unitary transformations. These correspond to the canonical transformations of classical mechanics. Of course, the present formulation, being equivalent to ordinary formulations, can be mathematically demonstrated to be invariant under these transformations. However, it has not been formulated in such a way that it is *physically* obvious that it is invariant. This incompleteness shows itself in a definite way. No direct procedure has been outlined to describe measurements of quantities other than position. Measurements of momentum, for example, of one particle, can be defined in terms of measurements of positions of other particles. The result of the analysis of such a situation does show the connection of momentum measurements to the Fourier transform of the wave function. But this is a rather roundabout method to obtain such an important physical result. It is to be expected that the postulates can be generalized by the replacement of the idea of "paths in a region of space-time R" to "paths of class R," or "paths having property R." But which properties correspond to which physical measurements has not been formulated in a general way.

12. A POSSIBLE GENERALIZATION

The formulation suggests an obvious generalization. There are interesting classical problems which satisfy a principle of least action but for which the action cannot be written as an integral of a function of positions and velocities. The action may involve accelerations, for example. Or, again, if interactions are not instantaneous, it may involve the product of coordinates at two different times, such as $\int x(t)x(t+T)dt$. The action, then, cannot be broken up into a sum of small contributions as in (10). As a consequence, no wave function is available to describe a state. Nevertheless, a transition probability can be defined for getting from a region R' into another R''. Most of the theory of the transition elements $\langle\chi_{t''}|F|\psi_{t'}\rangle_S$ can be carried over. One simply invents a symbol, such as $\langle R''|F|R'\rangle_S$ by an

[20] We did not immediately substitute p_i from (60) into (47) because (47) would then no longer have been valid to both zero order and the first order in ϵ. We could derive the commutation relations, but not the equations of motion. The two expressions in (60) represent the momenta at each end of the interval t_i to t_{i+1}. They differ by $-V'(x_{k+1})$ because of the force acting during the time ϵ.

equation such as (39) but with the expressions (19) and (20) for ψ and χ substituted, and the more general action substituted for S. Hamiltonian and momentum functionals can be defined as in section (10). Further details may be found in a thesis by the author.[21]

13. APPLICATION TO ELIMINATE FIELD OSCILLATORS

One characteristic of the present formulation is that it can give one a sort of bird's-eye view of the space-time relationships in a given situation. Before the integrations on the x_i are performed in an expression such as (39) one has a sort of format into which various F functionals may be inserted. One can study how what goes on in the quantum-mechanical system at different times is interrelated. To make these vague remarks somewhat more definite, we discuss an example.

In classical electrodynamics the fields describing, for instance, the interaction of two particles can be represented as a set of oscillators. The equations of motion of these oscillators may be solved and the oscillators essentially eliminated (Lienard and Wiechert potentials). The interactions which result involve relationships of the motion of one particle at one time, and of the other particle at another time. In quantum electrodynamics the field is again represented as a set of oscillators. But the motion of the oscillators cannot be worked out and the oscillators eliminated. It is true that the oscillators representing longitudinal waves may be eliminated. The result is instantaneous electrostatic interaction. The electrostatic elimination is very instructive as it shows up the difficulty of self-interaction very distinctly. In fact, it shows it up so clearly that there is no ambiguity in deciding what term is incorrect and should be omitted. This entire process is not relativistically invariant, nor is the omitted term. It would seem to be very desirable if the oscillators, representing transverse waves, could also be eliminated. This presents an almost insurmountable problem in the conventional quantum mechanics. We expect that the motion of a particle a at one time depends upon the motion of b at a previous time, and *vice versa*. A wave function $\psi(x_a, x_b; t)$, however, can only describe the behavior of both particles at one time. There is no way to keep track of what b did in the past in order to determine the behavior of a. The only way is to specify the state of the set of oscillators at t, which serve to "remember" what b (and a) had been doing.

The present formulation permits the solution of the motion of all the oscillators and their complete elimination from the equations describing the particles. This is easily done. One must simply solve for the motion of the oscillators before one integrates over the various variables x_i for the particles. It is the integration over x_i which tries to condense the past history into a single state function. This we wish to avoid. Of course, the result depends upon the initial and final states of the oscillator. If they are specified, the result is an equation for $\langle \chi_{t''} | 1 | \psi_{t'} \rangle$ like (38), but containing as a factor, besides $\exp(iS/\hbar)$ another functional G depending only on the coordinates describing the paths of the particles.

We illustrate briefly how this is done in a very simple case. Suppose a particle, coordinate $x(t)$, Lagrangian $L(\dot{x}, x)$ interacts with an oscillator, coordinate $q(t)$, Lagrangian $\frac{1}{2}(\dot{q}^2-\omega^2 q^2)$, through a term $\gamma(x, t)q(t)$ in the Lagrangian for the system. Here $\gamma(x, t)$ is any function of the coordinate $x(t)$ of the particle and the time.[22] Suppose we desire the probability of a transition from a state at time t', in which the particle's wave function is $\psi_{t'}$ and the oscillator is in energy level n, to a state at t'' with the particle in $\chi_{t''}$ and oscillator in level m. This is the square of

$$\langle \chi_{t''}\varphi_m | 1 | \psi_{t'}\varphi_n \rangle_{S_p+S_0+S_I} = \int\cdots\int \varphi_m^*(q_j)\chi_{t''}^*(x_j) \times \exp\frac{i}{\hbar}(S_p+S_0+S_I)\psi_{t'}(x_0)\varphi_n(q_0) \cdot \frac{dx_0}{A}\frac{dq_0}{a}\cdots\frac{dx_{j-1}}{A}\frac{dq_{j-1}}{a}dx_j dq_j. \quad (61)$$

[21] The theory of electromagnetism described by J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. 17, 157 (1945) can be expressed in a principle of least action involving the coordinates of particles alone. It was an attempt to quantize this theory, without reference to the fields, which led the author to study the formulation of quantum mechanics given here. The extension of the ideas to cover the case of more general action functions was developed in his Ph.D. thesis, "The principle of least action in quantum mechanics" submitted to Princeton University, 1942.

[22] The generalization to the case that γ depends on the velocity, $\dot{x}$, of the particle presents no problem.

Here $\varphi_n(q)$ is the wave function for the oscillator in state n, S_p is the action

$$\sum_{i=0}^{j-1} S_p(x_{i+1}, x_i)$$

calculated for the particle as though the oscillator were absent,

$$S_0 = \sum_{i=0}^{j-1}\left[\frac{\epsilon}{2}\left(\frac{q_{i+1}-q_i}{\epsilon}\right)^2 - \frac{\epsilon\omega^2}{2}q_{i+1}^2\right]$$

that of the oscillator alone, and

$$S_I = \sum_{i=0}^{j-1}\gamma_i q_i$$

(where $\gamma_i=\gamma(x_i, t_i)$) is the action of interaction between the particle and the oscillator. The normalizing constant, a, for the oscillator is $(2\pi\epsilon i/\hbar)^{-\frac{1}{2}}$. Now the exponential depends quadratically upon all the q_i. Hence, the integrations over all the variables q_i for $0<i<j$ can easily be performed. One is integrating a sequence of Gaussian integrals.

The result of these integrations is, writing $T=t''-t'$, $(2\pi i\hbar \sin\omega T/\omega)^{-\frac{1}{2}}\exp i(S_p+Q(q_j, q_0))/\hbar$, where $Q(q_j, q_0)$ turns out to be just the classical action for the forced harmonic oscillator (see reference 15). Explicitly it is

$$Q(q_j, q_0) = \frac{\omega}{2\sin\omega T}\Big[(\cos\omega T)(q_j^2+q_0^2) - 2q_jq_0$$
$$+\frac{2q_0}{\omega}\int_{t'}^{t''}\gamma(t)\sin\omega(t-t')dt$$
$$+\frac{2q_j}{\omega}\int_{t'}^{t''}\gamma(t)\sin\omega(t''-t)dt$$
$$-\frac{2}{\omega^2}\int_{t'}^{t''}\int_{t'}^{t}\gamma(t)\gamma(s)\sin\omega(t''-t)$$
$$\times\sin\omega(s-t')dsdt\Big].$$

It has been written as though $\gamma(t)$ were a continuous function of time. The integrals really should be split into Riemann sums and the quantity $\gamma(x_i, t_i)$ substituted for $\gamma(t_i)$. Thus, Q depends on the coordinates of the particle at all times through the $\gamma(x_i, t_i)$ and on that of the oscillator at times t' and t'' only. Thus, the quantity (61) becomes

$$\langle\chi_{t''}\varphi_m|1|\psi_{t'}\varphi_n\rangle_{S_p+S_0+S_I} = \int\cdots\int\chi_{t''}^*(x_j)G_{mn}$$
$$\times\exp\left(\frac{iS_p}{\hbar}\right)\psi_{t'}(x_0)\frac{dx_0}{A}\cdots\frac{dx_{j-1}}{A}dx_j$$
$$=\langle\chi_{t''}|G_{mn}|\psi_{t'}\rangle_{S_p}$$

which now contains the coordinates of the particle only, the quantity G_{mn} being given by

$$G_{mn} = (2\pi i\hbar\sin\omega T/\omega)^{-\frac{1}{2}}\int\int\varphi_m^*(q_j)$$
$$\times\exp(iQ(q_j, q_0)/\hbar)\varphi_n(q_0)dq_jdq_0.$$

Proceeding in an analogous manner one finds that all of the oscillators of the electromagnetic field can be eliminated from a description of the motion of the charges.

14. STATISTICAL MECHANICS

Spin and Relativity

Problems in the theory of measurement and statistical quantum mechanics are often simplified when set up from the point of view described here. For example, the influence of a perturbing measuring instrument can be integrated out in principle as we did in detail for the oscillator. The statistical density matrix has a fairly obvious and useful generalization. It results from considering the square of (38). It is an expression similar to (38) but containing integrations over two sets of variables dx_i and dx_i'. The exponential is replaced by $\exp i(S-S')/\hbar$, where S' is the same function of the x_i' as S is of x_i. It is required, for example, to describe the result of the elimination of the field oscillators where, say, the final state of the oscillators is unspecified and one desires only the sum over all final states m.

Spin may be included in a formal way. The Pauli spin equation can be obtained in this way:

One replaces the vector potential interaction term in $S(x_{i+1}, x_i)$,

$$\frac{e}{2c}(\mathbf{x}_{i+1}-\mathbf{x}_i)\cdot\mathbf{A}(\mathbf{x}_i)+\frac{e}{2c}(\mathbf{x}_{i+1}-\mathbf{x}_i)\cdot\mathbf{A}(\mathbf{x}_{i+1})$$

arising from expression (13) by the expression

$$\frac{e}{2c}(\boldsymbol{\sigma}\cdot(\mathbf{x}_{i+1}-\mathbf{x}_i))(\boldsymbol{\sigma}\cdot\mathbf{A}(\mathbf{x}_i)) + \frac{e}{2c}(\boldsymbol{\sigma}\cdot\mathbf{A}(\mathbf{x}_{i+1}))(\boldsymbol{\sigma}\cdot(\mathbf{x}_{i+1}-\mathbf{x}_i)).$$

Here $\mathbf{A}$ is the vector potential, $\mathbf{x}_{i+1}$ and $\mathbf{x}_i$ the vector positions of a particle at times t_{i+1} and t_i and $\boldsymbol{\sigma}$ is Pauli's spin vector matrix. The quantity Φ must now be expressed as $\prod_i \exp iS(x_{i+1}, x_i)/h$ for this differs from the exponential of the sum of $S(x_{i+1}, x_i)$. Thus, Φ is now a spin matrix.

The Klein Gordon relativistic equation can also be obtained formally by adding a fourth coordinate to specify a path. One considers a "path" as being specified by four functions $x^{(\mu)}(\tau)$ of a parameter τ. The parameter τ now goes in steps ϵ as the variable t went previously. The quantities $x^{(1)}(t)$, $x^{(2)}(t)$, $x^{(3)}(t)$ are the space coordinates of a particle and $x^{(4)}(t)$ is a corresponding time. The Lagrangian used is

$$\sum_{\mu=1}^{4}{}' [(dx^\mu/d\tau)^2+(e/c)(dx^\mu/d\tau)A_\mu],$$

where A_μ is the 4-vector potential and the terms in the sum for $\mu=1, 2, 3$ are taken with reversed sign. If one seeks a wave function which depends upon τ periodically, one can show this must satisfy the Klein Gordon equation. The Dirac equation results from a modification of the Lagrangian used for the Klein Gordon equation, which is analagous to the modification of the non-relativistic Lagrangian required for the Pauli equation. What results directly is the square of the usual Dirac operator.

These results for spin and relativity are purely formal and add nothing to the understanding of these equations. There are other ways of obtaining the Dirac equation which offer some promise of giving a clearer physical interpretation to that important and beautiful equation.

The author sincerely appreciates the helpful advice of Professor and Mrs. H. C. Corben and of Professor H. A. Bethe. He wishes to thank Professor J. A. Wheeler for very many discussions during the early stages of the work.

Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction

R. P. FEYNMAN*
Department of Physics, Cornell University, Ithaca, New York
(Received June 8, 1950)

The validity of the rules given in previous papers for the solution of problems in quantum electrodynamics is established. Starting with Fermi's formulation of the field as a set of harmonic oscillators, the effect of the oscillators is integrated out in the Lagrangian form of quantum mechanics. There results an expression for the effect of all virtual photons valid to all orders in $e^2/\hbar c$. It is shown that evaluation of this expression as a power series in $e^2/\hbar c$ gives just the terms expected by the aforementioned rules.

In addition, a relation is established between the amplitude for a given process in an arbitrary unquantized potential and in a quantum electrodynamical field. This relation permits a simple general statement of the laws of quantum electrodynamics.

A description, in Lagrangian quantum-mechanical form, of particles satisfying the Klein-Gordon equation is given in an Appendix. It involves the use of an extra parameter analogous to proper time to describe the trajectory of the particle in four dimensions.

A second Appendix discusses, in the special case of photons, the problem of finding what real processes are implied by the formula for virtual processes.

Problems of the divergences of electrodynamics are not discussed.

1. INTRODUCTION

IN two previous papers[1] rules were given for the calculation of the matrix element for any process in electrodynamics, to each order in $e^2/\hbar c$. No complete proof of the equivalence of these rules to the conventional electrodynamics was given in these papers. Secondly, no closed expression was given valid to all orders in $e^2/\hbar c$. In this paper these formal omissions will be remedied.[2]

In paper **II** it was pointed out that for many problems in electrodynamics the Hamiltonian method is not advantageous, and might be replaced by the over-all space-time point of view of a direct particle interaction. It was also mentioned that the Lagrangian form of quantum mechanics[3] was useful in this connection. The rules given in paper **II** were, in fact, first deduced in this form of quantum mechanics. We shall give this derivation here.

The advantage of a Lagrangian form of quantum mechanics is that in a system with interacting parts it permits a separation of the problem such that the motion of any part can be analyzed or solved first, and the results of this solution may then be used in the solution of the motion of the other parts. This separation is especially useful in quantum electrodynamics which represents the interaction of matter with the electromagnetic field. The electromagnetic field is an especially simple system and its behavior can be analyzed completely. What we shall show is that the net effect of the field is a delayed interaction of the particles. It is possible to do this easily only if it is not necessary at the same time to analyze completely the motion of the particles. The only advantage in our problems of the form of quantum mechanics in **C** is to permit one to separate these aspects of the problem. There are a number of disadvantages, however, such as a lack of familiarity, the apparent (but not real) necessity for dealing with matter in non-relativistic approximation, and at times a cumbersome mathematical notation and method, as well as the fact that a great deal of useful information that is known about operators cannot be directly applied.

It is also possible to separate the field and particle aspects of a problem in a manner which uses operators and Hamiltonians in a way that is much more familiar. One abandons the notation that the order of action of operators depends on their written position on the paper and substitutes some other convention (such that the order of operators is that of the time to which they refer). The increase in manipulative facility which accompanies this change in notation makes it easier to represent and to analyze the formal problems in electrodynamics. The method requires some discussion, however, and will be described in a succeeding paper. In this paper we shall give the derivations of the formulas of **II** by means of the form of quantum mechanics given in **C**.

The problem of interaction of matter and field will be analyzed by first solving for the behavior of the field in terms of the coordinates of the matter, and finally discussing the behavior of the matter (by matter is actually meant the electrons and positrons). That is to say, we shall first eliminate the field variables from the equations of motion of the electrons and then discuss the behavior of the electrons. In this way all of the rules given in the paper **II** will be derived.

Actually, the straightforward elimination of the field

* Now at the California Institute of Technology, Pasadena, California.

[1] R. P. Feynman, Phys. Rev. **76**, 749 (1949), hereafter called **I**, and Phys. Rev. **76**, 769 (1949), hereafter called **II**.

[2] See in this connection also the papers of S. Tomonaga, Phys. Rev. **74**, 224 (1948); S. Kanesawa and S. Tomonaga, Prog. Theoret. Phys. **3**, 101 (1948); J. Schwinger, Phys. Rev. **76**, 790 (1949); F. Dyson, Phys. Rev. **75**, 1736 (1949); W. Pauli and F. Villars, Rev. Mod. Phys. **21**, 434 (1949). The papers cited give references to previous work.

[3] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948), hereafter called **C**.

variables will lead at first to an expression for the behavior of an arbitrary number of Dirac electrons. Since the number of electrons might be infinite, this can be used directly to find the behavior of the electrons according to hole theory by imagining that nearly all the negative energy states are occupied by electrons. But, at least in the case of motion in a fixed potential, it has been shown that this hole theory picture is equivalent to one in which a positron is represented as an electron whose space-time trajectory has had its time direction reversed. To show that this same picture may be used in quantum electrodynamics when the potentials are not fixed, a special argument is made based on a study of the relationship of quantum electrodynamics to motion in a fixed potential. Finally, it is pointed out that this relationship is quite general and might be used for a general statement of the laws of quantum electrodynamics.

Charges obeying the Klein-Gordon equation can be analyzed by a special formalism given in Appendix A. A fifth parameter is used to specify the four-dimensional trajectory so that the Lagrangian form of quantum mechanics can be used. Appendix B discusses in more detail the relation of real and virtual photon emission. An equation for the propagation of a self-interacting electron is given in Appendix C.

In the demonstration which follows we shall restrict ourselves temporarily to cases in which the particle's motion is non-relativistic, but the transition of the final formulas to the relativistic case is direct, and the proof could have been kept relativistic throughout.

The transverse part of the electromagnetic field will be represented as an assemblage of independent harmonic oscillators each interacting with the particles, as suggested by Fermi.[4] We use the notation of Heitler.[5]

2. QUANTUM ELECTRODYNAMICS IN LAGRANGIAN FORM

The Hamiltonian for a set of non-relativistic particles interacting with radiation is, classically, $H=H_p+H_I+H_c+H_{tr}$, where $H_p+H_I=\sum_n \frac{1}{2}m_n^{-1}(\mathbf{p}_n-e_n\mathbf{A}^{tr}(\mathbf{x}_n))^2$ is the Hamiltonian of the particles of mass m_n, charge e_n, coordinate $\mathbf{x}_n$ and momentum $\mathbf{p}_n$ and their interaction with the transverse part of the electromagnetic field. This field can be expanded into plane waves

$$\mathbf{A}^{tr}(\mathbf{x})=(8\pi)^{\frac{1}{2}}\sum_{\mathbf{K}}[\mathbf{e}_1(q_{\mathbf{K}}^{(1)}\cos(\mathbf{K}\cdot\mathbf{x})+q_{\mathbf{K}}^{(3)}\sin(\mathbf{K}\cdot\mathbf{x}))+\mathbf{e}_2(q_{\mathbf{K}}^{(2)}\cos(\mathbf{K}\cdot\mathbf{x})+q_{\mathbf{K}}^{(4)}\sin(\mathbf{K}\cdot\mathbf{x}))] \quad (1)$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are two orthogonal polarization vectors at right angles to the propagation vector $\mathbf{K}$, magnitude k. The sum over $\mathbf{K}$ means, if normalized to unit volume, $\frac{1}{2}\int d^3\mathbf{K}/8\pi^3$, and each $q_{\mathbf{K}}^{(r)}$ can be considered as the coordinate of a harmonic oscillator. (The factor $\frac{1}{2}$ arises for the mode corresponding to $\mathbf{K}$ and to $-\mathbf{K}$ is the same.) The Hamiltonian of the transverse field represented as oscillators is

$$H_{tr}=\frac{1}{2}\sum_{\mathbf{K}}\sum_{r=1}^{4}((p_{\mathbf{K}}^{(r)})^2+k^2(q_{\mathbf{K}}^{(r)})^2)$$

where $p_{\mathbf{K}}^{(r)}$ is the momentum conjugate to $q_{\mathbf{K}}^{(r)}$. The longitudinal part of the field has been replaced by the Coulomb interaction,[6]

$$H_c=\tfrac{1}{2}\textstyle\sum_n\sum_m e_ne_m/r_{nm}$$

where $r_{nm}^2=(\mathbf{x}_n-\mathbf{x}_m)^2$. As is well known,[4] when this Hamiltonian is quantized one arrives at the usual theory of quantum electrodynamics. To express these laws of quantum electrodynamics one can equally well use the Lagrangian form of quantum mechanics to describe this set of oscillators and particles. The classical Lagrangian equivalent to this Hamiltonian is $L=L_p+L_I+L_c+L_{tr}$ where

$$L_p=\tfrac{1}{2}\textstyle\sum_n m_n\dot{\mathbf{x}}_n^2 \quad (2a)$$

$$L_I=\textstyle\sum_n e_n\dot{\mathbf{x}}_n\cdot\mathbf{A}^{tr}(\mathbf{x}_n) \quad (2b)$$

$$L_{tr}=\tfrac{1}{2}\textstyle\sum_{\mathbf{K}}\sum_r((\dot{q}_{\mathbf{K}}^{(r)})^2-k^2(q_{\mathbf{K}}^{(r)})^2) \quad (2c)$$

$$L_c=-\tfrac{1}{2}\textstyle\sum_n\sum_m e_ne_m/r_{mn}. \quad (2d)$$

When this Lagrangian is used in the Lagrangian forms of quantum mechanics of **C**, what it leads to is, of course, mathematically equivalent to the result of using the Hamiltonian H in the ordinary way, and is therefore equivalent to the more usual forms of quantum electrodynamics (at least for non-relativistic particles). We may, therefore, proceed by using this Lagrangian form of quantum electrodynamics, with the assurance that the results obtained must agree with those obtained from the more usual Hamiltonian form.

The Lagrangian enters through the statement that the functional which carries the system from one state to another is $\exp(iS)$ where

$$S=\int L\,dt=S_p+S_I+S_c+S_{tr}. \quad (3)$$

The time integrals must be written as Riemann sums with some care; for example,

$$S_I=\sum_n\int e_n\dot{\mathbf{x}}_n(t)\cdot\mathbf{A}^{tr}(\mathbf{x}_n(t))\,dt \quad (4)$$

becomes according to **C**, Eq. (19)

$$S_I=\textstyle\sum_n\sum_i\frac{1}{2}e_n(\mathbf{x}_{n,i+1}-\mathbf{x}_{n,i})\cdot(\mathbf{A}^{tr}(\mathbf{x}_{n,i+1})+\mathbf{A}^{tr}(\mathbf{x}_{n,i})) \quad (5)$$

so that the velocity $\dot{\mathbf{x}}_{n,i}$ which multiplies $\mathbf{A}^{tr}(\mathbf{x}_{n,i})$ is

$$\dot{\mathbf{x}}_{n,i}=\tfrac{1}{2}\epsilon^{-1}(\mathbf{x}_{n,i+1}-\mathbf{x}_{n,i})+\tfrac{1}{2}\epsilon^{-1}(\mathbf{x}_{n,i}-\mathbf{x}_{n,i-1}). \quad (6)$$

[4] E. Fermi, Rev. Mod. Phys. 4, 87 (1932).

[5] W. Heitler, *The Quantum Theory of Radiation*, second edition (Oxford University Press, London, 1944).

[6] The term in the sum for $n=m$ is obviously infinite but must be included for relativistic invariance. Our problem here is to re-express the usual (and divergent) form of electrodynamics in the form given in **II**. Modifications for dealing with the divergences are discussed in **II** and we shall not discuss them further here.

In the Lagrangian form it is possible to eliminate the transverse oscillators as is discussed in **C**, Section 13. One must specify, however, the initial and final state of all oscillators. We shall first choose the special, simple case that all oscillators are in their ground states initially and finally, so that all photons are virtual. Later we do the more general case in which real quanta are present initially or finally. We ask, then, for the amplitude for finding no quanta present and the particles in state $\chi_{t''}$ at time t'', if at time t' the particles were in state $\psi_{t'}$ and no quanta were present.

The method of eliminating field oscillators is described in Section 13 of **C**. We shall simply carry out the elimination here using the notation and equations of **C**. To do this, for simplicity, we first consider in the next section the case of a particle or a system of particles interacting with a single oscillator, rather than the entire assemblage of the electromagnetic field.

3. FORCED HARMONIC OSCILLATOR

We consider a harmonic oscillator, coordinate q, Lagrangian $L=\frac{1}{2}(\dot{q}^2-\omega^2q^2)$ interacting with a particle or system of particles, action S_p, through a term in the Lagrangian $q(t)\gamma(t)$ where $\gamma(t)$ is a function of the coordinates (symbolized as x) of the particle. The precise form of $\gamma(t)$ for each oscillator of the electromagnetic field is given in the next section. We ask for the amplitude that at some time t'' the particles are in state $\chi_{t''}$ and the oscillator is in, say, an eigenstate m of energy $\omega(m+\frac{1}{2})$ (units are chosen such that $\hbar=c=1$) when it is given that at a previous time t' the particles were in state $\psi_{t'}$ and the oscillator in n. The amplitude for this is the transition amplitude [see **C**, Eq. (61)]

$$\langle\chi_{t''}\varphi_m|1|\psi_{t'}\varphi_n\rangle_{S_p+S_0+S_I}$$
$$=\int\int\chi_{t''}^*(x_{t''})\varphi_m^*(q_{t''})\exp i(S_p+S_0+S_I)$$
$$\varphi_n(q_{t'})\psi_{t'}(x_{t'})dx_{t''}dx_{t'}dq_{t''}dq_{t'}\mathfrak{D}x(t)\mathfrak{D}q(t) \quad (7)$$

where x represents the variables describing the particle, S_p is the action calculated classically for the particles for a given path going from coordinate $x_{t'}$ at t' to $x_{t''}$ at t'', S_0 is the action $\int\frac{1}{2}(\dot{q}^2-\omega^2q^2)dt$ for any path of the oscillator going from $q_{t'}$ at t' to $q_{t''}$ at t'', while

$$S_I=\int q(t)\gamma(t)dt, \quad (8)$$

the action of interaction, is a functional of both $q(t)$ and $x(t)$, the paths of oscillator and particles. The symbols $\mathfrak{D}x(t)$ and $\mathfrak{D}q(t)$ represent a summation over all possible paths of particles and oscillator which go between the given end points in the sense defined in **C**, Eq. (9). (That is, assuming time to proceed in infinitesimal steps, ϵ, an integral over all values of the coordinates x and q corresponding to each instant in time, suitably normalized.)

The problem may be broken in two. The result can be written as an integral over all paths of the particles only, of $(\exp iS_p)\cdot G_{mn}$:

$$\langle\chi_{t''}\varphi_m|1|\psi_{t'}\varphi_n\rangle_{S_p+S_0+S_I}=\langle\chi_{t''}|G_{mn}|\psi_{t'}\rangle_{S_p} \quad (9)$$

where G_{mn} is a functional of the path of the particles alone (since it depends on $\gamma(t)$) given by

$$G_{mn}=\left\langle\varphi_m\left|\exp i\int q(t)\gamma(t)dt\right|\varphi_n\right\rangle_{S_0}$$
$$=\int\varphi_m^*(q_{t''})\exp i(S_0+S_I)\varphi_n(q_{t'})dq_{t'}dq_{t''}\mathfrak{D}q(t)$$
$$=\int\varphi_m^*(q_j)\exp i\epsilon\sum_{i=0}^{j-1}\left[\tfrac{1}{2}\epsilon^{-2}(q_{i+1}-q_i)^2-\tfrac{1}{2}\omega^2q_i^2+q_i\gamma_i\right]$$
$$\cdot\varphi_n(q_0)dq_0a^{-1}dq_1a^{-1}dq_2\cdots a^{-1}dq_j \quad (10)$$

where we have written the $\mathfrak{D}q(t)$ out explicitly (and have set $a=(2\pi i\epsilon)^{\frac{1}{2}}$, $t''-t'=j\epsilon$, $q_{t'}=q_0$, $q_{t''}=q_j$). The last form can be written as

$$G_{mn}=\int\varphi_m^*(q_j)k(q_j,t'';q_0,t')\varphi_n(q_0)dq_0dq_j \quad (11)$$

where $k(q_j,t'';q_0,t')$ is the kernel [as in **I**, Eq. (2)] for a forced harmonic oscillator giving the amplitude for arrival at q_j at time t'' if at time t' it was known to be at q_0. According to **C** it is given by

$$k(q_j,t'';q_0,t')=(2\pi i\omega^{-1}\sin\omega(t''-t'))^{-\frac{1}{2}}$$
$$\times\exp iQ(q_j,t'';q_0,t') \quad (12)$$

where $Q(q_j,t'';q_0,t')$ is the action calculated along the classical path between the end points $q_j,t'';q_0,t'$, and is given explicitly in **C**.[7] It is

[7] That (12) is correct, at least insofar as it depends on q_0, can be seen directly as follows. Let $\bar{q}(t)$ be the classical path which satisfies the boundary condition $\bar{q}(t')=q_0$, $\bar{q}(t'')=q_j$. Then in the integral defining k replace each of the variables q_i by $q_i=\bar{q}_i+y_i$, $(\bar{q}_i=\bar{q}(t_i))$, that is, use the displacement y_i from the classical path $\bar{q}_i$ as the coordinate rather than the absolute position. With the substitution $q_i=\bar{q}_i+y_i$ in the action

$$S_0+S_I=\int(\tfrac{1}{2}\dot{q}^2-\tfrac{1}{2}\omega^2q^2+\gamma q)dt$$
$$=\int(\tfrac{1}{2}\dot{\bar{q}}^2-\tfrac{1}{2}\omega^2\bar{q}^2+\gamma\bar{q})dt+\int(\tfrac{1}{2}\dot{y}^2-\tfrac{1}{2}\omega^2y^2)dt$$

the terms linear in y drop out by integrations by parts using the equation of motion $\ddot{\bar{q}}=-\omega^2\bar{q}+\gamma(t)$ for the classical path, and the boundary conditions $y(t')=y(t'')=0$. That this should occur should occasion no surprise, for the action functional is an extremum at $q(t)=\bar{q}(t)$ so that it will only depend to second order in the displacements y from this extremal orbit $\bar{q}(t)$. Further, since the action functional is quadratic to begin with, it cannot depend on y more than quadratically. Hence

$$S_0+S_I=Q+\int(\tfrac{1}{2}\dot{y}^2-\tfrac{1}{2}\omega^2y^2)dt$$

so that since $dq_i=dy_i$,

$$k(q_i,t'';q_0,t')=\exp(iQ)\int\exp\left(i\int\tfrac{1}{2}(\dot{y}^2-\omega^2y^2)dt\right)\mathfrak{D}y(t).$$

The factor following the $\exp iQ$ is the amplitude for a free oscillator to proceed from $y=0$ at $t=t'$ to $y=0$ at $t=t''$ and does not there-

$$Q=\frac{\omega}{2\sin\omega(t''-t')}\Big[(q_j^2+q_0^2)\cos\omega(t''-t')-2q_jq_0$$
$$+\frac{2q_j}{\omega}\int_{t'}^{t''}\gamma(t)\sin\omega(t-t')dt$$
$$+\frac{2q_0}{\omega}\int_{t'}^{t''}\gamma(t)\sin\omega(t''-t)dt$$
$$-\frac{2}{\omega^2}\int_{t'}^{t''}\int_{t'}^{t}\gamma(t)\gamma(s)\sin\omega(t''-t)$$
$$\times\sin\omega(s-t')dsdt\Big]. \quad (13)$$

The solution of the motion of the oscillator can now be completed by substituting (12) and (13) into (11) and performing the integrals. The simplest case is for $m, n=0$ for which case[3]

$$\varphi_0(q_0)=(\omega/\pi)^{\frac{1}{4}}\exp(-\tfrac{1}{2}\omega q_0^2)\exp(-\tfrac{1}{2}i\omega t')$$

so that the integrals on q_0, q_j are just Gaussian integrals. There results

$$G_{00}=\exp\Big(-\frac{1}{2}\omega^{-1}\int_{t'}^{t''}\int_{t'}^{t}\exp(-i\omega(t-s))\gamma(t)\gamma(s)dtds\Big)$$

a result of fundamental importance in the succeeding developments. By replacing $t-s$ by its absolute value $|t-s|$ we may integrate both variables over the entire range and divide by 2. We will henceforth make the results more general by extending the limits on the integrals from $-\infty$ to $+\infty$. Thus if one wishes to study the effect on a particle of interaction with an oscillator for just the period t' to t'' one may use

$$G_{00}=\exp\Big(-\frac{1}{4\omega}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}$$
$$\times\exp(-i\omega|t-s|)\gamma(t)\gamma(s)dtds\Big) \quad (14)$$

imagining in this case that the interaction $\gamma(t)$ is zero outside these limits. We defer to a later section the discussion of other values of m, n.

Since G_{00} is simply an exponential, we can write it as $\exp(iI)$, consider that the complete "action" for the system of particles is $S=S_p+I$ and that one computes transition elements with this "action" instead of S_p (see C, Sec. 12). The functional I, which is given by

$$I=\tfrac{1}{4}i\omega^{-1}\int\int\exp(-i\omega|t-s|)\gamma(s)\gamma(t)dsdt \quad (15)$$

is complex, however; we shall speak of it as the complex action. It describes the fact that the system at one time can affect itself at a different time by means of a temporary storage of energy in the oscillator. When there are several independent oscillators with different interactions, the effect, if they are all in the lowest state at t' and t'', is the product of their separate G_{00} contributions. Thus the complex action is additive, being the sum of contributions like (15) for each of the several oscillators.

4. VIRTUAL TRANSITIONS IN THE ELECTROMAGNETIC FIELD

We can now apply these results to eliminate the transverse field oscillators of the Lagrangian (2). At first we can limit ourselves to the case of purely virtual transitions in the electromagnetic field, so that there is no photon in the field at t' and t''. That is, all of the field oscillators are making transitions from ground state to ground state.

The $\gamma_{\mathbf{K}}^{(r)}$ corresponding to each oscillator $q_{\mathbf{K}}^{(r)}$ is found from the interaction term L_I [Eq. (2b)], substituting the value of $\mathbf{A}^{tr}(\mathbf{x})$ given in (1). There results, for example,

$$\gamma_{\mathbf{K}}^{(1)}=(8\pi)^{\frac{1}{2}}\textstyle\sum_n e_n(\mathbf{e}_1\cdot\mathbf{x}'_n)\cos(\mathbf{K}\cdot\mathbf{x}_n)$$
$$\gamma_{\mathbf{K}}^{(3)}=(8\pi)^{\frac{1}{2}}\textstyle\sum_n e_n(\mathbf{e}_1\cdot\mathbf{x}'_n)\sin(\mathbf{K}\cdot\mathbf{x}_n) \quad (16)$$

the corresponding results for $\gamma_{\mathbf{K}}^{(2)}$, $\gamma_{\mathbf{K}}^{(4)}$ replace $\mathbf{e}_1$ by $\mathbf{e}_2$.

The complex action resulting from oscillator of coordinate $q_{\mathbf{K}}^{(1)}$ is therefore

$$I_{\mathbf{K}}^{(1)}=\frac{8\pi i}{4k}\sum_n\sum_m\int\int e_ne_m\exp(-ik|t-s|)(\mathbf{e}_1\cdot\mathbf{x}'_n(t))$$
$$\times(\mathbf{e}_1\cdot\mathbf{x}'_m(s))\cdot\cos(\mathbf{K}\cdot\mathbf{x}_n(t))\cos(\mathbf{K}\cdot\mathbf{x}_m(s))dsdt.$$

The term $I_{\mathbf{K}}^{(3)}$ exchanges the cosines for sines, so in the sum $I_{\mathbf{K}}^{(1)}+I_{\mathbf{K}}^{(3)}$ the product of the two cosines, $\cos A\cdot\cos B$ is replaced by $(\cos A\cos B+\sin A\sin B)$ or $\cos(A-B)$. The terms $I_{\mathbf{K}}^{(2)}+I_{\mathbf{K}}^{(4)}$ give the same result with $\mathbf{e}_2$ replacing $\mathbf{e}_1$. The sum $(\mathbf{e}_1\cdot\mathbf{V})(\mathbf{e}_1\cdot\mathbf{V}')+(\mathbf{e}_2\cdot\mathbf{V})(\mathbf{e}_2\cdot\mathbf{V}')$ is $(\mathbf{V}\cdot\mathbf{V}')-k^{-2}(\mathbf{K}\cdot\mathbf{V})(\mathbf{K}\cdot\mathbf{V}')$ since it is the sum of the products of vector components in two orthogonal directions, so that if we add the product in the third direction (that of $\mathbf{K}$) we construct the complete scalar product. Summing over all $\mathbf{K}$ then, since $\sum_{\mathbf{K}}=\frac{1}{2}\int d^3\mathbf{K}/8\pi^3$ we find for the total complex action of all of the transverse oscillators,

$$I_{tr}=i\sum_n\sum_m\int_{t'}^{t''}dt\int_{t'}^{t''}ds\int e_ne_m\exp(-ik|t-s|)$$
$$\times[\mathbf{x}'_n(t)\cdot\mathbf{x}'_m(s)-k^{-2}(\mathbf{K}\cdot\mathbf{x}'_n(t))(\mathbf{K}\cdot\mathbf{x}'_m(s))]$$
$$\cdot\cos(\mathbf{K}\cdot(\mathbf{x}_n(t)-\mathbf{x}_m(s)))d^3\mathbf{K}/8\pi^2k. \quad (17)$$

---

fore depend on q_0, q_j, or $\gamma(t)$, being a function only of $t''-t'$. [That it is actually $(2\pi i\omega^{-1}\sin\omega(t''-t'))^{-\frac{1}{2}}$ can be demonstrated either by direct integration of the y variables or by using some normalizing property of the kernels k, for example that G_{00} for the case $\gamma=0$ must equal unity.] The expression for Q given in C on page 386 is in error, the quantities q_0 and q_j should be interchanged.

[3] It is most convenient to define the state φ_n with the phase factor $\exp[-i\omega(n+\frac{1}{2})t']$ and the final state with the factor $\exp[-i\omega(m+\frac{1}{2})t'']$ so that the results will not depend on the particular times t', t'' chosen.

This is to be added to S_p+S_e to obtain the complete action of the system with the oscillators removed.

The term in $(\mathbf{K}\cdot\mathbf{x}'_n(t))(\mathbf{K}\cdot\mathbf{x}'_m(s))$ can be simplified by integration by parts with respect to t and with respect to s [note that $\exp(-ik|t-s|)$ has a discontinuous slope at $t=s$, or break the integration up into two regions]. One finds

$$I_{tr}=R-I_c+I_{\text{transient}} \tag{18}$$

where

$$R=-i\sum_n\sum_m\int_{t'}^{t''}dt\int_{t'}^{t''}ds\int e_ne_m \times\exp(-ik|t-s|)(1-\mathbf{x}'_n(t)\cdot\mathbf{x}'_m(s)) \cdot\cos\mathbf{K}\cdot(\mathbf{x}_n(t)-\mathbf{x}_m(s))d^3\mathbf{K}/8\pi^2k \tag{19}$$

and

$$I_c=-\sum_n\sum_m\int_{t'}^{t''}dt\int e_ne_m \times\cos\mathbf{K}\cdot(\mathbf{x}_n(t)-\mathbf{x}_m(t))d^3\mathbf{K}/4\pi^2k^2 \tag{20}$$

comes from the discontinuity in slope of $\exp(-ik|t-s|)$ at $t=s$. Since

$$\int\cos(\mathbf{K}\cdot\mathbf{R})d^3\mathbf{K}/4\pi^2k^2=\int_0^\infty(kr)^{-1}\sin(kr)dk/\pi=(2r)^{-1}$$

this term I_c just cancels the Coulomb interaction term $S_c=\int L_cdt$. The term

$$\begin{aligned}I_{\text{transient}}=&-\sum_n\sum_m e_ne_m\int\frac{d^3\mathbf{K}}{4\pi^2k^2}\\&\times\Big\{\int_{t'}^{t''}[\exp(-ik(t''-t))\cos\mathbf{K}\cdot(\mathbf{x}_n(t'')-\mathbf{x}_m(t))\\&+\exp(-ik(t-t'))\cos\mathbf{K}\cdot(\mathbf{x}_n(t)-\mathbf{x}_m(t'))]dt\\&+(2k)^{-1}i[\cos\mathbf{K}\cdot(\mathbf{x}_n(t'')-\mathbf{x}_m(t''))\\&+\cos\mathbf{K}\cdot(\mathbf{x}_n(t')-\mathbf{x}_m(t'))\\&-2\exp(-ik(t''-t'))\cos\mathbf{K}\cdot(\mathbf{x}_n(t')-\mathbf{x}_m(t''))]\Big\}.\end{aligned} \tag{21}$$

is one which comes from the limits of integration at t' and t'', and involves the coordinates of the particle at either one of these times or the other. If t' and t'' are considered to be exceedingly far in the past and future, there is no correlation to be expected between these temporally distant coordinates and the present ones, so the effects of $I_{\text{transient}}$ will cancel out quantum mechanically by interference. This transient was produced by the sudden turning on of the interaction of field and particles at t' and its sudden removal at t''. Alternatively we can imagine the charges to be turned on after t' adiabatically and turned off slowly before t'' (in this case, in the term L_c, the charges should also be considered as varying with time). In this case, in the limit, $I_{\text{transient}}$ is zero.[9] Hereafter we shall drop the transient term and consider the range of integration of t to be from $-\infty$ to $+\infty$, imagining, if one needs a definition, that the charges vary with time and vanish in the direction of either limit.

To simplify R we need the integral

$$J=\int\exp(-ik|t|)\cos(\mathbf{K}\cdot\mathbf{R})d^3\mathbf{K}/8\pi^2k =\int_0^\infty\exp(-ik|t|)\sin(kr)dk/2\pi r \tag{22}$$

where r is the length of the vector $\mathbf{R}$. Now

$$\int_0^\infty\exp(-ikx)dk=\lim_{\epsilon\to0}(-i(x-i\epsilon)^{-1}) =-ix^{-1}+\pi\delta(x)=\pi\delta_+(x)$$

where the equation serves to define $\delta_+(x)$ [as in **II**, Eq. (3)]. Hence, expanding $\sin(kr)$ in exponentials find

$$\begin{aligned}J=&-(4\pi r)^{-1}((|t|-r)^{-1}-(|t|+r)^{-1})\\&+(4ir)^{-1}(\delta(|t|-r)-\delta(|t|+r))\\=&-(2\pi)^{-1}(t^2-r^2)^{-1}+(2i)^{-1}\delta(t^2-r^2)\\=&-\tfrac{1}{2}i\delta_+(t^2-r^2)\end{aligned} \tag{23}$$

where we have used the fact that

$$\delta(t^2-r^2)=(2r)^{-1}(\delta(|t|-r)+\delta(|t|+r))$$

and that $\delta(|t|+r)=0$ since both $|t|$ and r are necessarily positive.

Substitution of these results into (19) gives finally,

$$R=-\frac{1}{2}\sum_n\sum_m\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}e_ne_m(1-\mathbf{x}'_n(t)\cdot\mathbf{x}'_m(s)) \times\delta_+((t-s)^2-(\mathbf{x}_n(t)-\mathbf{x}_m(s))^2)dtds. \tag{24}$$

The total complex action of the system is then[10] S_p+R. Or, what amounts to the same thing; to obtain

[9] One can obtain the final result, that the total interaction is just R, in a formal manner starting from the Hamiltonian from which the longitudinal oscillators have not yet been eliminated. There are for each $\mathbf{K}$ and cos or sin, four oscillators $q_{\mu\mathbf{K}}$ corresponding to the three components of the vector potential ($\mu=1$, 2, 3) and the scalar potential ($\mu=4$). It must then be assumed that the wave functions of the initial and final state of the $\mathbf{K}$ oscillators is the function $(k/\pi)\exp[-\frac{1}{2}k(q_{1\mathbf{K}}^2+q_{2\mathbf{K}}^2+q_{3\mathbf{K}}^2-q_{4\mathbf{K}}^2)]$. The wave function suggested here has only formal significance, of course, because the dependence on $q_{4\mathbf{K}}$ is not square integrable, and cannot be normalized. If each oscillator were assumed actually in the ground state, the sign of the $q_{4\mathbf{K}}$ term would be changed to positive, and the sign of the frequency in the contribution of these oscillators would be reversed (they would have negative energy).

[10] The classical action for this problem is just $Sp+R'$ where R' is the real part of the expression (24). In view of the generalization of the Lagrangian formulation of quantum mechanics suggested in Section 12 of C, one might have anticipated that R would have been simply R'. This corresponds, however, to boundary conditions other than no quanta present in past and future. It is harder to interpret physically. For a system enclosed in a light tight box, however, it appears likely that both R and R' lead to the same results.

transition amplitudes including the effects of the field we must calculate the transition element of $\exp(iR)$:

$$\langle\chi_{t''}|\exp iR|\psi_{t'}\rangle S_p \tag{25}$$

under the action S_p of the particles, excluding interaction. Expression (24) for R must be considered to be written in the usual manner as a Riemann sum and the expression (25) interpreted as defined in **C** [Eq. (39)]. Expression (6) must be used for $\mathbf{x}'_n$ at time t.

Expression (25), with (24), then contains all the effects of virtual quanta on a (at least non-relativistic) system according to quantum electrodynamics. It contains the effects to all orders in $e^2/\hbar c$ in a single expression. If expanded in a power series in $e^2/\hbar c$, the various terms give the expressions to the corresponding order obtained by the diagrams and methods of **II**. We illustrate this by an example in the next section.

5. EXAMPLE OF APPLICATION OF EXPRESSION (25)

We shall not be much concerned with the non-relativistic case here, as the relativistic case given below is as simple and more interesting. It is, however, very similar and at this stage it is worth giving an example to show how expressions resulting from (25) are to be interpreted according to the rules of **C**. For example, consider the case of a single electron, coordinate $\mathbf{x}$, either free or in an external given potential (contained for simplicity in S_p, not in[11] R). Its interaction with the field produces a reaction back on itself given by R as in (24) but in which we keep only a single term corresponding to $m=n$. Assume the effect of R to be small and expand $\exp(iR)$ as $1+iR$. Let us find the amplitude at time t'' of finding the electron in a state ψ with no quanta emitted, if at time t' it was in the same state. It is

$$\langle\psi_{t''}|1+iR|\psi_{t'}\rangle S_p=\langle\psi_{t''}|1|\psi_{t'}\rangle S_p+i\langle\psi_{t''}|R|\psi_{t'}\rangle S_p$$

where $\langle\psi_{t''}|1|\psi_{t'}\rangle S_p=\exp[-iE(t''-t')]$ if E is the energy of the state, and

$$\langle\psi_{t''}|R|\psi_{t'}\rangle S_p=-\tfrac{1}{2}e^2\int_{t'}^{t''}dt\int_{t'}^{t''}ds\langle\psi_{t''}|(1-\mathbf{x}'_t\cdot\mathbf{x}'_s)\times\delta_+((t-s)^2-(\mathbf{x}_t-\mathbf{x}_s)^2)|\psi_{t'}\rangle S_p. \tag{26}$$

Here $\mathbf{x}_s=\mathbf{x}(s)$, etc. In (26) we shall limit the range of integrations by assuming $s<t$, and double the result.

The expression within the brackets $\langle\ \rangle S_p$ on the right-hand side of (26) can be evaluated by the methods described in **C** [Eq. (29)]. An expression such as (26) can also be evaluated directly in terms of the propagation kernel $K(2,1)$ [see **I**, Eq. (2)] for an electron moving in the given potential.

The term $\mathbf{x}'_s\cdot\mathbf{x}'_t$ in the non-relativistic case produces an interesting complication which does not have an analog for the relativistic case with the Dirac equation. We discuss it below, but for a moment consider in further detail expression (26) but with the factor $(1-\mathbf{x}'_s\cdot\mathbf{x}'_t)$ replaced simply by unity.

The kernel $K(2,1)$ is defined and discussed in **I**. From its definition as the amplitude that the electron be found at $\mathbf{x}_2$ at time t_2, if at t_1 it was at $\mathbf{x}_1$, we have

$$K(\mathbf{x}_2,t_2;\mathbf{x}_1,t_1)=\langle\delta(\mathbf{x}-\mathbf{x}_2)t_2|1|\delta(\mathbf{x}-\mathbf{x}_1)t_1\rangle S_p \tag{27}$$

that is, more simply $K(2,1)$ is the sum of $\exp(iS_p)$ over all paths which go from space time point 1 to 2.

In the integrations over all paths implied by the symbol in (26) we can first integrate over all the $\mathbf{x}_i$ variables corresponding to times t_i from t' to s, not inclusive, the result being a factor $K(\mathbf{x}_s,s;\mathbf{x}_{t'},t')$ according to (27). Next we integrate on the variables between s and t not inclusive, giving a factor $K(\mathbf{x}_t,t;\mathbf{x}_s,s)$ and finally on those between t and t'' giving $K(\mathbf{x}_{t''},t'';\mathbf{x}_t,t)$. Hence the left-hand term in (26) excluding the $\mathbf{x}'_t\cdot\mathbf{x}'_s$ factor is

$$-e^2\int dt\int ds\int\psi^*(\mathbf{x}_{t''},t'')K(\mathbf{x}_{t''},t'';\mathbf{x}_t,t)\delta_+((t-s)^2-(\mathbf{x}_t-\mathbf{x}_s)^2)\cdot K(\mathbf{x}_t,t;\mathbf{x}_s,s)K(\mathbf{x}_s,s;\mathbf{x}_{t'},t')\times\psi(\mathbf{x}_{t'},t')d^3\mathbf{x}_{t''}d^3\mathbf{x}_td^3\mathbf{x}_sd^3\mathbf{x}_{t'} \tag{28}$$

which in improved notation and in the relativistic case is essentially the result given in **II**.

We have made use of a special case of a principle which may be stated more generally as

$$\langle\chi_{t''}|F(\mathbf{x}_1,t_1;\mathbf{x}_2,t_2;\cdots\mathbf{x}_k,t_k)|\psi_{t'}\rangle S_p=\int\chi^*(\mathbf{x}_{t''},t'';\mathbf{x}_1,t_1)\cdot K(\mathbf{x}_1,t_1;\mathbf{x}_2,t_2)\cdots\times K(\mathbf{x}_{k-1},t_{k-1};\mathbf{x}_k,t_k)K(\mathbf{x}_k,t_k;\mathbf{x}_{t'},t')\cdot F(\mathbf{x}_1,t_1;\mathbf{x}_2,t_2;\cdots\mathbf{x}_k,t_k)\psi(\mathbf{x}_{t'})\times d^3\mathbf{x}_{t''}d^3\mathbf{x}_1d^3\mathbf{x}_2\cdots d^3\mathbf{x}_kd^3\mathbf{x}_{t'} \tag{29}$$

where F is any function of the coordinate $\mathbf{x}_1$ at time t_1, $\mathbf{x}_2$ at t_2 up to $\mathbf{x}_k$, t_k, and, it is important to notice, we have assumed $t''>t_1>t_2>\cdots t_k>t'$.

Expressions of higher order arising for example from R^2 are more complicated as there are quantities referring to several different times mixed up, but they all can be interpreted readily. One simply breaks up the ranges of integrations of the time variables into parts such that in each the order of time of each variable is definite. One then interprets each part by formula (29).

[11] One can show from (25) how the correlated effect of many atoms at a distance produces on a given system the effects of an external potential. Formula (24) yields the result that this potential is that obtained from Liénard and Wiechert by retarded waves arising from the charges and currents resulting from the distant atoms making transitions. Assume the wave functions χ and ψ can be split into products of wave functions for system and distant atoms and expand $\exp(iR)$ assuming the effect of any individual distant atom is small. Coulomb potentials arise even from nearby particles if they are moving slowly.

As a simple example we may refer to the problem of the transition element

$$\left\langle \chi_{t''} \middle| \int U(\mathbf{x}(t), t)dt \int V(\mathbf{x}(s), s)ds \middle| \psi_{t'} \right\rangle$$

arising, say, in the cross term in U and V in an ordinary second order perturbation problem (disregarding radiation) with perturbation potential $U(\mathbf{x}, t)+V(\mathbf{x}, t)$. In the integration on s and t which should include the entire range of time for each, we can split the range of s into two parts, $s<t$ and $s>t$. In the first case, $s<t$, the potential V acts earlier than U, and in the other range, vice versa, so that

$$\begin{aligned}
&\left\langle \chi_{t''} \middle| \int U(\mathbf{x}_t, t)dt \int V(\mathbf{x}_s, s)ds \middle| \psi_{t'} \right\rangle \\
&= \int_{t'}^{t''} dt \int_{t'}^{t} ds \int \chi^*(\mathbf{x}_{t''})K(\mathbf{x}_{t''}, t''; \mathbf{x}_t, t) \\
&\quad \times U(\mathbf{x}_t, t)K(\mathbf{x}_t, t; \mathbf{x}_s, s)V(\mathbf{x}_s, s) \\
&\quad \cdot K(\mathbf{x}_s, s; \mathbf{x}_{t'}, t')\psi(\mathbf{x}_{t'})d^3\mathbf{x}_{t''}d^3\mathbf{x}_t d^3\mathbf{x}_s d^3\mathbf{x}_{t'} \\
&\quad + \int_{t'}^{t''} dt \int_{t}^{t''} ds \int \chi^*(\mathbf{x}_{t''})K(\mathbf{x}_{t''}, t''; \mathbf{x}_s, s) \\
&\quad \times V(\mathbf{x}_s, s)K(\mathbf{x}_s, s; \mathbf{x}_t, t)U(\mathbf{x}_t, t) \\
&\quad \cdot K(\mathbf{x}_t, t; \mathbf{x}_{t'}, t')\psi(\mathbf{x}_{t'})d^3\mathbf{x}_{t''}d^3\mathbf{x}_s d^3\mathbf{x}_t d^3\mathbf{x}_{t'} \quad (30)
\end{aligned}$$

so that the single expression on the left is represented by two terms analogous to the two terms required in analyzing the Compton effect. It is in this way that the several terms and their corresponding diagrams corresponding to each process arise when an attempt is made to represent the transition elements of single expressions involving time integrals in terms of the propagation kernels K.

It remains to study in more detail the term in (26) arising from $\mathbf{x}'(t)\cdot\mathbf{x}'(s)$ in the interaction. The interpretation of such expressions is considered in detail in **C**, and we must refer to Eqs. (39) through (50) of that paper for a more thorough analysis. A similar type of term also arises in the Lagrangian formulation in simpler problems, for example the transition element

$$\left\langle \chi_{t''} \middle| \int \mathbf{x}'(t)\cdot\mathbf{A}(\mathbf{x}(t), t)dt \int \mathbf{x}'(s)\cdot\mathbf{B}(\mathbf{x}(s), s)ds \middle| \psi_{t'} \right\rangle$$

arising say, in the cross term in **A** and **B** in a second-order perturbation problem for a particle in a perturbing vector potential $\mathbf{A}(\mathbf{x}, t)+\mathbf{B}(\mathbf{x}, t)$. The time integrals must first be written as Riemannian sums, the velocity (see (6)) being replaced by $\mathbf{x}'=\frac{1}{2}\epsilon^{-1}(\mathbf{x}_{i+1}-\mathbf{x}_i)+\frac{1}{2}\epsilon^{-1}(\mathbf{x}_i-\mathbf{x}_{i-1})$ so that we ask for the transition element of

$$\sum_i\sum_j [\tfrac{1}{2}(\mathbf{x}_{i+1}-\mathbf{x}_i)+\tfrac{1}{2}(\mathbf{x}_i-\mathbf{x}_{i-1})]\cdot\mathbf{A}(\mathbf{x}_i, t_i) \times [\tfrac{1}{2}(\mathbf{x}_{j+1}-\mathbf{x}_j)+\tfrac{1}{2}(\mathbf{x}_j-\mathbf{x}_{j-1})]\cdot\mathbf{B}(\mathbf{x}_j, t_j). \quad (31)$$

In **C** it is shown that when converted to operator notation the quantity $(x_{i-1}-x_i)/\epsilon$ is equivalent (nearly, see below) to an operator,

$$(x_{i+1}-x_i)/\epsilon \rightarrow i(Hx-xH) \quad (32)$$

operating in order indicated by the time index i (that is after x_l's for $l\leq i$ and before all x_l's for $l>i$). In non-relativistic mechanics $i(Hx-xH)$ is the momentum operator p_x divided by the mass m. Thus in (31) the expression $[\frac{1}{2}(\mathbf{x}_{i+1}-\mathbf{x}_i)+\frac{1}{2}(\mathbf{x}_i-\mathbf{x}_{i-1})]\cdot\mathbf{A}(\mathbf{x}_i, t_i)$ becomes $\epsilon(\mathbf{p}\cdot\mathbf{A}+\mathbf{A}\cdot\mathbf{p})/2m$. Here again we must split the sum into two regions $j<i$ and $j>i$ so the quantities in the usual notation will operate in the right order such that eventually (31) becomes identical with the right-hand side of Eq. (30) but with $U(\mathbf{x}_t, t)$ replaced by the operator

$$\frac{1}{2m}\left(\frac{1}{i}\frac{\partial}{\partial\mathbf{x}_t}\cdot\mathbf{A}(\mathbf{x}_t, t)+\mathbf{A}(\mathbf{x}_t, t)\cdot\frac{1}{i}\frac{\partial}{\partial\mathbf{x}_t}\right)$$

standing in the same place, and with the operator

$$\frac{1}{2m}\left(\frac{1}{i}\frac{\partial}{\partial\mathbf{x}_s}\cdot\mathbf{B}(\mathbf{x}_s, s)+\frac{1}{i}\mathbf{B}(\mathbf{x}_s, s)\cdot\frac{\partial}{\partial\mathbf{x}_s}\right)$$

standing in the place of $V(\mathbf{x}_s, s)$. The sums and factors ϵ have now become $\int dt\int ds$.

This is nearly but not quite correct, however, as there is an additional term coming from the terms in the sum corresponding to the special values, $j=i$, $j=i+1$ and $j=i-1$. We have tacitly assumed from the appearance of the expression (31) that, for a given i, the contribution from just three such special terms is of order ϵ^2. But this is not true. Although the expected contribution of a term like $(x_{i+1}-x_i)(x_{j+1}-x_j)$ for $j\neq i$ is indeed of order ϵ^2, the expected contribution of $(x_{i+1}-x_i)^2$ is $+i\epsilon m^{-1}$ [**C**, Eq. (50)], that is, of order ϵ. In non-relativistic mechanics the velocities are unlimited and in very short times ϵ the amplitude diffuses a distance proportional to the square root of the time. Making use of this equation then we see that the additional contribution from these terms is essentially

$$im^{-1}\epsilon\sum_i \mathbf{A}(\mathbf{x}_i, t_i)\cdot\mathbf{B}(\mathbf{x}_i, t_i) = im^{-1}\int \mathbf{A}(\mathbf{x}(t), t)\cdot\mathbf{B}(\mathbf{x}(t), t)dt$$

when summed on all i. This has the same effect as a first-order perturbation due to a potential $\mathbf{A}\cdot\mathbf{B}/m$. Added to the term involving the momentum operators

we therefore have an additional term[12]

$$\frac{i}{m}\int_{t'}^{t''} dt \int \chi^*(\mathbf{x}_{t''})K(\mathbf{x}_{t''}, t''; \mathbf{x}_t, t)\mathbf{A}(\mathbf{x}_t, t)\cdot\mathbf{B}(\mathbf{x}_t, t) \cdot K(\mathbf{x}_t, t; \mathbf{x}_{t'}, t')\psi(\mathbf{x}_{t'})d^3\mathbf{x}_{t''}d^3\mathbf{x}_t d^3\mathbf{x}_{t'}. \quad (33)$$

In the usual Hamiltonian theory this term arises, of course, from the term $\mathbf{A}^2/2m$ in the expansion of the Hamiltonian

$$H=(2m)^{-1}(\mathbf{p}-\mathbf{A})^2=(2m)^{-1}(\mathbf{p}^2-\mathbf{p}\cdot\mathbf{A}-\mathbf{A}\cdot\mathbf{p}+\mathbf{A}^2)$$

while the other term arises from the second-order action of $\mathbf{p}\cdot\mathbf{A}+\mathbf{A}\cdot\mathbf{p}$. We shall not be interested in non-relativistic quantum electrodynamics in detail. The situation is simpler for Dirac electrons. For particles satisfying the Klein-Gordon equation (discussed in Appendix A) the situation is very similar to a four-dimensional analog of the non-relativistic case given here.

6. EXTENSION TO DIRAC PARTICLES

Expressions (24) and (25) and their proof can be readily generalized to the relativistic case according to the one electron theory of Dirac. We shall discuss the hole theory later. In the non-relativistic case we began with the proposition that the amplitude for a particle to proceed from one point to another is the sum over paths of $\exp(iS_p)$, that is, we have for example for a transition element

$$\langle\chi|1|\psi\rangle=\lim_{\epsilon\to 0}\int\cdots\int\chi^*(\mathbf{x}_N)\Phi_p(\mathbf{x}_N, \mathbf{x}_{N-1}, \cdots\mathbf{x}_0) \cdot\psi(\mathbf{x}_0)d^3\mathbf{x}_0 d^3\mathbf{x}_1\cdots d^3\mathbf{x}_N \quad (34)$$

where for $\exp(iS_p)$ we have written Φ_p, that is more precisely,

$$\Phi_p=\Pi_i A^{-1}\exp iS(\mathbf{x}_{i+1}, \mathbf{x}_i).$$

As discussed in C this form is related to the usual form of quantum mechanics through the observation that

$$(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon=A^{-1}\exp[iS(\mathbf{x}_{i+1}, \mathbf{x}_i)] \quad (35)$$

where $(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon$ is the transformation matrix from a representation in which $\mathbf{x}$ is diagonal at time t_i to one in which $\mathbf{x}$ is diagonal at time $t_{i+1}=t_i+\epsilon$ (so that it is identical to $K_0(\mathbf{x}_{i+1}, t_{i+1}; \mathbf{x}_i, t_i)$ for the small time interval ϵ). Hence the amplitude for a given path can also be written

$$\Phi_p=\Pi_i(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon \quad (36)$$

for which form, of course, (34) is exact irrespective of whether $(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon$ can be expressed in the simple form (35).

For a Dirac electron the $(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon$ is a 4×4 matrix (or $4^N\times4^N$ if we deal with N electrons) but the expression (34) with (36) is still correct (as it is in fact for any quantum-mechanical system with a sufficiently general definition of the coordinate $\mathbf{x}_i$). The product (36) now involves operators, the order in which the factors are to be taken is the order in which the terms appear in time.

For a Dirac particle in a vector and scalar potential (times the electron charge e) $\mathbf{A}(\mathbf{x}, t)$, $A_4(\mathbf{x}, t)$, the quantity $(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon^{(A)}$ is related to that of a free particle to the first order in ϵ as

$$(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon^{(A)}=(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon^{(0)}\exp[-i(\epsilon A_4(\mathbf{x}_i, t_i) -(\mathbf{x}_{i+1}-\mathbf{x}_i)\cdot\mathbf{A}(\mathbf{x}_i, t_i))]. \quad (37)$$

This can be verified directly by substitution into the Dirac equation.[13] It neglects the variation of $\mathbf{A}$ and A_4 with time and space during the short interval ϵ. This produces errors only of order ϵ^2 in the Dirac case for the expected square velocity $(\mathbf{x}_{i+1}-\mathbf{x}_i)^2/\epsilon^2$ during the interval ϵ is finite (equaling the square of the velocity of light) rather than being of order $1/\epsilon$ as in the non-relativistic case. [This makes the relativistic case somewhat simpler in that it is not necessary to define the velocity as carefully as in (6); $(\mathbf{x}_{i+1}-\mathbf{x}_i)/\epsilon$ is sufficiently exact, and no term analogous to (33) arises.]

Thus $\Phi_p^{(A)}$ differs from that for a free particle, $\Phi_p^{(0)}$, by a factor $\Pi_i\exp-i(\epsilon A_4(\mathbf{x}_i, t_i)-(\mathbf{x}_{i+1}-\mathbf{x}_i)\cdot\mathbf{A}(\mathbf{x}_i, t_i))$ which in the limit can be written as

$$\exp\left\{-i\int[A_4(\mathbf{x}(t), t)-\mathbf{x}^\cdot(t)\cdot\mathbf{A}(\mathbf{x}(t), t)]dt\right\} \quad (38)$$

exactly as in the non-relativistic case.

The case of a Dirac particle interacting with the quantum-mechanical oscillators representing the field may now be studied. Since the dependence of $\Phi_p^{(A)}$ on $\mathbf{A}$, A_4 is through the same factor as in the non-relativistic case, when $\mathbf{A}$, A_4 are expressed in terms of the oscillator coordinates q, the dependence of Φ on the oscillator coordinates q is unchanged. Hence the entire analysis of the preceding sections which concern the results of the integration over oscillator coordinates can be carried through unchanged and the results will be expression (25) with formula (24) for R. Expression (25) is now interpreted as

$$\langle\chi_{t''}|\exp iR|\psi_{t'}\rangle=\lim_{\epsilon\to 0}\int\chi^*(\mathbf{x}_{t''}{}^{(1)}, \mathbf{x}_{t''}{}^{(2)}\cdots) \times\prod_n(\Phi_{p,n}{}^{(0)}d^3\mathbf{x}_{t''}{}^{(n)}d^3\mathbf{x}_{t''-\epsilon}{}^{(n)}\cdots d^3\mathbf{x}_{t'}{}^{(n)}) \cdot\exp(iR)\psi(\mathbf{x}_{t'}{}^{(1)}, \mathbf{x}_{t'}{}^{(2)}\cdots) \quad (39)$$

[12] The term corresponding to this for the self-energy expression (26) would give an integral over $\delta_+((t-t)^2-(\mathbf{x}_t-\mathbf{x}_t)^2)$ which is evidently infinite and leads to the quadratically divergent self-energy. There is no such term for the Dirac electron, but there is for Klein-Gordon particles. We shall not discuss the infinities in this paper as they have already been discussed in II.

[13] Alternatively, note that Eq. (37) is exact for arbitrarily large ϵ if the potential A_μ is constant. For if the potential in the Dirac equation is the gradient of a scalar function $A_\mu=\partial\chi/\partial x_\mu$ the potential may be removed by replacing the wave function by $\psi=\exp(-i\chi)\psi'$ (gauge transformation). This alters the kernel by a factor $\exp[-i(\chi(2)-\chi(1))]$ owing to the change in the initial and final wave functions. A constant potential A_μ is the gradient of $\chi=A_\mu x_\mu$ and can be completely removed by this gauge transformation, so that the kernel differs from that of a free particle by the factor $\exp[-i(A_\mu x_{\mu 2}-A_\mu x_{\mu 1})]$ as in (37).

where $\Phi_{p,n}^{(0)}$, the amplitude for a particular path for particle n is simply the expression (36) where $(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon$ is the kernel $K_{0,n}(\mathbf{x}_{i+1}^{(n)}, t_{i+1}; \mathbf{x}_i^{(n)}, t_i)$ for a free electron according to the one electron Dirac theory, with the matrices which appear operating on the spinor indices corresponding to particle (n) and the order of all operations being determined by the time indices.

For calculational purposes we can, as before, expand R as a power series and evaluate the various terms in the same manner as for the non-relativistic case. In such an expansion the quantity $\mathbf{x}^{\cdot}(t)$ is replaced, as we have seen in (32), by the operator $i(H\mathbf{x}-\mathbf{x}H)$, that is, in this case by α operating at the corresponding time. There is no further complicated term analogous to (33) arising in this case, for the expected value of $(x_{i+1}-x_i)^2$ is now of order ϵ^2 rather than ϵ.

For example, for self-energy one sees that expression (28) will be (with other terms coming from those with $\mathbf{x}^{\cdot}(t)$ replaced by α and with the usual β in back of each K_0 because of the definition of K_0 in relativity theory)

$$\langle\psi_{t''}|R|\psi_{t'}\rangle_{S_p} = -e^2\int\psi^*(\mathbf{x}_{t''})K_0(\mathbf{x}_{t''}, t''; \mathbf{x}_t, t)\beta\alpha_\mu \cdot\delta_+((t-s)^2-(\mathbf{x}_t-\mathbf{x}_s)^2)K_0(\mathbf{x}_t, t; \mathbf{x}_s, s)\beta\alpha_\mu \cdot K_0(\mathbf{x}_s, s; \mathbf{x}_{t'}, t')\beta\psi(\mathbf{x}_{t'})d^3\mathbf{x}_{t''}d^3\mathbf{x}_t d^3\mathbf{x}_s d^3\mathbf{x}_{t'}dtds, \quad (40)$$

where $\alpha_4=1$, $\alpha_{1,2,3}=\alpha_{x,y,z}$ and a sum on the repeated index μ is implied in the usual way; $a_\mu b_\mu = a_4b_4-a_1b_1-a_2b_2-a_3b_3$. One can change $\beta\alpha_\mu$ to γ_μ and ψ^* to $\bar{\psi}\beta$. In this manner all of the rules referring to virtual photons discussed in II are deduced; but with the difference that K_0 is used instead of K_+ and we have the Dirac one electron theory with negative energy states (although we may have any number of such electrons).

7. EXTENSION TO POSITRON THEORY

Since in (39) we have an arbitrary number of electrons, we can deal with the hole theory in the usual manner by imagining that we have an infinite number of electrons in negative energy states.

On the other hand, in paper I on the theory of positrons, it was shown that the results of the hole theory in a system with a given external potential A_μ were equivalent to those of the Dirac one electron theory if one replaced the propagation kernel, K_0, by a different one, K_+, and multiplied the resultant amplitude by factor C_v involving A_μ. We must now see how this relation, derived in the case of external potentials, can also be carried over in electrodynamics to be useful in simplifying expressions involving the infinite sea of electrons.

To do this we study in greater detail the relation between a problem involving virtual photons and one involving purely external potentials. In using (25) we shall assume in accordance with the hole theory that the number of electrons is infinite, but that they all have the same charge, e. Let the states $\psi_{t'}$, $\chi_{t''}$, represent the vacuum plus perhaps a number of real electrons in positive energy states and perhaps also some empty negative energy states. Let us call the amplitude for the transition in an external potential B_μ, but *excluding virtual photons*, $T_0[B]$, a functional of $B_\mu(1)$. We have seen (38)

$$T_0[B]=\langle\chi_{t''}|\exp iP|\psi_{t'}\rangle \quad (41)$$

where

$$P=-\sum_n\int[B_4(\mathbf{x}^{(n)}(t), t)-\mathbf{x}^{\cdot(n)}(t)\cdot\mathbf{B}(\mathbf{x}^{(n)}(t), t)]dt$$

by (38). We can write this as

$$P=-\sum_n\int B_\mu(x_\nu^{(n)}(t))\dot{x}_\mu^{(n)}(t)dt$$

where $x_4(t)=t$ and $\dot{x}_4=1$, the other values of μ corresponding to space variables. The corresponding amplitude for the same process in the same potential, but *including* all the virtual photons we may call,

$$T_{e^2}[B]=\langle\chi_{t''}|\exp(iR)\exp(iP)|\psi_{t'}\rangle. \quad (42)$$

Now let us consider the effect on $T_{e^2}[B]$ of changing the coupling e^2 of the virtual photons. Differentiating (42) with respect to e^2 which appears only [14] in R we find

$$dT_{e^2}[B]/d(e^2)=\left\langle\chi_{t''}\middle|-\frac{i}{2}\sum_n\sum_m\int\int dtds\dot{x}_\mu^{(n)}(t)\dot{x}_\mu^{(m)}(s) \cdot\delta_+((x_\nu^{(n)}(t)-x_\nu^{(m)}(s))^2)\exp i(R+P)\middle|\psi_{t'}\right\rangle. \quad (43)$$

We can also study the first-order effect of a change of B_μ:

$$\delta T_{e^2}[B]/\delta B_\mu(1)=-i\left\langle\chi_{t''}\middle|\sum_n\int dt\dot{x}_\mu^{(n)}\delta^4(x_\alpha^{(n)}(t)-x_{\alpha,1}) \cdot\exp i(R+P)\middle|\psi_{t'}\right\rangle \quad (44)$$

where $x_{\alpha,1}$ is the field point at which the derivative with respect to B_μ is taken[15] and the term (current density) $-\sum_n\int dt\dot{x}_\mu^{(n)}(t)\delta^4(x_\alpha^{(n)}(t)-x_{\alpha,1})$ is just $\delta P/\delta B_\mu(1)$. The function $\delta^4(x_\alpha^{(n)}-x_{\alpha,1})$ means $\delta(x_4^{(n)}-x_{4,1})$

[14] In changing the charge e^2 we mean to vary only the degree to which virtual photons are important. We do not contemplate changes in the influence of the external potentials. If one wishes, as e is raised the strength of the potential is decreased proportionally so that B_μ, the potential times the charge e, is held constant.

[15] The functional derivative is defined such that if $T[B]$ is a number depending on the functions $B_\mu(1)$, the first order variation in T produced by a change from B_μ to $B_\mu+\Delta B_\mu$ is given by

$$T[B+\Delta B]-T[B]=\int(\delta T[B]/\delta B_\mu(1))\Delta B_\mu(1)d\tau_1$$

the integral extending over all four-space $x_{\alpha,1}$.

$\times\delta(x_3^{(n)}-x_{3,1})\delta(x_2^{(n)}-x_{2,1})\delta(x_1^{(n)}-x_{1,1})$ that is, $\delta(2,1)$ with $x_{\alpha,2}=x_\alpha^{(n)}(t)$. A second variation of T gives, by differentiation of (44) with respect to $B_\nu(2)$,

$$\delta^2T_{e2}[B]/\delta B_\mu(1)\delta B_\nu(2)$$
$$=-\left\langle\chi_{t'}\left|\sum_n\sum_m\int dt\int ds\dot{x}_\mu^{(n)}(t)\dot{x}_\nu^{(m)}(s)\right.\right.$$
$$\cdot\delta^4(x_\alpha^{(n)}(t)-x_{\alpha,1})\delta^4(x_\beta^{(m)}(s)-x_{\beta,2})$$
$$\left.\left.\times\exp i(R+P)\right|\psi_{t'}\right\rangle.$$

Comparison of this with (43) shows that

$$dT_{e2}[B]/d(e^2)=\tfrac{1}{2}i\int\int(\delta^2T_{e2}[B]/\delta B_\mu(1)\delta B_\mu(2))$$
$$\times\delta_+(s_{12}^2)d\tau_1d\tau_2 \quad (45)$$

where $s_{12}^2=(x_{\mu,1}-x_{\mu,2})(x_{\mu,1}-x_{\mu,2})$.

We now proceed to use this equation to prove the validity of the rules given in **II** for electrodynamics. This we do by the following argument. The equation can be looked upon as a differential equation for $T_{e2}[B]$. It determines $T_{e2}[B]$ uniquely if $T_0[B]$ is known. We have shown it is valid for the hole theory of positrons. But in **I** we have given formulas for calculating $T_0[B]$ whose correctness relative to the hole theory we have there demonstrated. Hence we have shown that the $T_{e2}[B]$ obtained by solving (45) with the initial condition $T_0[B]$ as given by the rules in **I** will be equal to that given for the same problem by the second quantization theory of the Dirac matter field coupled with the quantized electromagnetic field. But it is evident (the argument is given in the next paragraph) that the rules[16] given in **II** constitute a solution in power series in e^2 of the Eq. (45) [which for $e^2=0$ reduce to the $T_0[B]$ given in **I**]. Hence the rules in **II** must give, to each order in e^2, the matrix element for any process that would be calculated by the usual theory of second quantization of the matter and electromagnetic fields. This is what we aimed to prove.

That the rules of **II** represent, in a power series expansion, a solution of (45) is clear. For the rules there given may be stated as follows: Suppose that we have a process to order k in e^2 (i.e., having k virtual photons) and order n in the external potential B_μ. Then, *the matrix element for the process with one more virtual photon and two less potentials is that obtained from the previous matrix by choosing from the n potentials a pair, say $B_\mu(1)$ acting at 1 and $B_\nu(2)$ acting at 2, replacing them by $ie^2\delta_{\mu\nu}\delta_+(s_{12}^2)$, adding the results for each way of choosing the pair, and dividing by $k+1$, the present number of photons.* The matrix with no virtual photons ($k=0$) being given to any n by the rules of **I**, this permits terms to all orders in e^2 to be derived by recursion. It is evident that the rule in italics is that of **II**, and equally evident that it is a word expression of Eq. (45). [The factor $\frac{1}{2}$ in (45) arises since in integrating over all $d\tau_1$ and $d\tau_2$ we count each pair twice. The division by $k+1$ is required by the rules of **II** for, there, each diagram is to be taken only once, while in the rule given above we say what to do to add one extra virtual photon to k others. But which one of the $k+1$ is to be identified at the last photon added is irrelevant. It agrees with (45) of course for it is canceled on differentiating with respect to e^2 the factor $(e^2)^{k+1}$ for the $(k+1)$ photons.]

8. GENERALIZED FORMULATION OF QUANTUM ELECTRODYNAMICS

The relation implied by (45) between the formal solution for the amplitude for a process in an arbitrary unquantized external potential to that in a quantized field appears to be of much wider generality. We shall discuss the relation from a more general point of view here (still limiting ourselves to the case of no photons in initial or final state).

In earlier sections we pointed out that as a consequence of the Lagrangian form of quantum mechanics the aspects of the particles' motions and the behavior of the field could be analyzed separately. What we did was to integrate over the field oscillator coordinates first. We could, in principle, have integrated over the particle variables first. That is, we first solve the problem with the action of the particles and their interaction with the field and then multiply by the exponential of the action of the field and integrate over all the field oscillator coordinates. (For simplicity of discussion let us put aside from detailed special consideration the questions involving the separation of the longitudinal and transverse parts of the field.[9]) Now the integral over the particle coordinates for a given process is precisely the integral required for the analysis of the motion of the particles in an unquantized potential. With this observation we may suggest a generalization to all types of systems.

Let us suppose the formal solution for the amplitude for some given process with matter in an external potential $B_\mu(1)$ is some numerical quantity T_0. We mean matter in a more general sense now, for the motion of the matter may be described by the Dirac equation, or by the Klein-Gordon equation, or may involve charged or neutral particles other than electrons and positrons in any manner whatsoever. The quantity T_0 depends of course on the potential function $B_\mu(1)$; that is, it is a functional $T_0[B]$ of this potential. We

[16] That is, of course, those rules of **II** which apply to the unmodified electrodynamics of Dirac electrons. (The limitation excluding real photons in the initial and final states is removed in Sec. 8.) The same arguments clearly apply to nucleons interacting via neutral vector mesons, vector coupling. Other couplings require a minor extension of the argument. The modification to the $(\mathbf{x}_{i+1}|\mathbf{x}_i)_\epsilon$, as in (37), produced by some couplings cannot very easily be written without using operators in the exponents. These operators can be treated as numbers if their order of operation is maintained to be always their order in time. This idea will be discussed and applied more generally in a succeeding paper.

assume we have some expression for it in terms of B_μ (exact, or to some desired degree of approximation in the strength of the potential).

Then the answer $T_{e2}[B]$ to the corresponding problem in quantum electrodynamics is $T_0[A_\mu(1)+B_\mu(1)] \times \exp(iS_0)$ summed over all possible distributions of field $A_\mu(1)$, wherein S_0 is the action for the field $S_0 = -(8\pi e^2)^{-1}\sum_\mu \int((\partial A_\mu/\partial t)^2-(\nabla A_\mu)^2)d^3\mathbf{x}dt$ the sum on μ carrying the usual minus sign for space components.

If $F[A]$ is any functional of $A_\mu(1)$ we shall represent by ${}_0|F[A]|_0$ this superposition of $F[A]\exp(iS_0)$ over distributions of A_μ for the case in which there are no photons in initial or final state. That is, we have

$$T_{e2}[B] = {}_0|T_0[A+B]|_0. \tag{46}$$

The evaluation of ${}_0|F[A]|_0$ directly from the definition of the operation ${}_0|\ \ |_0$ is not necessary. We can give the result in another way. We first note that the operation is linear,

$${}_0|F_1[A]+F_2[A]|_0 = {}_0|F_1[A]|_0 + {}_0|F_2[A]|_0 \tag{47}$$

so that if F is represented as a sum of terms each term can be analyzed separately. We have studied essentially the case in which $F[A]$ is an exponential function. In fact, what we have done in Section 4 may be repeated with slight modification to show that

$$\left|\exp\left(-i\int j_\mu(1)A_\mu(1)d\tau_1\right)\right|_0 = \exp\left(-\tfrac{1}{2}ie^2\int\int j_\mu(1)j_\mu(2)\delta_+(s_{12}^2)d\tau_1 d\tau_2\right) \tag{48}$$

where $j_\mu(1)$ is an arbitrary function of position and time for each value of μ.

Although this gives the evaluation of ${}_0|\ \ |_0$ for only a particular functional of A_μ the appearance of the arbitrary function $j_\mu(1)$ makes it sufficiently general to permit the evaluation for any other functional. For it is to be expected that any functional can be represented as a superposition of exponentials with different functions $j_\mu(1)$ (by analogy with the principle of Fourier integrals for ordinary functions). Then, by (47), the result of the operation is the corresponding superposition of expressions equal to the right-hand side of (48) with the various j's substituted for j_μ.

In many applications $F[A]$ can be given as a power series in A_μ:

$$F[A] = f_0 + \int f_\mu(1)A_\mu(1)d\tau_1 + \int\int f_{\mu\nu}(1,2)A_\mu(1)A_\nu(2)d\tau_1 d\tau_2 + \cdots \tag{49}$$

where f_0, $f_\mu(1)$, $f_{\mu\nu}(1,2)\cdots$ are known numerical functions independent of A_μ. Then by (47)

$${}_0|F[A]|_0 = f_0 + \int f_\mu(1){}_0|A_\mu(1)|_0 d\tau_1 + \int\int f_{\mu\nu}(1,2){}_0|A_\mu(1)A_\nu(2)|_0 d\tau_1 d\tau_2 + \cdots \tag{50}$$

where we set ${}_0|1|_0 = 1$ (from (48) with $j_\mu = 0$). We can work out expressions for the successive powers of A_μ by differentiating both sides of (48) successively with respect to j_μ and setting $j_\mu = 0$ in each derivative. For example, the first variation (derivative) of (48) with respect to $j_\mu(3)$ gives

$${}_0\left|-iA_\mu(3)\exp\left(-i\int j_\nu(1)A_\nu(1)d\tau_1\right)\right|_0 = -ie^2\int \delta_+(s_{31}^2)j_\mu(4)d\tau_4 \times \exp\left(-\tfrac{1}{2}ie^2\int\int j_\nu(1)j_\nu(2)\delta_+(s_{12}^2)d\tau_1 d\tau_2\right). \tag{51}$$

Setting $j_\mu = 0$ gives

$${}_0|A_\mu(3)|_0 = 0.$$

Differentiating (51) again with respect to $j_\nu(4)$ and setting $j_\nu = 0$ shows

$${}_0|A_\mu(3)A_\nu(4)|_0 = ie^2\delta_{\mu\nu}\delta_+(s_{34}^2) \tag{52}$$

and so on for higher powers. These results may be substituted into (50). Clearly therefore when $T_0[B+A]$ in (46) is expanded in a power series and the successive terms are computed in this way, we obtain the results given in **II**.

It is evident that (46), (47), (48) imply that $T_{e2}[B]$ satisfies the differential equation (45) and conversely (45) with the definition (46) implies (47) and (48). For if $T_0[B]$ is an exponential

$$T_0[B] = \exp\left(-i\int j_\mu(1)B_\mu(1)d\tau_1\right) \tag{53}$$

we have from (46), (48) that

$$T_{e2}[B] = \exp\left[-\tfrac{1}{2}ie^2\int\int j_\mu(1)j_\mu(2)\delta_+(s_{12}^2)d\tau_1 d\tau_2\right] \cdot \exp\left[-i\int j_\nu(1)B_\nu(1)d\tau_1\right]. \tag{54}$$

Direct substitution of this into Eq. (45) shows it to be a solution satisfying the boundary condition (53). Since the differential equation (45) is linear, if $T_0[B]$ is a superposition of exponentials, the corresponding superposition of solutions (54) is also a solution.

Many of the formal representations of the matter system (such as that of second quantization of Dirac electrons) represent the interaction with a fixed potential in a formal exponential form such as the left-hand side of (48), except that $j_\mu(1)$ is an operator instead of a numerical function. Equation (48) may still be used if care is exercised in defining the order of the operators on the right-hand side. The succeeding paper will discuss this in more detail.

Equation (45) or its solution (46), (47), (48) constitutes a very general and convenient formulation of the laws of quantum electrodynamics for virtual processes. Its relativistic invariance is evident if it is assumed that the unquantized theory giving $T_0[B]$ is invariant. It has been proved to be equivalent to the usual formulation for Dirac electrons and positrons (for Klein-Gordon particles see Appendix A). It is suggested that it is of wide generality. It is expressed in a form which has meaning even if it is impossible to express the matter system in Hamiltonian form; in fact, it only requires the existence of an amplitude for fixed potentials which obeys the principle of superposition of amplitudes. If $T_0[B]$ is known in power series in B, calculations of $T_{e^2}[B]$ in a power series of e^2 can be made directly using the italicized rule of Sec. 7. The limitation to virtual quanta is removed in the next section.

On the other hand, the formulation is unsatisfactory because for situations of importance it gives divergent results, even if $T_0[B]$ is finite. The modification proposed in **II** of replacing $\delta_+(s_{12}^2)$ in (45), (48) by $f_+(s_{12}^2)$ is not satisfactory owing to the loss of the theorems of conservation of energy or probability discussed in **II** at the end of Sec. 6. There is the additional difficulty in positron theory that even $T_0[B]$ is infinite to begin with (vacuum polarization). Computational ways of avoiding these troubles are given in **II** and in the references of footnote 2.

9. CASE OF REAL PHOTONS

The case in which there are real photons in the initial or the final state can be worked out from the beginning in the same manner.[17] We first consider the case of a system interacting with a single oscillator. From this result the generalization will be evident. This time we shall calculate the transition element between an initial state in which the particle is in state $\psi_{t'}$ and the oscillator is in its nth eigenstate (i.e., there are n photons in the field) to a final state with particle in $\chi_{t''}$, oscillator in mth level. As we have already discussed, when the coordinates of the oscillator are eliminated the result is the transition element $\langle\chi_{t''}|G_{mn}|\psi_{t'}\rangle$ where

$$G_{mn}=\int \varphi_m^*(q_j)k(q_j,t'';q_0,t')\varphi_n(q_0)dq_0dq_j \quad (11)$$

where φ_m, φ_n are the wave functions[8] for the oscillator in state m, n and k is given in (12). The G_{mn} can be evaluated most easily by calculating the generating function

$$g(X,Y)=\sum_m\sum_n G_{mn}X^mY^n(m!n!)^{-\frac{1}{2}} \quad (55)$$

for arbitrary X, Y. If expression (11) is substituted in the left-hand side of (55), the expression can be simplified by use of the generating function relation for the eigenfunctions[8] of the harmonic oscillator

$$\sum_n \varphi_n(q_0)Y^n(n!)^{-\frac{1}{2}}=(\omega/\pi)^{\frac{1}{4}}\exp(-\tfrac{1}{2}i\omega t')\times\exp\tfrac{1}{2}[\omega q_0^2-(Y\exp[-i\omega t']-(2\omega)^{\frac{1}{2}}q_0)^2]$$

Using a similar expansion for the φ_m^* one is left with the exponential of a quadratic function of q_0 and q_j. The integration on q_0 and q_j is then easily performed to give

$$g(X,Y)=G_{00}\exp(XY+i\beta^*X+i\beta Y) \quad (56)$$

from which expansion in powers of X and Y and comparison to (11) gives the final result

$$G_{mn}=G_{00}(m!n!)^{-\frac{1}{2}}\sum_r \frac{m!}{(m-r)!r!}\frac{n!}{(n-r)!r!}\times r!(i\beta^*)^{m-r}(i\beta)^{n-r} \quad (57)$$

where G_{00} is given in (14) and

$$\beta=(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\gamma(t)\exp(-i\omega t)dt,$$
$$\beta^*=(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\gamma(t)\exp(+i\omega t)dt, \quad (58)$$

and the sum on r is to go from 0 to m or to n whichever is the smaller. (The sum can be expressed as a Laguerre polynomial but there is no advantage in this.)

Formula (57) is readily understandable. Consider first a simple case of absorption of one photon. Initially we have one photon and finally none. The amplitude for this is the transition element of $G_{01}=i\beta G_{00}$ or $\langle\chi_{t''}|i\beta G_{00}|\psi_{t'}\rangle$. This is the same as would result if we asked for the transition element for a problem in which all photons are virtual but there was present a perturbing potential $-(2\omega)^{-\frac{1}{2}}\gamma(t)\exp(-i\omega t)$ and we required the first-order effect of this potential. Hence photon absorption is like the first order action of a potential varying in time as $\gamma(t)\exp(-i\omega t)$ that is with a positive frequency (i.e., the sign of the coefficient of t in the exponential corresponds to positive energy). The amplitude for emission of one photon involves $G_{10}=i\beta^*G_{00}$, which is the same result except that the potential has negative frequency. Thus we begin by interpreting $i\beta^*$ as the amplitude for emission of one photon $i\beta$ as the amplitude for absorption of one.

Next for the general case of n photons initially and m finally we may understand (57) as follows. We first

[17] For an alternative method starting directly from the formula (24) for virtual photons, see Appendix B.

neglect Bose statistics and imagine the photons as individual distinct particles. If we start with n and end with m this process may occur in several different ways. The particle may absorb in total $n-r$ of the photons and the final m photons will represent r of the photons which were present originally plus $m-r$ new photons emitted by the particle. In this case the $n-r$ which are to be absorbed may be chosen from among the original n in $n!/(n-r)!r!$ different ways, and each contributes a factor $i\beta$, the amplitude for absorption of a photon. Which of the $m-r$ photons from among the m are emitted can be chosen in $m!/(m-r)!r!$ different ways and each photon contributes a factor $i\beta^*$ in amplitude. The initial r photons which do not interact with the particle can be re-arranged among the final r in $r!$ ways. We must sum over the alternatives corresponding to different values of r. Thus the form of G_{mn} can be understood. The remaining factor $(m!)^{-\frac{1}{2}}(n!)^{-\frac{1}{2}}$ may be interpreted as saying that in computing probabilities (which therefore involves the square of G_{mn}) the photons may be considered as independent but that if m are actually equal the statistical weight of each of the states which can be made by rearranging the m equal photons is only $1/m!$. This is the content of Bose statistics; that m equal particles in a given state represents just one state, i.e., has statistical weight unity, rather than the $m!$ statistical weight which would result if it is imagined that the particles and states can be identified and rearranged in $m!$ different ways. This holds for both the initial and final states of course. From this rule about the statistical weights of states the derivation of the blackbody distribution law follows.

The actual electromagnetic field is represented as a host of oscillators each of which behaves independently and produces its own factor such as G_{mn}. Initial or final states may also be linear combinations of states in which one or another oscillator is excited. The results for this case are of course the corresponding linear combination of transition elements.

For photons of a given direction of polarization and for sin or cos waves the explicit expression for β can be obtained directly from (58) by substituting the formulas (16) for the γ's for the corresponding oscillator. It is more convenient to use the linear combination corresponding to running waves. Thus we find the amplitude for absorption of a photon of momentum $\mathbf{K}$, frequency $k=(\mathbf{K}\cdot\mathbf{K})^{\frac{1}{2}}$ polarized in direction $\mathbf{e}$ is given by including a factor i times

$$\beta_{\mathbf{K},\mathbf{e}}=(4\pi)^{\frac{1}{2}}(2k)^{-\frac{1}{2}}\sum_n e_n\int_{t'}^{t''}\exp(-ikt)\times\exp(i\mathbf{K}\cdot\mathbf{x}_n(t))\mathbf{e}\cdot\dot{\mathbf{x}}_n(t)dt \quad (59)$$

in the transition element (25). The density of states in momentum space is now $(2\pi)^{-3}d^3\mathbf{K}$. The amplitude for emission is just i times the complex conjugate of this expression, or what amounts to the same thing, the same expression with the sign of the four vector k_μ reversed. Since the factor (59) is exactly the first-order effect of a vector potential

$$\mathbf{A}^{PH}=(2\pi/k)^{\frac{1}{2}}\mathbf{e}\exp(-i(kt-\mathbf{K}\cdot\mathbf{x}))$$

of the corresponding classical wave, we have derived the rules for handling real photons discussed in **II**.

We can express this directly in terms of the quantity $T_{\epsilon 2}[B]$, the amplitude for a given transition without emission of a photon. What we have said is that the amplitude for absorption of just one photon whose classical wave form is $A_\mu^{PH}(1)$ (time variation $\exp(-ikt_1)$ corresponding to positive energy k) is proportional to the first order (in ϵ) change produced in $T_{\epsilon 2}[B]$ on changing B to $B+\epsilon A^{PH}$. That is, more exactly,

$$\int(\delta T_{\epsilon 2}[B]/\delta B_\mu(1))A_\mu^{PH}(1)d\tau_1 \quad (60)$$

is the amplitude for absorption by the particle system of one photon, A^{PH}. (A superposition argument shows the expression to be valid not only for plane waves, but for spherical waves, etc., as given by the form of A^{PH}.) The amplitude for emission is the same expression but with the sign of the frequency reversed in A^{PH}. The amplitude that the system absorbs two photons with waves $A_\mu^{PH_1}$ and $A_\nu^{PH_2}$ is obtained from the next derivative,

$$\int\int(\delta^2 T_{\epsilon 2}[B]/\delta B_\mu(1)\delta B_\nu(2))A_\mu^{PH_1}(1)A_\nu^{PH_2}(2)d\tau_1 d\tau_2,$$

the same expression holding for the absorption of one and emission of the other, or emission of both depending on the sign of the time dependence of A^{PH_1} and A^{PH_2}. Larger photon numbers correspond to higher derivatives, absorption of l_1 emission of l_2 requiring the (l_1+l_2) derivatics. When two or more of the photons are exactly the same (e.g., $A^{PH_1}=A^{PH_2}$) the same expression holds for the amplitude that l_1 are absorbed by the system while l_2 are emitted. However, the statement that initially n of a kind are present and m of this kind are present finally, does not imply $l_1=n$ and $l_2=m$. It is possible that only $n-r=l_1$ were absorbed by the system and $m-r=l_2$ emitted, and that r remained from initial to final state without interaction. This term is weighed by the combinatorial coefficient $(m!n!)^{-\frac{1}{2}}\binom{m}{r}\binom{n}{r}r!$ and summed over the possibilities for r as explained in connection with (57). Thus once the amplitude for virtual processes is known, that for real photon processes can be obtained by differentiation.

It is possible, of course, to deal with situations in which the electromagnetic field is not in a definite state after the interaction. For example, we might ask for the total probability of a given process, such as a scattering, without regard for the number of photons emitted. This is done of course by squaring the ampli-

tude for the emission of m photons of a given kind and summing on all m. Actually the sums and integrations over the oscillator momenta can usually easily be performed analytically. For example, the amplitude, starting from vacuum and ending with m photons of a given kind, is by (56) just

$$G_{m0}=(m!)^{-\frac{1}{2}}G_{00}(i\beta^*)^m. \tag{61}$$

The square of the amplitude summed on m requires the product of two such expressions (the $\gamma(t)$ in the β of one and in the other will have to be kept separately) summed on m:

$$\sum_m G_{m0}^*G_{m0}'=\sum_m G_{00}^*G_{00}'(m!)^{-1}\beta^m(\beta'^*)^m = G_{00}^*G_{00}'\exp(\beta\beta'^*).$$

In the resulting expression the sum over all oscillators is easily done. Such expressions can be of use in the analysis in a direct manner of problems of line width, of the Bloch-Nordsieck infra-red problem, and of statistical mechanical problems, but no such applications will be made here.

The author appreciates his opportunities to discuss these matters with Professor H. A. Bethe and Professer J. Ashkin, and the help of Mr. M. Baranger with the manuscript.

APPENDIX A. THE KLEIN-GORDON EQUATION

In this Appendix we describe a formulation of the equations for a particle of spin zero which was first used to obtain the rules given in II for such particles. The complete physical significance of the equations has not been analyzed thoroughly so that it may be preferable to derive the rules directly from the second quantization formulation of Pauli and Weisskopf. This can be done in a manner analogous to the derivation of the rules for the Dirac equation given in I or from the Schwinger-Tomonaga formulation[2] in a manner described, for example, by Rohrlich.[18] The formulation given here is therefore not necessary for a description of spin zero particles but is given only for its own interest as an alternative to the formulation of second quantization.

We start with the Klein-Gordon equation

$$(i\partial/\partial x_\mu-A_\mu)^2\psi=m^2\psi \tag{1A}$$

for the wave function ψ of a particle of mass m in a given external potential A_μ. We shall try to represent this in a manner analogous to the formulation of quantum mechanics in C. That is, we try to represent the amplitude for a particle to get from one point to another as a sum over all trajectories of an amplitude $\exp(iS)$ where S is the classical action for a given trajectory. To maintain the relativistic invariance in evidence the idea suggests itself of describing a trajectory in space-time by giving the four variables $x_\mu(u)$ as functions of some fifth parameter u (rather than expressing x_1, x_2, x_3 in terms of x_4). As we expect to represent paths which may reverse themselves in time (to represent pair production, etc., as in I) this is certainly a more convenient representation, for all four functions $x_\mu(u)$ may be considered as functions of a parameter u (somewhat analogous to proper time) which increase as we go along the trajectory, whether the trajectory is proceeding forward ($dx_4/du>0$) or backward ($dx_4/du<0$) in time.[19] We shall then have a new type of wave function $\varphi(x, u)$ a function of five variables, x standing for the four x_μ. It gives the amplitude for arrival at point x_μ with a certain value of the parameter u. We shall suppose that this wave function satisfies the equation

$$i\partial\varphi/\partial u=-\tfrac{1}{2}(i\partial/\partial x_\mu-A_\mu)^2\varphi \tag{2A}$$

which is seen to be analogous to the time-dependent Schrödinger equation, u replacing the time and the four coordinates of space-time x_μ replacing the usual three coordinates of space.

Since the potentials $A_\mu(x)$ are functions only of coordinates x_μ and are independent of u, the equation is separable in u and we can write a special solution in the form $\varphi=\exp(\frac{1}{2}im^2u)\psi(x)$ where $\psi(x)$, a function of the coordinates x_μ only, satisfies (1A) and the eigenvalue $\frac{1}{2}m^2$ conjugate to u is related to the mass m of the particle. Equation (2A) is therefore equivalent to the Klein-Gordon Eq. (1A) provided we ask in the end only for the solution of (1A) corresponding to the eigenvalue $\frac{1}{2}m^2$ for the quantity conjugate to u.

We may now proceed to represent Eq. (2A) in Lagrangian form in general and without regard to this eigenvalue condition. Only in the final solutions need we apply the eigenvalue condition. That is, if we have some special solution $\varphi(x, u)$ of (2A) we can select that part corresponding to the eigenvalue $\frac{1}{2}m^2$ by calculating

$$\psi(x)=\int_{-\infty}^{\infty}\exp(-\tfrac{1}{2}im^2u)\,\varphi(x, u)\,du$$

and thereby obtain a solution ψ of Eq. (1A).

Since (2A) is so closely analogous to the Schrödinger equation, it is easily written in the Lagrangian form described in C, simply by working by analogy. For example if $\varphi(x, u)$ is known at one value of u its value at a slightly larger value $u+\epsilon$ is given by

$$\varphi(x, u+\epsilon)=\int \exp i\epsilon\left[-\frac{(x_\mu-x_\mu')^2}{2\epsilon^2}-\frac{1}{2}\left(\frac{x_\mu-x_\mu'}{\epsilon}\right)(A_\mu(x)+A_\mu(x'))\right] \cdot\varphi(x', u)\,d^4\tau_{x'}(2\pi i\epsilon)^{-2}(-2\pi i\epsilon)^{-2} \tag{3A}$$

where $(x_\mu-x_\mu')^2$ means $(x_\mu-x_\mu')(x_\mu-x_\mu')$, $d^4\tau_{x'}=dx_1'dx_2'dx_3'dx_4'$ and the sign of the normalizing factor is changed for the x_4 component since the component has the reversed sign in its quadratic coefficient in the exponential, in accordance with our summation convention $a_\mu b_\mu=a_4b_4-a_1b_1-a_2b_2-a_3b_3$. Equation (3A), as can be verified readily as described in C, Sec. 6, is equivalent to first order in ϵ, to Eq. (2A). Hence, by repeated use of this equation the wave function at $u_0=n\epsilon$ can be represented in terms of that at $u=0$ by:

$$\varphi(x_{\nu, n}, u_0)=\int \exp-\frac{i\epsilon}{2}\sum_{i=1}^{n}\left[\left(\frac{x_{\mu, i}-x_{\mu, i-1}}{\epsilon}\right)^2 +\epsilon^{-1}(x_{\mu, i}-x_{\mu, i-1})(A_\mu(x_i)+A_\mu(x_{i-1}))\right] \cdot\varphi(x_{\nu, 0}, 0)\prod_{i=0}^{n-1}(d^4\tau_i/4\pi^2\epsilon^2 i). \tag{4A}$$

That is, roughly, the amplitude for getting from one point to another with a given value of u_0 is the sum over all trajectories of $\exp(iS)$ where

$$S=-\int_0^{u_0}[\tfrac{1}{2}(dx_\mu/du)^2+(dx_\mu/du)A_\mu(x)]du, \tag{5A}$$

when sufficient care is taken to define the quantities, as in C. This completes the formulation for particles in a fixed potential but a few words of description may be in order.

In the first place in the special case of a free particle we can define a kernal $k^{(0)}(x, u_0; x', 0)$ for arrival from x_μ', 0 to x_μ at u_0 as the sum over all trajectories between these points of $\exp-i\int_0^{u_0}\frac{1}{2}(dx_\mu/du)^2du$. Then for this case we have

$$\varphi(x, u_0)=\int k^{(0)}(x, u_0; x', 0)\varphi(x', 0)d^4\tau_{x'} \tag{6A}$$

and it is easily verified that k_0 is given by

$$k^{(0)}(x, u_0; x', 0)=(4\pi^2u_0^2i)^{-1}\exp-i(x_\mu-x_\mu')^2/2u_0 \tag{7A}$$

for $u_0>0$ and by 0, by definition, for $u_0<0$. The corresponding

[18] F. Rohrlich (to be published).

[19] The physical ideas involved in such a description are discussed in detail by Y. Nambu, Prog. Theor. Phys. 5, 82 (1950). An equation of type (2A) extended to the case of Dirac electrons has been studied by V. Fock, Physik Zeits. Sowjetunion 12, 404 (1937).

kernel of importance when we select the eigenvalue $\frac{1}{2}m^2$ is[20]

$$2iI_+(x,x')=\int_{-\infty}^{\infty}k^{(0)}(x,u_0;x',0)\exp(-\tfrac{1}{2}im^2u_0)du_0$$
$$=\int_0^{\infty}du_0(4\pi^2u_0^2i)^{-1}\exp-\tfrac{1}{2}i(m^2u_0+u_0^{-1}(x_\mu-x_\mu')^2) \quad (8A)$$

(the last extends only from $u_0=0$ since k_0 is zero for negative u_0) which is identical to the I_+ defined in II.[21] This may be seen readily by studying the Fourier transform, for the transform of the integrand on the right-hand side is

$$\int(4\pi^2u_0^2i)^{-1}\exp(ip\cdot x)\exp-\tfrac{1}{2}i(m^2u_0+x_\mu^2/u_0)d^4\tau_x$$
$$=\exp-\tfrac{1}{2}iu_0(m^2-p_\mu^2)$$

so that the u_0 integration gives for the transform of I_+ just $1/(p_\mu^2-m^2)$ with the pole defined exactly as in II. Thus we are automatically representing the positrons as trajectories with the time sense reversed.

If $\Phi^{(0)}[x(u)]=\exp-i\int_0^{u_0}\frac{1}{2}(dx_\mu/du)^2du$ is the amplitude for a given trajectory $x_\nu(u)$ for a free particle, then the amplitude in a potential is

$$\Phi^{(A)}[x(u)]=\Phi^{(0)}[x(u)]\exp-i\int_0^{u_0}(dx_\mu/du)A_\mu(x)du. \quad (9A)$$

If desired this may be studied by perturbation methods by expanding the exponential in powers of A_μ.

For interpretation, the integral in (9A) must be written as a Riemann sum, and if a perturbation expansion is made, care must be taken with the terms quadratic in the velocity, for the effect of $(x_{\mu,i+1}-x_{\mu,i})(x_{\nu,i+1}-x_{\nu,i})$ is not of order ϵ^2 but is $-i\delta_{\mu\nu}\epsilon$. The "velocity" dx_μ/du becomes the momentum operator $p_\mu=+i\partial/\partial x_\mu$ operating half before and half after A_μ, just as in the non-relativistic Schrödinger equation discussed in Sec. 5. Furthermore, in exactly the same manner as in that case, but here in four dimensions, a term quadratic in A_μ arises in the second-order perturbation terms from the coincidence of two velocities for the same value of u.

As an example, the kernal $k^{(A)}(x,u_0;x',0)$ for proceeding from $x_\mu',0$ to x_μ,u_0 in a potential A_μ differs from $k^{(0)}$ to first order in A_μ by a term

$$-i\int_0^{u_0}duk^{(0)}(x,u_0;y,u)\tfrac{1}{2}(p_\mu A_\mu(y)+A_\mu(y)p_\mu)k^{(0)}(y,u;x',0)d\tau_y$$

the p_μ here meaning $+i\partial/\partial y_\mu$. The kernel of importance on selecting the eigenvalue $\frac{1}{2}m^2$ is obtained by multiplying this by $\exp(-\frac{1}{2}im^2u_0)$ and integrating u_0 from 0 to ∞. The kernel $k^{(0)}(x,u_0;y,u)$ depends only on $u'=u_0-u$ and in the integrals on u and u_0; $\int_0^\infty du_0\int_0^{u_0}du\exp(-\frac{1}{2}im^2u_0)\cdots$, can be written, on interchanging the order of integration and changing variables to u and u', $\int_0^\infty du\int_0^\infty du'\exp(-\frac{1}{2}im^2(u+u'))\cdots$. Now the integral on u' converts $k^{(0)}(x,u_0;y,u)$ to $2iI_+(x,y)$ by (8A), while that on u converts $k^{(0)}(y,u;x',0)$ to $2iI_+(y,x')$, so the result becomes

$$\int 2iI_+(x,y)(p_\mu A_\mu+A_\mu p_\mu)I_+(y,x')d^4\tau_y$$

as expected. The same principle works to any order so that the rules for a single Klein-Gordon particle in external potentials given in II, Section 9, are deduced.

The transition to quantum electrodynamics is simple for in (5A) we already have a transition amplitude represented as a sum (over trajectories, and eventually u_0) of terms, in each of which the potential appears in exponential form. We may make use of the general relation (54). Hence, for example, one finds for the case of no photons in the initial and final states, in the presence of an external potential B_μ, the amplitude that a particle proceeds from $(x_\mu',0)$ to (x_μ,u_0) is the sum over all trajectories of the quantity

$$\exp-i\Big[\frac{1}{2}\int_0^{u_0}\Big(\frac{dx_\mu}{du}\Big)^2du+\int_0^{u_0}\frac{dx_\mu}{du}B_\mu(x(u))du$$
$$+\frac{e^2}{2}\int_0^{u_0}\int_0^{u_0}\frac{dx_\mu(u)}{du}\frac{dx_\nu(u')}{du'}\delta_+((x_\mu(u)-x_\mu(u'))^2)dud u'\Big]. \quad (10A)$$

This result must be multiplied by $\exp(-\frac{1}{2}im^2u_0)$ and integrated on u_0 from zero to infinity to express the action of a Klein-Gordon particle acting on itself through virtual photons. The integrals are interpreted as Riemann sums, and if perturbation expansions are made, the necessary care is taken with the terms quadratic in velocity. When there are several particles (other than the virtual pairs already included) one use a separate u for each, and writes the amplitude for each set of trajectories as the exponental of $-i$ times

$$\frac{1}{2}\sum_n\int_0^{u_0^{(n)}}\Big(\frac{dx_\mu^{(n)}}{du}\Big)^2du+\sum_n\int_0^{u_0^{(n)}}\frac{dx_\mu^{(n)}}{du}B_\mu(x_\mu^{(n)}(u))du$$
$$+\frac{e^2}{2}\sum_n\sum_m\int_0^{u_0^{(n)}}\int_0^{u_0^{(m)}}\frac{dx_\nu^{(n)}(u)}{du}\frac{dx_\nu^{(m)}(u')}{du'}$$
$$\times\delta_+((x_\mu^{(n)}(u)-x_\mu^{(m)}(u'))^2)dud u', \quad (11A)$$

where $x_\mu^{(n)}(u)$ are the coordinates of the trajectory of the nth particle.[22] The solution should depend on the $u_0^{(n)}$ as $\exp(-\frac{1}{2}im^2\sum_n u_0^{(n)})$.

Actually, knowledge of the motion of a single charge implies a great deal about the behavior of several charges. For a pair which eventually may turn out to be a virtual pair may appear in the short run as two "other particles." As a virtual pair, that is, as the reverse section of a very long and complicated single track we know its behavior by (10A). We can assume that such a section can be looked at equally well, for a limited duration at least, as being due to other unconnected particles. This then implies a definite law of interaction of particles if the self-action (10A) of a single particle is known. (This is similar to the relation of real and virtual photon processes discussed in detail in Appendix B.) It is possible that a detailed analysis of this could show that (10A) implied that (11A) was correct for many particles. There is even reason to believe that the law of Bose-Einstein statistics and the expression for contributions from closed loops could be deduced by following this argument. This has not yet been analyzed completely, however, so we must leave this formulation in an incomplete form. The expression for closed loops should come out to be $C_v=\exp+L$ where L, the contribution from a

[20] The factor $2i$ in front of I_+ is simply to make the definition of I_+ here agree with that in I and II. In II it operates with $\mathbf{p}\cdot\mathbf{A}+\mathbf{A}\cdot\mathbf{p}$ as a perturbation. But the perturbation coming from (3A) in a natural way by expansion of the exponential is $-\frac{1}{2}i(\mathbf{p}\cdot\mathbf{A}+\mathbf{A}\cdot\mathbf{p})$.

[21] Expression (8A) is closely related to Schwinger's parametric integral representation of these functions. For example, (8A) becomes formula (45) of F. Dyson, Phys. Rev. **75**, 486 (1949) for $\Delta_F\equiv\Delta^{(1)}-2i\bar{\Delta}\equiv 2iI_+$ if $(2\alpha)^{-1}$ is substituted for u_0.

[22] The form (10A) suggests another interesting possibility for avoiding the divergences of quantum electrodynamics in this case. The divergences arise from the δ_+ function when $u=u'$. We might restrict the integration in the double integral such that $|u-u'|>\delta$ where δ is some finite quantity, very small compared with m^{-2}. More generally, we could keep the region $u=u'$ from contributing by including in the integrand a factor $F(u-u')$ where $F(x)\to 1$ for x large compared to some δ, and $F(0)=0$ (e.g., $F(x)$ acts qualitatively like $1-\exp(-x^2\delta^{-2})$. (Another way might be to replace u by a discontinuous variable, that is, we do not use the limit in (4A) as $\epsilon\to 0$ but set $\epsilon=\delta$.) The idea is that two interactions would contribute very little in amplitude if they followed one another too rapidly in u. It is easily verified that this makes the otherwise divergent integrals finite. But whether the resulting formulas make good physical sense is hard to see. The action of a potential would now depend on the value of u so that Eq. (2A), or its equivalent, would not be separable in u so that $\frac{1}{2}m^2$ would no longer be a strict eigenvalue for all disturbances. High energy potentials could excite states corresponding to other eigenvalues, possibly thereby corresponding to other masses. This note is meant only as a speculation, for not enough work has been done in this direction to make sure that a reasonable physical theory can be developed along these lines. (What little work has been done was not promising.) Analogous modifications can also be made for Dirac electrons.

single loop, is

$$L=2\int_0^\infty l(u_0)\exp(-\tfrac{1}{2}im^2u_0)du_0/u_0$$

where $l(u_0)$ is the sum over all trajectories which close on themselves ($x_\mu(u_0)=x_\mu(0)$) of $\exp(iS)$ with S given in (5A), and a final integration $d\tau_{x(0)}$ on $x_\mu(0)$ is made. This is equivalent to putting

$$l(u_0)=\int(k^{(A)}(x,u_0;x,0)-k^{(0)}(x,u_0;x,0))d\tau_x.$$

The term $k^{(0)}$ is subtracted only to simplify convergence problems (as adding a constant independent of A_μ to L has no effect).

APPENDIX B. THE RELATION OF REAL AND VIRTUAL PROCESSES

If one has a general formula for all virtual processes he should be able to find the formulas and states involved in real processes. That is to say, we should be able to deduce the formulas of Section 9 directly from the formulation (24), (25) (or its generalized equivalent such as (46), (48)) without having to go all the way back to the more usual formulation. We discuss this problem here.

That this possibility exists can be seen from the consideration that what looks like a real process from one point of view may appear as a virtual process occurring over a more extended time.

For example, if we wish to study a given real process, such as the scattering of light, we can, if we wish, include in principle the source, scatterer, and eventual absorber of the scattered light in our analysis. We may imagine that no photon is present initially, and that the source then emits light (the energy coming say from kinetic energy in the source). The light is then scattered and eventually absorbed (becoming kinetic energy in the absorber). From this point of view the process is virtual; that is, we start with no photons and end with none. Thus we can analyze the process by means of our formula for virtual processes, and obtain the formulas for real processes by attempting to break the analysis into parts corresponding to emission, scattering, and absorption.[23]

To put the problem in a more general way, consider the amplitude for some transition from a state empty of photons far in the past (time t') to a similar one far in the future ($t=t''$). Suppose the time interval to be split into three regions a, b, c in some convenient manner, so that region b is an interval $t_2>t>t_1$ around the present time that we wish to study. Region a, $(t_1>t>t')$, precedes b, and c, $(t''>t>t_2)$, follows b. We want to see how it comes about that the phenomena during b can be analyzed by a study of transitions $g_{ji}(b)$ between some initial state i at time t_1 (which no longer need be photon-free), to some other final state j at time t_2. The states i and j are members of a large class which we will have to find out how to specify. (The single index i is used to represent a large number of quantum numbers, so that different values of i will correspond to having various numbers of various kinds of photons in the field, etc.) Our problem is to represent the over-all transition amplitude, $g(a,b,c)$, as a sum over various values of i,j of a product of three amplitudes,

$$g(a,b,c)=\Sigma_i\Sigma_j g_{0j}(c)g_{ji}(b)g_{i0}(a); \qquad (1B)$$

first the amplitude that during the interval a the vacuum state makes transition to some state i, then the amplitude that during b the transition to j is made, and finally in c the amplitude that the transition from j to some photon-free state 0 is completed.

The mathematical problem of splitting $g(a,b,c)$ is made definite by the further condition that $g_{ji}(b)$ for given i,j must not involve the coordinates of the particles for times corresponding to regions a or c, $g_{i0}(a)$ must involve those only in region a, and $g_{0j}(c)$ only in c.

To become acquainted with what is involved, suppose first that we do not have a problem involving virtual photons, but just the transition of a one-dimensional Schrödinger particle going in a long time interval from, say, the origin o to the origin o, and ask what states i we shall need for intermediary time intervals. We must solve the problem (1B) where $g(a,b,c)$ is the sum over all trajectories going from o at t' to o at t'' of $\exp iS$ where $S=\int L dt$. The integral may be split into three parts $S=S_a+S_b+S_c$ corresponding to the three ranges of time. Then $\exp(iS)=\exp(iS_c)\cdot\exp(iS_b)\cdot\exp(iS_a)$ and the separation (1B) is accomplished by taking for $g_{i0}(a)$ the sum over all trajectories lying in a from o to some end point x_{t_1} of $\exp(iS_a)$, for $g_{ji}(b)$ the sum over trajectories in b of $\exp(iS_b)$ between end points x_{t_1} and x_{t_2}, and for $g_{0j}(c)$ the sum of $\exp(iS_c)$ over the section of the trajectory lying in c and going from x_{t_2} to o. Then the sum on i and j can be taken to be the integrals on x_{t_1}, x_{t_2} respectively. Hence the various states i can be taken to correspond to particles being at various coordinates x. (Of course any other representation of the states in the sense of Dirac's transformation theory could be used equally well. Which, one, whether coordinate, momentum, or energy level representation, is of course just a matter of convenience and we cannot determine that simply from (1B).)

We can consider next the problem including virtual photons. That is, $g(a,b,c)$ now contains an additional factor $\exp(iR)$ where R involves a double integral $\int\int$ over all time. Those parts of the index i which correspond to the particle states can be taken in the same way as though R were absent. We study now the extra complexities in the states produced by splitting the R. Let us first (solely for simplicity of the argument) take the case that there are only two regions a, c separated by time t_0 and try to expand

$$g(a,c)=\Sigma_i g_{0i}(c)g_{i0}(a).$$

The factor $\exp(iR)$ involves R as a double integral which can be split into three parts $\int_a\int_a+\int_c\int_c+\int_a\int_c$ for the first of which both t, s are in a, for the second both are in c, for the third one is in a the other in c. Writing $\exp(iR)$ as $\exp(iR_{cc})\cdot\exp(iR_{aa})\cdot\exp(iR_{ac})$ shows that the factors R_{cc} and R_{aa} produce no new problems for they can be taken bodily into $g_{0i}(c)$ and $g_{i0}(a)$ respectively. However, we must disentangle the variables which are mixed up in $\exp(iR_{ac})$.

The expression for R_{ac} is just twice (24) but with the integral on s extending over the range a and that for t extending over c. Thus $\exp(iR_{ac})$ contains the variables for times in a and in c in a quite complicated mixture. Our problem is to write $\exp(iR_{ac})$ as a sum over possibly a vast class of states i of the product of two parts, like $h_i'(c)h_i(a)$, each of which involves the coordinates in one interval alone.

This separation may be made in many different ways, corresponding to various possible representations of the state of the electromagnetic field. We choose a particular one. First we can expand the exponential, $\exp(iR_{ac})$, in a power series, as $\Sigma_n i^n(n!)^{-1}(R_{ac})^n$. The states i can therefore be subdivided into subclasses corresponding to an integer n which we can interpret as the number of quanta in the field at time t_0. The amplitude for the case $n=0$ clearly just involves $\exp(iR_{aa})$ and $\exp(iR_{cc})$ in the way that it should if we interpret these as the amplitudes for regions a and c, respectively, of making a transition between a state of zero photons and another state of zero photons.

Next consider the case $n=1$. This implies an additional factor in the transitional element; the factor R_{ac}. The variables are still mixed up. But an easy way to perform the separation suggests itself. Namely, expand the $\delta_+((t-s)^2-(\mathbf{x}_n(t)-\mathbf{x}_m(s))^2)$ in R_{ac} as a Fourier integral as

$$i\int\exp(-ik|t-s|)\exp(-i\mathbf{K}\cdot(\mathbf{x}_n(t)-\mathbf{x}_m(s))d^3\mathbf{K}/4\pi^2k.$$

[23] The formulas for real processes deduced in this way are strictly limited to the case in which the light comes from sources which are originally dark, and that eventually all light emitted is absorbed again. We can only extend it to the case for which these restrictions do not hold by hypothesis, namely, that the details of the scattering process are independent of these characteristics of the light source and of the eventual disposition of the scattered light. The argument of the text gives a method for discovering formulas for real processes when no more than the formula for virtual processes is at hand. But with this method belief in the general validity of the resulting formulas must rest on the physical reasonableness of the above-mentioned hypothesis.

For the exponential can be written immediately as a product of $\exp+i(\mathbf{K}\cdot\mathbf{x}_m(s))$, a function only of coordinates for times s in a (suppose $s<t$), and $\exp-i\mathbf{K}\cdot\mathbf{x}_n(t)$ (a function only of coordinates during interval c). The integral on $d^3\mathbf{K}$ can be symbolized as a sum over states i characterized by the value of $\mathbf{K}$. Thus the state with $n=1$ must be further characterized by specifying a vector $\mathbf{K}$, interpreted as the momentum of the photon. Finally the factor $(1-\mathbf{x}'_n(t)\cdot\mathbf{x}'_m(s))$ in R_{ac} is simply the sum of four parts each of which is already split (namely 1, and each of the three components in the vector scalar product). Hence each photon of momentum $\mathbf{K}$ must still be characterized by specifying it as one of four varieties; that is, there are four polarizations.[24] Thus in trying to represent the effect of the past a on the future c we are lead to invent photons of four polarizations and characterized by a propagation vector $\mathbf{K}$.

The term for a given polarization and value of $\mathbf{K}$ (for $n=1$) is clearly just $-\beta_c\beta_a^*$ where the β_a is defined in (59) but with the time integral extending just over region a, while β_c is the same expression with the integration over region c. Hence the amplitude for transition during interval a from a state with no quanta to a state with one in a given state of polarization and momentum is calculated by inclusion of an extra factor $i\beta_a^*$ in the transition element. Absorption in region c corresponds to a factor $i\beta_c$.

We next turn to the case $n=2$. This requires analysis of R_{ac}^2. The δ_+ can be expanded again as a Fourier integral, but for each of the two δ_+ in $\frac{1}{2}R_{ac}^2$ we have a value of $\mathbf{K}$ which may be different. Thus we say, we have two photons, one of momentum $\mathbf{K}$ and one momentum $\mathbf{K}'$ and we sum over all values of $\mathbf{K}$ and $\mathbf{K}'$. (Similarly each photon is characterized by its own independent polarization index.) The factor $\frac{1}{2}$ can be taken into account neatly by asserting that we count each possible pair of photons as constituting just one state at time t_0. Then the $\frac{1}{2}$ arises for the sum over all $\mathbf{K}$, $\mathbf{K}'$ (and polarizations) counts each pair twice. On the other hand, for the terms representing two identical photons ($\mathbf{K}=\mathbf{K}'$) of like polarization, the $\frac{1}{2}$ cannot be so interpreted. Instead we invent the rule that a state of two like photons has statistical weight $\frac{1}{2}$ as great as that calculated as though the photons were different. This, generalized to n identical photons, is the rule of Bose statistics.

The higher values of n offer no problem. The $1/n!$ is interpreted combinatorially for different photons, and as a statistical factor when some are identical. For example, for all n identical one obtains a factor $(n!)^{-1}(-\beta_c\beta_a^*)^n$ so that $(n!)^{-\frac{1}{2}}(i\beta_a^*)^n$ can be interpreted as the amplitude for emission (from no initial photons) of n identical photons, in complete agreement with (61) for G_{m0}.

To obtain the amplitude for transitions in which neither the initial nor the final state is empty of photons we must consider the more general case of the division into three time regions (1B). This time we see that the factor which involves the coordinates in an entangled manner is $\exp i(R_{ab}+R_{bc}+R_{ac})$. It is to be expanded in the form $\Sigma_i\Sigma_j h_i''(c)h_{ij}'(b)h_j(a)$. Again the expansion in power series and development in Fourier series with a polarization sum will solve the problem. Thus the exponential is $\Sigma_r\Sigma_{l_1}\Sigma_{l_2}(iR_{ac})^r(iR_{ab})^{l_1}(iR_{bc})^{l_2}(l_1!)^{-1}(l_2!)^{-1}(r!)^{-1}$. Now the R are written as Fourier series, one of the terms containing l_1+l_2+r variables $\mathbf{K}$. Since l_1+r involve a, l_2+r involve c and l_1+l_2 involve b, this term will give the amplitude that l_1+r photons are emitted during the interval a, of those l_1 are absorbed during b but the remaining r, along with l_2 new ones emitted during b go on to be absorbed during the interval c. We have therefore $n=l_1+r$ photons in the state at time t_1 when b begins, and $m=l_2+r$ at t_2 when b is over. They each are characterized by momentum vectors and polarizations. When these are different the factors $(l_1!)^{-1}(l_2!)^{-1}(r!)^{-1}$ are absorbed combinatorially. When some are equal we must invoke the rule of the statistical weights. For example, suppose all l_1+l_2+r photons are identical. Then $R_{ab}=i\beta_b\beta_a^*$, $R_{bc}=i\beta_c\beta_b^*$, $R_{ac}=i\beta_c\beta_a^*$ so that our sum is

$$\Sigma_{l_1}\Sigma_{l_2}\Sigma_r(l_1!l_2!r!)^{-1}(i\beta_c)^{l_2+r}(i\beta_b)^{l_1}(i\beta_b^*)^{l_2}(i\beta_a^*)^{l_1+r}.$$

Putting $m=l_2+r$, $n=l_1+r$, this is the sum on n and m of

$$(i\beta_c)^m(m!)^{-\frac{1}{2}}[\Sigma_r(m!n!)^{\frac{1}{2}}((m-r)!(n-r)!r!)^{-1}\times(i\beta_b^*)^{m-r}(i\beta_b)^{n-r}](n!)^{-\frac{1}{2}}(i\beta_a^*)^n.$$

The last factor we have seen is the amplitude for emission of n photons during interval a, while the first factor is the amplitude for absorption of m during c. The sum is therefore the factor for transition from n to m identical photons, in accordance with (57). We see the significance of the simple generating function (56).

We have therefore found rules for real photons in terms of those for virtual. The real photons are a way of representing and keeping track of those aspects of the past behavior which may influence the future.

If one starts from a theory involving an arbitrary modification of the direct interaction δ_+ (or in more general situations) it is possible in this way to discover what kinds of states and physical entities will be involved if one tries to represent in the present all the information needed to predict the future. With the Hamiltonian method, which begins by assuming such a representation, it is difficult to suggest modifications of a general kind, for one cannot formulate the problem without having a complete representation of the characteristics of the intermediate states, the particles involved in interaction, etc. It is quite possible (in the author's opinion, it is very likely) that we may discover that in nature the relation of past and future is so intimate for short durations that no simple representation of a present may exist. In such a case a theory could not find expression in Hamiltonian form.

An exactly similar analysis can be made just as easily starting with the general forms (46), (48). Also a coordinate representation of the photons could have been used instead of the familiar momentum one. One can deduce the rules (60), (61). Nothing essentially different is involved physically, however, so we shall not pursue the subject further here. Since they imply[23] all the rules for real photons, Eqs. (46), (47), (48) constitute a compact statement of all the laws of quantum electrodynamics. But they give divergent results. Can the result *after* charge and mass renormalization also be expressed to all orders in $e^2/\hbar c$ in a simple way?

APPENDIX C. DIFFERENTIAL EQUATION FOR ELECTRON PROPAGATION

An attempt has been made to find a differential wave equation for the propagation of an electron interacting with itself, analogous to the Dirac equation, but containing terms representing the self-action. Neglecting all effects of closed loops, one such equation has been found, but not much has been done with it. It is reported here for whatever value it may have.

An electron acting upon itself is, from one point of view, a complex system of a particle and a field of an indefinite number of photons. To find a differential law of propagation of such a system we must ask first what quantities known at one instant will permit the calculation of these same quantities an instant later. Clearly, a knowledge of the position of the particle is not enough. We should need to specify: (1) the amplitude that the electron is at x and there are no photons in the field, (2) the amplitude the electron is at x and there is one photon of such and such a kind in the field, (3) the amplitude there are two photons, etc. That is, a series of functions of ever increasing numbers of variables. Following this view, we shall be led to the wave equation of the theory of second quantization.

We may also take a different view. Suppose we know a quantity $\Phi_e 2[B, x]$, a spinor function of x_μ, and functional of $B_\mu(1)$, defined as the amplitude that an electron arrives at x_μ with no photon in the field when it moves in an arbitrary external unquantized potential $B_\mu(1)$. We allow the electron also to interact with itself,

[24] Usually only two polarizations transverse to the propagation vector $\mathbf{K}$ are used. This can be accomplished by a further rearrangement of terms corresponding to the reverse of the steps leading from (17) to (19). We omit the details here as it is well-known that either formulation gives the same results. See II, Section 8.

but $\Phi_\epsilon 2$ is the amplitude at a given instant that there happens to be no photons present. As we have seen, a complete knowledge of this functional will also tell us the amplitude that the electron arrives at x and there is just one photon, of form $A_\mu^{PH}(1)$ present. It is, from (60), $\int(\delta\Phi_\epsilon 2[B, x]/\delta B_\mu(1))A_\mu^{PH}(1)d\tau_1$.

Higher numbers of photons correspond to higher functional derivatives of $\Phi_\epsilon 2$. Therefore, $\Phi_\epsilon 2[B, x]$ contains all the information requisite for describing the state of the electron-photon system, and we may expect to find a differential equation for it. Actually it satisfies ($\nabla = \gamma_\mu \partial/\partial x_\mu$, $B = \gamma_\mu B_\mu$),

$$(i\nabla - m)\Phi_\epsilon 2[B, x] = B(x)\Phi_\epsilon 2[B, x] + ie^2\gamma_\mu \int \delta_+(s_{x1}^2)(\delta\Phi_\epsilon 2[B, x]/\delta B_\mu(1))d\tau_1 \quad (1C)$$

as may be seen from a physical argument.[25] The operator $(i\nabla - m)$ operating on the x coordinate of $\Phi_\epsilon 2$ should equal, from Dirac's equation, the changes in $\Phi_\epsilon 2$ as we go from one position x to a neighboring position due to the action of vector potentials. The term $B(x)\Phi_\epsilon 2$ is the effect of the external potential. But $\Phi_\epsilon 2$ may also change for at the first position x we may have had a photon present (amplitude that it was emitted at another point 1 is $\delta\Phi_\epsilon 2/\delta B_\mu(1)$) which was absorbed at x (amplitude photon released at 1 gets to x is $\delta_+(s_{x1}^2)$ where s_{x1}^2 is the squared invariant distance from 1 to x) acting as a vector potential there (factor γ_μ). Effects of vacuum polarization are left out.

Expansion of the solution of (1C) in a power series in B and e^2 starting from a free particle solution for a single electron, produces a series of terms which agree with the rules of II for action of potentials and virtual photons to various orders. It is another matter to use such an equation for the practical solution of a problem to all orders in e^2. It might be possible to represent the self-energy problem as the variational problem for m, stemming from (1C). The δ_+ will first have to be modified to obtain a *convergent result*.

We are not in need of the general solution of (1C). (In fact, we have it in (46), (48) in terms of the solution $T_0[B] = \Phi_0[B, x]$ of the ordinary Dirac equation $(i\nabla - m)\Phi_0[B, x] = B\Phi_0[B, x]$. The general solution is too complicated, for complete knowledge of the motion of a self-acting electron in an arbitrary potential is essentially all of electrodynamics (because of the kind of relation of real and virtual processes discussed for photons in Appendix B, extended also to real and virtual pairs). Furthermore, it is easy to see that other quantities also satisfy (1C). Consider a system of many electrons, and single out some one for consideration, supposing all the others go from some definite initial state i to some definite final state f. Let $\Phi_\epsilon 2[B, x]$ be the amplitude that the special electron arrives at x, there are no photons present, and the other electrons go from i to f when there is an external potential B_μ present (which B_μ also acts on the other electrons). Then $\Phi_\epsilon 2$ also satisfies (1C). Likewise the amplitude with closed loops (all other electrons go vacuum to vacuum) also satisfies (1C) including all vacuum polarization effects. The various problems correspond to different assumptions as to the dependence of $\Phi_\epsilon 2[B, x]$ on B_μ in the limit of zero e^2. The Eq. (1C) without further boundary conditions is probably too general to be useful.

[25] Its general validity can also be demonstrated mathematically from (45). The amplitude for arriving at x with no photons in the field with virtual photon coupling e^2 is a transition amplitude. It must, therefore, satisfy (45) with $T_\epsilon 2[B] = \Phi_\epsilon 2[B, x]$ for any x. Hence show that the quantity

$$C_\epsilon 2[B, x] = (i\nabla - m - B(x))\Phi_\epsilon 2[B, x] - ie^2\gamma_\mu \int \delta_+(s_{x1}^2)(\delta\Phi_\epsilon 2[B, x]/\delta B_\mu(1))d\tau_1$$

also satisfies Eq. (45) by substituting $C_\epsilon 2[B, x]$ for $T_\epsilon 2[B]$ in (45) and using the fact that $\Phi_\epsilon 2[B, x]$ satisfies (45). Hence if $C_0[B, x] = 0$ then $C_\epsilon 2[B, x] = 0$ for all e^2. But $C_\epsilon 2[B, x] = 0$ means that $\Phi_\epsilon 2[B, x]$ satisfies (1C). Therefore, that solution $\Phi_\epsilon 2[B, x]$ of (45) which also satisfies $(i\nabla - m - B(x))\Phi_0[B, x] = 0$ (the propagation of a free electron without virtual photons) is a solution of (1C) as we wished to show. Equation (1C) may be more convenient than (45) for some purposes for it does not involve differentiation with respect to the coupling constant, and is more analogous to a wave equation.

PHYSICAL REVIEW VOLUME 84, NUMBER 1 OCTOBER 1, 1951

An Operator Calculus Having Applications in Quantum Electrodynamics

RICHARD P. FEYNMAN*
California Institute of Technology, Pasadena, California
(Received May 23, 1951)

An alteration in the notation used to indicate the order of operation of noncommuting quantities is suggested. Instead of the order being defined by the position on the paper, an ordering subscript is introduced so that $A_sB_{s'}$ means AB or BA depending on whether s exceeds s' or *vice versa*. Then A_s can be handled as though it were an ordinary numerical function of s. An increase in ease of manipulating some operator expressions results. Connection to the theory of functionals is discussed in an appendix. Illustrative applications to quantum mechanics are made. In quantum electrodynamics it permits a simple formal understanding of the interrelation of the various present day theoretical formulations.

The operator expression of the Dirac equation is related to the author's previous description of positrons. An attempt is made to interpret the operator ordering parameter in this case as a fifth coordinate variable in an extended Dirac equation. Fock's parametrization, discussed in an appendix, seems to be easier to interpret.

In the last section a summary of the numerical constants appearing in formulas for transition probabilities is given.

IN this paper we suggest an alteration in the mathematical notation for handling operators. This new notation permits a considerable increase in the ease of manipulation of complicated expressions involving operators. No results which are new are obtained in this way, but it does permit one to relate various formulas of operator algebra in quantum mechanics in a simpler manner than is often available. In particular, it is applied to quantum electrodynamics to permit an easier way of seeing the relationships among the conventional formulations, that of Schwinger and Tomanaga,[1] and that of the author.[2] These relationships have already been discussed by many people, particularly Dyson.[3] The connection was shown by means of a re-ordering of operators in each term of a perturbation power series. Here, the same end is achieved in much the same way without having to resort to such an expansion.

It is felt, in the face of daily experimental surprises for meson theory, that it might be worth while to spend one's time expressing electrodynamics in every physical and mathematical way possible. There may be some hope that a thorough understanding of electrodynamics might give a clue as to the possible structure of the more complete theory to which it is an approximation. This is one reason that this paper is published, even though it is little more than a mathematical re-expression of old material. A second reason is the desire to describe a mathematical method which may be useful in other fields.

The mathematics is not completely satisfactory. No attempt has been made to maintain mathematical rigor. The excuse is not that it is expected that rigorous demonstrations can be easily supplied. Quite the contrary, it is believed that to put the present methods on a rigorous basis may be quite a difficult task, beyond the abilities of the author.

The mathematical ideas are described and are illustrated with simple applications to quantum mechanics, in the first four sections. Some possible mathematical relations between the operator calculus described here and the theory of functionals is described in Appendix A, with further specific mathematical applications in Appendixes B and C. Section 5, and more particularly Secs. 6 to 9, apply specifically to quantum electrodynamics and may be omitted without loss by those whose interest is limited to mathematical questions. The use of a fifth variable to parametrize the Dirac equation is discussed in Secs. 8 and 9. An alternative procedure due to V. Fock[4] appears in Appendix D. Section 10 gives a summary of the rules for computing matrix elements.

1. DESCRIPTION OF THE NOTATION

The order of operation of operators is conventionally represented by the position in which the operators are written on the paper. Thus, the product AB of two operators A and B is to be distinguished from the product in reverse order BA. The algebra of operators is noncommutative, so that all of the ordinary algebra, calculus, and analysis with ordinary numbers becomes of small utility for operators. Thus, for a single operator, α, ordinary functions of this operator, such as $A=\exp\alpha$, can be defined, for example, by power series. These functions obey the rules of ordinary analysis even though α is an operator. But if another operator β is introduced with which α does not commute, the question of functions of the two variables α, β is beset with commutation difficulties and the simplest theorems of analysis are lost. For example, if $B=\exp\beta$, it is not true

* Absent on leave at the University of Brazil, Rio de Janeiro, Brazil.

[1] See J. Schwinger, Phys. Rev. **76**, 790 (1949), and S. Tomonaga, Phys. Rev. **74**, 224 (1948), where additional references to previous work may be found.

[2] The author's previous papers will hereafter be designated as follows: R. P. Feynman, Revs. Modern Phys. **20**, 367 (1948)—C; Phys. Rev. **76**, 749 (1949)—I; Phys. Rev. **76**, 769 (1949)—II; and Phys. Rev. **80**, 440 (1950)—III.

[3] F. Dyson, Phys. Rev. **75**, 486, 1736 (1949).

[4] V. Fock, Physik. Z. Sowjetunion **12**, 404 (1937).

that BA, that is, $\exp\beta\exp\alpha$, is equal to $\exp(\beta+\alpha)$. Thus, the law of addition of exponents fails. Consequently, the principles of elementary calculus are no longer operative in a simple way. For example, expand $\exp(\alpha+\beta)$ to first order in β, assuming β small. The zero-order term is $\exp\alpha$, but the first-order term is neither $\beta\exp\alpha$ nor $(\exp\alpha)\beta$ nor the average of the two. From the theory of time-dependent perturbations in quantum mechanics we learn that it is

$$\exp(\alpha+\beta) = \exp\alpha + \int_0^1 \exp[(1-s)\alpha]\beta\exp(s\alpha)ds + \cdots \quad (1)$$

The appearance of the integral in this analytic result appears surprising and its derivation does not indicate clearly how to differentiate or expand other functions of $\alpha+\beta$. Further, the simple integral on s is not easy to perform, although the results can be given in several ways as power series. That the integral cannot be done in a general fashion is clearly due to a weakness of notation, for in a representation in which α is diagonal with eigenvalues α_n we can of course verify directly the usual result,

$$\langle\exp(\alpha+\beta)\rangle_{mn} = (\exp\alpha_n)\delta_{mn} + (\exp\alpha_m - \exp\alpha_n)\beta_{mn}(\alpha_m-\alpha_n)^{-1} + \cdots \quad (2)$$

of the perturbation theory of stationary states.

We shall change the usual notation of the theory of operators and indicate the order in which operators are to operate by a different device. We attach an index to the operator with the rule that the operator with higher index operates later. Thus, BA may be written B_1A_0 or A_0B_1. The order no longer depends on the position on the paper, so that all of the ordinary processes of analysis may be applied as though A_0 and B_1 were commuting numbers. It is only at the end of a calculation, when the quantities are to be interpreted as operators, that the indices 0 and 1 are of importance if one wishes to reconvert an expression to the usual notation. Thus, if $A=\exp\alpha$ and $B=\exp\beta$, we can now safely write $BA=\exp(\alpha_0+\beta_1)$, as there is only one way to interpret the latter expression. Other analytic processes then become valid. For example,

$$\exp(\alpha_0+\beta_1) = 1+(\alpha_0+\beta_1)+\frac{1}{2!}(\alpha_0+\beta_1)^2+\cdots = 1+\alpha+\beta+\tfrac{1}{2}(\alpha^2+2\beta\alpha+\beta^2)+\cdots$$

in the conventional notation. For on squaring

$$(\alpha_0+\beta_1)^2 = \alpha_0^2+2\alpha_0\beta_1+\beta_1^2,$$

we must interpret the quantity $\alpha_0\beta_1$ as $\beta\alpha$ in accordance with our convention. The quantity β_1^2 alone (that is, not multiplied by any other expression with an index, such as α_0) is simply β^2, since the index is no longer necessary to define the order of operations, there being only one operator in the term.

The notation is to be extended so that the index need not be integral, for example, $A_{-\frac{1}{2}}B_{3.2}=BA$, since $3.1>-\frac{1}{2}$, and in general $A_sB_{s'}=BA$ if $s'>s$ and AB if $s>s'$ and is undefined if $s=s'$.

How can we work with an expression such as $\exp(\alpha+\beta)$ so as to free the α and β of their noncommutative aspects and thus utilize the theory of functions for rearranging the expressions? Take a quantity N very large and write

$$\exp(\alpha+\beta)\approx\left[\exp\frac{1}{N}(\alpha+\beta)\right]^N\approx\left[1+\frac{1}{N}(\alpha+\beta)\right]^N = \left[1+\frac{1}{N}(\alpha+\beta)\right]\left[1+\frac{1}{N}(\alpha+\beta)\right]\cdots \times\left[1+\frac{1}{N}(\alpha+\beta)\right] \text{ for } N \text{ factors.}$$

In each factor we replace $\alpha+\beta$ by $\alpha_i+\beta_i$, where i is an index running to N, and write

$$\exp(\alpha+\beta) = \lim_{N\to\infty}\prod_{i=1}^{N}\left[1+\frac{1}{N}(\alpha_i+\beta_i)\right] = \lim_{N\to\infty}\exp\left[\frac{1}{N}\sum_{i=1}^{N}(\alpha_i+\beta_i)\right],$$

where the last expression is written in accordance with the new convention that the index i controls the order of operation. (The ambiguity arising from α_i and β_i with the same index can only cause trouble in a product $\alpha_i\beta_i$, and such products are of vanishing importance as $N\to\infty$.) More simply, calling $s=i/N$, we can take the limit and write

$$\exp(\alpha+\beta) = \exp\left[\int_0^1(\alpha_s+\beta_s)ds\right]. \quad (3)$$

That this is valid is, of course, evident, since we could call $\alpha_s+\beta_s=\gamma_s$ with γ a definite operator operating at order s, so that

$$\exp\left(\int_0^1\gamma_s ds\right) = \exp\left(\int_0^1\gamma ds\right),$$

for in this expression the order index is unnecessary, only one operator γ being involved. The integral is just γ,

$$\int_0^1\gamma ds = \gamma\int_0^1 ds = \gamma,$$

since γ does not now depend on s. Therefore, Eq. (3) is trivial as it stands; but what is not trivial is the fact

that the right-hand side of Eq. (3) may be manipulated just as though α_s and β_s were numerical functions of s, with the assurance that now the order of operations will always be automatically specified by the index. For example, from Eq. (3) we have the legitimate relation

$$\exp(\alpha+\beta)=\exp\left(\int_0^1 \alpha_s ds\right)\exp\left(\int_0^1 \beta_s ds\right). \quad (4)$$

As an example, showing that such manipulations do not destroy the validity of equations, consider the term of first order in both α and β on both sides of Eq. (4). Expanding the left side as $1+(\alpha+\beta)+\frac{1}{2}(\alpha+\beta)^2+\cdots$, we see that the term in question is $\frac{1}{2}(\alpha\beta+\beta\alpha)$, while expansion of the right side gives for the corresponding term $(\int_0^1 \alpha_s ds)(\int_0^1 \beta_s ds)$. This can be simplified by being written as

$$\int_0^1\int_0^1 \alpha_s ds\beta_{s'}ds'=\int_0^1\int_0^s \alpha_s\beta_{s'}dsds'+\int_0^1\int_s^1 \alpha_s\beta_{s'}dsds'.$$

In the first integral we have $s>s'$, so that $\alpha_s\beta_{s'}$ is equal to $\alpha\beta$, while in the second $s<s'$, so it is $\beta\alpha$. Hence, there results $\alpha\beta\int_0^1\int_0^s dsds'+\beta\alpha\int_0^1\int_s^1 dsds'$; thus on performing the integrations we find finally

$$\left(\int_0^1 \alpha_s ds\right)\left(\int_0^1 \beta_s ds\right)=\tfrac{1}{2}(\alpha\beta+\beta\alpha).$$

This process of rearranging the form of expressions involving operators ordered by indices so that they may be written in conventional form we shall call disentangling the operators. The process is not always easy to perform and, in fact, is the central problem of this operator calculus. As a second example of disentangling, consider the problem of expanding Eq. (4) to the first order in β. It is evidently

$$\exp(\alpha+\beta)=\exp\left(\int_0^1 \alpha_s ds\right)+\exp\left(\int_0^1 \alpha_{s'}ds'\right)\cdot\int_0^1 \beta_s ds+\cdots. \quad (5)$$

The first term is simply $\exp\alpha$, for α_s is independent of s, as there is no other operator with which α_s does not commute in this term. The next is the integral over s of $\exp(\int_0^1 \alpha_{s'}ds')\beta_s$. In the integral on s' we can split the range, according to whether $s'<s$ or $s'>s$: $\exp(\int_s^1 \alpha_{s'}ds')\exp(\int_0^s \alpha_{s'}ds')\beta_s$. The $\alpha_{s'}$ in the first factor acts after the β_s and is otherwise independent of s', while the $\alpha_{s'}$ in the second factor is to act before the β_s. Hence, if we write these factors respectively after and before the β_s and imply the usual convention, the $\alpha_{s'}$ will be independent of s' in the range 0 to s and we may perform the integral. Hence, the result is the integral on s of $\exp[(1-s)\alpha]\beta\exp(s\alpha)$ in agreement with Eq. (1).

Incidentally, by applying new subscripts in another way the term may be also written as $\int_0^1 \exp[(1-s)\alpha_2]\beta_1\times\exp(s\alpha_0)ds$, in which case the integral may be immediately performed to give

$$\begin{aligned}[(d/d\epsilon)\exp(\alpha+\epsilon\beta)]_{\epsilon=0}&=\exp\left(\int_0^1 \alpha_{s'}ds'\right)\int_0^1 \beta_s ds\\&=\int_0^1 \exp[(1-s)\alpha]\beta\exp(s\alpha)ds\\&=(\exp\alpha_2-\exp\alpha_0)(\alpha_2-\alpha_0)^{-1}\beta_1. \quad (6)\end{aligned}$$

All the four expressions are equivalent as has been shown, but only the first and third are in a form in which the operators are disentangled so that the conventional expressions may be used. In the representation in which α is diagonal, it should be evident that the matrix element of the last expression is that given in Eq. (2)

Any operator function of $\alpha+\beta$ can, by replacing $\alpha+\beta$ by $\int_0^1 \alpha_s ds+\int_0^1 \beta_s ds$, be manipulated in a manifold of ways, many of which lead to useful formulas. In a like manner, more complicated operator expressions can be rewritten using ordering indices. They may then be manipulated using all of the results of ordinary analysis.

A word about notation: Inasmuch as in mathematics and physics there are already many uses of the subscript notation, very often we shall write $\alpha(s)$ for α_s. In a sense, $\alpha(s)$ is a function of s, namely, in the sense that although the operator α may be definite, its order of operation is not—so that the operator plus a prescription of where it is to operate, $\alpha(s)$, is a function of s. Furthermore, there will be many cases in which the operator actually depends explicitly on the parameter of order. In this case we should have strictly to write $\alpha_s(s)$ but will omit the subscript when no ambiguity will result from the change.

We may remark in a general sense about the mathematical character of our expressions. Given an expression such as $\int_0^1 \beta(s)ds$, we are not concerned with evaluating the integral, for the quantity when separated from other factors with which it might be multiplied is incompletely defined. Thus, although $\int_0^1 \beta_s ds$ standing alone is equivalent simply to β, this is far from true when $\int_0^1 \beta_s ds$ is multiplied by other expressions such as $\exp\int_0^1 \alpha_s ds$. Thus, we must consider the complete expression as a complete functional of the argument functions $\alpha(s)$, $\beta(s)$, etc. With each such functional we are endeavoring to associate an operator. The operator depends on the functional in a complex way (the operator is a functional of a functional) so that, for example, the operator corresponding to the product of two functionals is not (in general) the simple product of the operators corresponding to the separate factors. (The corresponding statement equating the sum of two

functionals and the sum of the corresponding operators is true, however.) Hence, we can consider the most complex expressions involving a number of operators M, N, as described by functionals $F[M(s), N(s)\cdots]$ of the argument functions $M(s)$, $N(s)\cdots(\equiv M_s, N_s\cdots)$. For each functional we are to find the corresponding operator in some simple form;[5] that is, we wish to disentangle the functional. One fact we know is that any analytic rearrangement may be performed which leaves the value of the functional unchanged for arbitrary $M(s)$, $N(s)\cdots$ considered as ordinary numerical functions. Besides, there are a few special operations which we may perform on $F[M(s), N(s)\cdots]$, to disentangle the expressions, which are valid only because the functional does represent an operator according to our rules. These special operations (such as extracting an exponential factor discussed in Sec. 3) are, of course, proper to the new calculus; and our powers of analysis in this field will increase as we develop more of them.

2. APPLICATIONS IN QUANTUM MECHANICS

The wave equation $i\partial\psi/\partial t=H\psi$ determines the wave function $\psi(t_2)$ at time t_2 in terms of that at time t_1, $\psi(t_1)$. In fact, they are related by a unitary transformation $\psi(t_2)=\Omega(t_2, t_1)\psi(t_1)$. The unitary operator $\Omega(t_2, t_1)$ can be expressed as $\Omega(t_2, t_1)=\exp(-i(t_2-t_1)H)$ in the case that H is independent of the time. In spite of the simple appearance of the analytic form of Ω in terms of H, little has been done except formally with this expression for the reasons outlined in the previous section. We may readily re-express it as

$$\Omega(t_2, t_1)=\exp\left(-i\int_{t_1}^{t_2} H_t dt\right)$$

and may then find the expression easy to utilize. Further, if H is an explicit function of the time $H(t)$, we can consider the Ω to be developed as a large number of small unitary transformations in succession, so that we have directly

$$\Omega(t_2, t_1)=\exp\left[-i\int_{t_1}^{t_2} H_t(t)dt\right]. \tag{7}$$

Hereafter in this section we shall make the convention that time is the ordering parameter and simply write $H(t)$ for $H_t(t)$.

We can use this expression to derive many results in quantum theory. Thus, if $H(t)$ can be written as the sum of two parts $H^{(0)}(t)$ and $U(t)$, we have

$$\Omega^{(U)}(t_2, t_1)=\exp\left(-i\int_{t_1}^{t_2} H^{(0)}(t)dt\right)\times\exp\left(-i\int_{t_1}^{t_2} U(t)dt\right). \tag{8}$$

[5] This point of view is discussed in further detail in Appendix A.

If $H^{(0)}$ is simple and U is small, an expansion in powers of U is simple. We call $\Omega^{(U)}$ the operator corresponding to the hamiltonian $H^{(0)}+U$ and $\Omega^{(0)}$ that corresponding to $H^{(0)}$. The first-order difference of $\Omega^{(U)}$and $\Omega^{(0)}$ is

$$-i\int_{t_1}^{t_2} U(t)dt\exp\left[-i\int_{t_1}^{t_2} H(t')dt'\right],$$

which may be disentangled as

$$-i\int_{t_1}^{t_2}\exp\left[-i\int_{t}^{t_2} H(t')dt'\right]U(t)\exp\left[-i\int_{t_1}^{t} H(t')dt'\right]dt = -i\int_{t_1}^{t_2}\Omega^{(0)}(t_2, t)U(t)\Omega^{(0)}(t, t_1)dt, \tag{9}$$

as explained in connection with Eq. (6). This is a standard result of time-dependent perturbation theory.

As a second example consider the perturbation term of first order in U and in V arising from the hamiltonian $H^{(0)}(t)+U(t)+V(t)$. It is

$$-\int_{t_1}^{t_2} U(t')dt'\int_{t_1}^{t_2} V(t'')dt''\exp\left[-i\int_{t_1}^{t_2} H(t)dt\right].$$

In order to disentangle this, we can break the t'' integral into two regions, $t''<t'$ and $t''>t'$. The term arising from the first region has V operating before U, while the reverse is true for the other region. (The integral on t for each region is then divided into three parts determined by the relation of t to t', t''.) Thus, the term becomes, when disentangled, the sum of two terms:

$$\begin{aligned}&-\int_{t_1}^{t_2}\int_{t_1}^{t'}\Omega^{(0)}(t_2, t')U(t')\Omega^{(0)}(t', t'')V(t'')\times\Omega^{(0)}(t'', t_1)dt''dt',\\ &-\int_{t_1}^{t_2}\int_{t'}^{t_2}\Omega^{(0)}(t_2, t'')V(t'')\Omega^{(0)}(t'', t')U(t')\times\Omega^{(0)}(t', t_1)dt''dt'.\end{aligned} \tag{10}$$

This is the way that the various terms corresponding to the different diagrams arise in quantum electrodynamics when an attempt is made to calculate explicitly a single operator expression arising in perturbation theory.

The results here are very similar to those derived from the lagrangian form of quantum mechanics as in III. Here we have the advantage of being able to use the more familiar operator concepts and to work in greater generality from the start. For it is not necessary that H be restricted to coordinate and momentum operators only. Equations (7) and (8) are correct for any H; for example, one containing creation and annihilation operators of second quantization, or Dirac matrices, etc.

The connection of these formulas to those given in I is simple. $K(x_2, t_2; x_1, t_1)$ is just a coordinate integral

kernel representation of the operator $\Omega(t_2, t_1)$ so that, for example, Eq. (9) gives directly

$$-i\int_{t_1}^{t_2} K^{(0)}(x_2, t_2; x, t)U(x, t)K^{(0)}(x, t; x_1, t_1)dxdt,$$

the expression (9) of I, while Eq. (10) translates immediately into the expression (30) of III.

As another type of application, consider two interacting systems whose hamiltonian is $H^{(a)}+H^{(b)}+U[x^{(a)}, x^{(b)}]$, where $H^{(a)}$ involves operators of system (a) only, $H^{(b)}$ involves only those of system (b), and U involves both. Then we may ask for the amplitude, if at t_1 system (a) is in state ψ_1 and system (b) in ϕ_1, that at t_2 they are in ψ_2, ϕ_2. This is the matrix element

$$m=\left\langle \psi_2\phi_2 \middle| \exp\left(-i\int_{t_1}^{t_2} H^{(a)}(t)dt - i\int_{t_1}^{t_2} H^{(b)}(t)dt - i\int_{t_1}^{t_2} U[x^{(a)}(t), x^{(b)}(t)]dt\right) \middle| \psi_1\phi_1 \right\rangle. \quad (11)$$

But this may be split into two problems. We may first find the matrix element

$$T^{(a)}[x^{(b)}(t)]=\left\langle \psi_2 \middle| \exp\left(-i\int_{t_1}^{t_2} H^{(a)}(t)dt - i\int_{t_1}^{t_2} U[x^{(a)}(t), x^{(b)}(t)]dt\right) \middle| \psi_1 \right\rangle \quad (12)$$

for the system (a) alone, considering that in the interaction potential $U[x^{(a)}, x^{(b)}]$, all operators referring to (b) are arbitrary numerical functions of t. (We have been writing as though U depends on (b) only through the coordinate, $x^{(b)}$; but the same method applies if it is also a function of momentum, or spin, or other operators on system (b).) Then the matrix element $T^{(a)}$ depends on the function $x^{(b)}(t)$. As we indicate, it is a functional of $x^{(b)}(t)$. The final answer, m, is then a matrix element $\langle\phi_2|M|\phi_1\rangle$ between the states ϕ_1 and ϕ_2;

$$m=\left\langle \phi_2 \middle| \exp\left(-i\int_{t_1}^{t_2} H^{(b)}(t)dt\right) T^{(a)}[x^{(b)}(t)] \middle| \phi_1 \right\rangle, \quad (13)$$

wherein now the quantities $x^{(b)}(t)$ are considered as ordered operators operating relative to each other and to $H^{(b)}(t)$ in accordance with the time parametrization.

In this way we can analyze one part of a pair of interacting systems without having yet analyzed the other. The influence of a on b is completely contained in the operator functional $T^{(a)}[x^{(b)}(t)]$. This separation may be useful in analysis of the theory of measurement and of quantum statistical mechanics. It is the possibility of such a separation which exists also in the case[2] of the lagrangian form of quantum mechanics, **C**, which makes that form useful in analyzing the quantum properties of the electromagnetic field. We may therefore expect that with the present operator notation it should be equally easy to make this analysis. That this is indeed true we show by example further on. Since this, the main advantage of the lagrangian form, can be so easily managed with the new notation for operators, this may well take the place of the lagrangian form in many applications. It is in some ways a more powerful and general form than the lagrangian. It is not restricted to the nonrelativistic mechanics in any way. A possible advantage of the other form at present might be a slight increase in *anschaulichkeit* offered for the interpretation of the nonrelativistic quantum mechanics.

3. DISENTANGLING AN EXPERIMENTAL FACTOR

There is one theorem which is very useful in disentangling operator expressions. We shall give it in this section. Suppose we have several operators M, N, etc. (which may also be functions of time, or more generally, the ordering parameter s), which are ordered in some way.

Let us say the functional $F[M(s), N(s)\cdots]$ defines the ordered operator. Now suppose we replace $M(s)$ by $M'(s)=U^{-1}M(s)U$, $N(s)$ by $N'=U^{-1}NU$, etc., where U is some constant operator. Then, as is well known, in $F[M(s), N(s)\cdots]$ in any product of successive operators, such as $M(s+ds)N(s)$, the UU^{-1} cancel out in between (that is, $MN=U^{-1}MUU^{-1}NU=U^{-1}MNU$, etc.), so that there results

$$F[M'(s), N'(s)\cdots]=U^{-1}F[M(s), N(s)\cdots]U, \quad (14)$$

where the U's are written to operate in the correct order. (If we wish to be more specific, we can imagine the range of the ordering parameter to be $s=0$ to 1 and write the right-hand side as $U_1^{-1}FU_0$.)

This is a simple rewriting of a well-known theorem of equivalence transformations. However, a much more interesting case is that in which $U(s)$ is actually a function of the ordering parameter. That is, we contemplate performing different transformations on the operators $M(s)$ depending on the value of s at which they are to operate. Then in a product of successive operators such as $M'(s+ds)N'(s)$, where

$$M'(s+ds)=U^{-1}(s+ds)M(s+ds)U(s+ds)$$

(operating in the order indicated by the position of U^{-1}, M and U) and $N'(s)=U^{-1}(s)N(s)U(s)$, the factors $U(s+ds)$ and $U^{-1}(s)$ will not cancel out, but we will find the operator $U(s+ds)U^{-1}(s)$ operating between times s and $s+ds$, say at $s+\frac{1}{2}ds$. If we assume U continuous, we can imagine $U(s)$ differs from $U(s+ds)$ to the first order in ds, and hence that $U(s+ds)U^{-1}(s)$ equals to first order in ds:

$$U(s+ds)U^{-1}(s)=1+P(s)ds,$$

where $P(s)$ is an operator defined by this relation in the

limit $ds \to 0$. We may write this relation

$$dU(s)/ds = P(s)U(s) \tag{15}$$

with positional ordering. Hence, between s and $s+ds$ there should operate an additional factor $1+P(s)ds$, which for convenience we may write, valid to first order, as $\exp[P(s+\frac{1}{2}ds)ds]$. The $s+\frac{1}{2}ds$ in $P(s+\frac{1}{2}ds)$ will automatically locate the factor in the correct order. But there is a factor of this kind appearing between the operators for each value of s, or multiplying the factors all together, we obtain the net factor $\exp\int_0^1 P(s)ds$, the product becoming a sum, or integral in the exponent. Hence, we have the general theorem:

$$F[M'(s), N'(s)\cdots] = U^{-1}(1)F[M(s), N(s)\cdots] \times \exp\left(\int_0^1 P(s)ds\right)U(0),$$

where

$$M'(s) = U^{-1}(s)M(s)U(s) \tag{16}$$

and

$$U(s) = \exp\left[\int_0^s P(s')ds'\right]U(0), \tag{17}$$

this last coming from integrating Eq. (15). We shall use the theorem by writing it in the form

$$\exp\int_0^1 P(s)ds F[M(s), N(s)\cdots] = U(1)F[M'(s), N'(s)\cdots]U^{-1}(0), \tag{18}$$

in which form it serves as a rule for disentangling an exponential factor from another expression. A word of caution is necessary in reading Eqs. (18), (16), and (17), for three different notations are used in the expressions. In Eq. (18) the new ordered notation is used in its complete form; for example, the s in $\exp\int_0^1 P(s)ds$ gives the order in which the P is to operate relative to the M, N of the functional F which it multiplies. In Eq. (16), however, all the operators are to operate at s, but the relative order in M' of U, M, and U^{-1} is as given by the usual position convention. Finally, Eq. (17) would be less ambiguous if it were replaced by the differential equation (15). For in the solution (17), the s' are to bear no relation to the s in Eq. (16) or Eq. (18). The operator $U(s)$ is to be computed from P by Eq. (17) first, then the whole operator $U(s)$ is to operate in Eq. (16), and then in Eq. (18) at the position s.

We shall use this theorem in several applications related to quantum electrodynamics. Most particularly, we shall find a certain special case useful enough to warrant special mention. It is the case that $P(s)$ is of the form $\alpha(s)P_s$, where $\alpha(s)$ is a simple numerical function, and P_s is an operator whose form does not depend on s but whose order of operation does. Then if we call $a(s) = \int_0^s \alpha(s')ds'$, so that $a(s)$ is also a numerical function, Eq. (17) gives $U(s) = \exp[a(s)P_s]U(0)$. We shall further choose to specialize $U(0) = 1$. (The more general case corresponds to a final simple constant equivalence transformation (14) with $U(0)$.) Then our theorem may be written

$$\exp\left[\int_0^1 \alpha(s)P_s ds\right]F[M(s), N(s)\cdots] = \exp\left[P_1\int_0^1 \alpha(s)ds\right]F[M'(s), N'(s)\cdots], \tag{19}$$

where

$$M'(s) = \exp\left[-P_s\int_0^s \alpha(s')ds'\right]M(s) \times \exp\left[+P_s\int_0^s \alpha(s')ds'\right]. \tag{20}$$

Further, since this theorem with Eq. (20) substituted into Eq. (19) is valid when $\alpha(s)$ is an *arbitrary* numerical function, it is also true if $\alpha(s)$ is any ordered operator $\alpha(s)$ commuting with P for all s, provided that in *all* expressions involving α, the parameter s or s' is consistently interpreted as giving the order in which the α operates.[6]

The mathematical proof of the theorems (18) and (19) offered here is admittedly very sketchy; but since the theorems are true, it should not be hard to supply them with more satisfactory demonstrations (see Appendix A for an alternative demonstration).

There are a number of other interesting relations which we may derive from Eq. (19), but which we shall not need in this paper. One is included here because it has been found useful in certain other applications. If $\alpha(s)$ is considered infinitesimal in Eqs. (19) and (20), expansion in first order in α gives the following result (or differentiate each side with respect to $\alpha(t)$ and set $\alpha(s) = 0$),

$$P_t F[M(s)] = P_1 F[M(s)] - \int_t^1 (PM - MP)_s \delta F/\delta M(s) ds$$

(if we assume F can be represented by a functional having a derivative $\delta F/\delta M(s)$). We have taken F to depend only on one operator $M(s)$, but the generalization is clear. Here, $(PM-MP)_s$ is conventional ordering is $PM(s) - M(s)P$ and is considered to act as an entity at s. The differential form

$$(dP_t/dt)F[M(s)] = (PM - MP)_t \delta F/\delta M(t) \tag{21}$$

is also useful.

[6] To simplify such descriptions, in a situation involving two sets of operators, any one of the first set commuting with any one of the second, it is often convenient to generalize to the use of two different ordering parameters—one for first set, and one for the second.

4. THE INTERACTION REPRESENTATION

As a first simple direct application of our theorem consider again the perturbation problem (8) of computing the operator

$$\Omega^{(U)}(t_2, t_1) = \exp\left[-i\int_{t_1}^{t_2} H^{(0)}(t)dt\right] \exp\left[-i\int_{t_1}^{t_2} U(t)dt\right]. \quad (8)$$

If we suppose the properties of $H^{(0)}$ to be known and simple, the right side of Eq. (8) may be disentangled by means of our theorem (18). We consider $-iH^{(0)}(t)$ as an operator $P(s)$ and

$$\exp\left[-i\int_{t_1}^{t_2} U(t)dt\right]$$

as the functional F from which the

$$\exp\left[-i\int H^{(0)}(t)dt\right]$$

is to be disentangled. Hence, a direct application of Eq. (18) gives

$$\exp\left[-i\int_{t_1}^{t_2} H^{(0)}(t)dt\right] \exp\left[-i\int_{t_1}^{t_2} U(t)dt\right] = S(t_2) \exp\left[-i\int_{t_1}^{t_2} U'(t)dt\right] S^{-1}(t_1), \quad (22)$$

where

$$S(t) = \exp\left[-i\int_a^t H^{(0)}(t')dt'\right]$$

(the lower limit a used in defining S is arbitrary; it may be taken as t_1 so $S(t_1)=1$ if that is convenient), and

$$U'(t) = S^{-1}(t)U(t)S(t) \quad (23)$$

(operating in positional order). If we take matrix elements not between states ψ_1 and ψ_2 but between $\psi_1' = S^{-1}(t_1)\psi_1$ and $\psi_2' = S^{-1}(t_2)\psi_2$, we may call the Ω-matrix Ω' and omit the $S(t_2)$ and $S^{-1}(t_1)$ factors in Eq. (22). These new time-dependent states ψ' are evidently states that would give rise to ψ_1 at t_1 and ψ_2 at t_2 (from some fixed reference time a) if the perturbation were not acting. Then the time-dependent perturbation theory simply comes to evaluating

$$\Omega'(t_2, t_1) = \exp\left(-i\int_{t_1}^{t_2} U'(t)dt\right). \quad (24)$$

Expansion in power series, substitution of U' from Eq. (23), and use of the relation $\Omega^{(0)}(t', t'') = S(t')S^{-1}(t'')$ leads immediately to the formulas (9) and (10), so that Eqs. (23) and (24) give the simplest form to the time-dependent perturbation theory. Of course, the same results may be obtained by a unitary transformation in the conventional way. Ordinarily, result (24) is not written in this way, for it involves the time convention on the ordering of the operators. (It is usually expressed as a differential equation for Ω'.) If the perturbation U represents an interaction between some systems described by $H^{(0)}$, the reduction of Eq. (8) to Eq. (24) is called passing to the interaction representation.

5. SYSTEM COUPLED TO AN HARMONIC OSCILLATOR

As a further example of the use of the notation we solve completely the problem of a particle or system of particles coupled linearly to an harmonic oscillator. This problem in greater generality is the main problem of quantum electrodynamics. It has been thoroughly studied in III, but we solve it again as an illustration of the new notation. Let the hamiltonian of the combined system be

$$H = H_p(t) + H_{\text{osc}} - \Gamma(t)q,$$

where H_{osc} is the hamiltonian of the oscillator alone,

$$H_{\text{osc}} = (1/2m)(p^2 + \omega^2 q^2),$$

where p is the momentum conjugate to q, the coordinate of the oscillator. Further, H_p, which may depend explicitly on time, is the hamiltonian of the particles, and Γ may contain any operators pertaining to the particles as well as possibly being an explicit function of the time. We ask for the matrix element for finding the particles in state $\chi_{t''}$ and the oscillator in some eigenstate m at time t'', if at a previous time t' the particles are in state $\chi_{t'}$, and the oscillator in its nth eigenstate. It is the matrix element

$$\left\langle \chi_{t''}\phi_m \middle| \exp -i\int_{t'}^{t''} [H_p(t) + H_{\text{osc}}(t) - 1\,(t)q(t)]dt \middle| \phi_n\chi_{t'} \right\rangle$$

using the time ordering convention. As already discussed in Sec. 2, this can be considered as the matrix element between states $\chi_{t'}$ and $\chi_{t''}$ of the matrix

$$M = \exp\left(-i\int_{t'}^{t''} H_p(t)dt\right) G_{mn}, \quad (25)$$

where G_{mn} (the analog of T of Sec. 2), a functional of $\Gamma(t)$, serves to define the net effect on the particles of their interaction with the oscillator. Calculation of G_{mn} means evaluating

$$G_{mn} = \left\langle \phi_m \middle| \exp -i\int_{t'}^{t''} H_{\text{osc}}(t)dt \cdot \exp i\int_{t'}^{t''} \Gamma(t)q(t)dt \middle| \phi_n \right\rangle \quad (26)$$

in a general way as a functional of $\Gamma(t)$. We are to consider $\Gamma(t)$ here as a simple numerical function, and

only later utilize the fact that it is an operator involving the particles when we go to evaluate the matrix M, Eq. (25), between the particle states $\chi_{t'}$ and $\chi_{t''}$. The evaluation of G_{mn} for an arbitrary numerical function $\Gamma(t)$ may be performed in a variety of ways. One is by the lagrangian form of quantum mechanics given explicitly in III, Sec. 3, with the sole difference (which is unessential for this part of the problem) that there $\Gamma(t)$ was called $\gamma(t)$ and was a functional of the coordinates $x(t)$ of the particles, while here we see we are in a more general position as $\Gamma(t)$ may be a functional of any ordered operators referring to the particles. We find, for example [III, Eq. (14)],

$$G_{00}=\exp\left[-(1/4\omega)\int_{t'}^{t''}\int_{t'}^{t''}\exp(-i\omega|t-s|)\times\Gamma(t)\Gamma(s)dtds\right]. \quad (27)$$

The same result may also be obtained by a direct solution of the Schrödinger equation for the forced oscillator. The great advantage of the operator notation is to allow this formal solution for an oscillator forced by an *arbitrary potential function* $\Gamma(t)$ to be equally useful when the oscillator is actually in interaction with a quantum-mechanical system!

Thus, we have the answer for G_{mn} in III, Eq. (57), using Γ for γ. It is, however, interesting to see how this expression for G_{mn} could be worked out directly using the methods of the ordered operator calculus. We want to disentangle the operator

$$G=\exp\left(-i\int_{t'}^{t''}H_{\mathrm{osc},t}dt\right)\exp\left(i\int_{t'}^{t''}\Gamma(t)q_t dt\right). \quad (26')$$

Let us call, in the usual way, $Q^*=(\frac{1}{2}\omega)^{\frac{1}{2}}(q-i\omega^{-1}p)$ and $Q=(\frac{1}{2}\omega)^{\frac{1}{2}}(q+i\omega^{-1}p)$ the creation and annihilation operators. They satisfy the commutation relation

$$QQ^*-Q^*Q=1. \quad (28)$$

In terms of them $H_{\mathrm{osc}}=\frac{1}{2}\omega(Q^*Q+QQ^*)$ and

$$q=(2\omega)^{-\frac{1}{2}}(Q+Q^*). \quad (29)$$

Now as a first step we pass to the interaction representation (Sec. 4). We use the theorem (20) with $P_s=-iH_{\mathrm{osc}}$, $\alpha(s)=1$, to disentangle the $\exp[-i\int H_{\mathrm{osc}}(t)dt]$ factor, obtaining

$$G=S(t'')\exp\left\{i(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\Gamma(t)[Q'(t)+Q'^*(t)]dt\right\}S^{-1}(t'),$$

where $Q'(t)=S^{-1}(t)QS(t)$, $Q'^*(t)=S^{-1}(t)Q^*S(t)$, and $S(t)=\exp(-itH_{\mathrm{osc}})$. By redefining the wave functions so they contain $S(t)$, or, for eigenstates, the energy factors $\exp(-iE_nt)$ for the free oscillator, we can eliminate the $S(t'')$ and $S^{-1}(t')$ and need merely calculate the matrix element of

$$G'=\exp i(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\Gamma(t)[Q'(t)+Q'^*(t)]dt.$$

From Eq. (28) we readily calculate that[7]

$$Q'(t)=Qe^{-i\omega t}\quad\text{and}\quad Q'^*(t)=Q^*e^{+i\omega t}, \quad (30)$$

so that the problem becomes the disentanglement of

$$G'=\exp\left[i(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\Gamma(t)e^{+i\omega t}Q_t^*dt\right]\times\exp\left[i(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\Gamma(s)e^{-i\omega s}Q_s ds\right].$$

We shall find it most convenient to disentangle this into a form in which all the annihilation operators operate first, and then come the creation operators (since the nth state cannot suffer more than n annihilation operators, $Q^{n+1}\phi_n$ vanishing, the expression will be easy to evaluate in this form). To this end let us use theorem (19) again, this time with $P=Q^*$,

$$\alpha(s)=i(2\omega)^{-\frac{1}{2}}\Gamma(t)e^{+i\omega t}.$$

Calling, temporarily,

$$A(t)=i(2\omega)^{-\frac{1}{2}}\int_{t'}^{t}\Gamma(t)e^{+i\omega t}dt,$$

we find

$$G'=\exp[A(t'')Q_{t''}^*]\exp\left[i(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\Gamma(t)e^{-i\omega t}Q''(t)dt\right],$$

where

$$Q''(t)=\exp[-A(t)Q^*]Q\exp[+A(t)Q^*]. \quad (31)$$

The commutation relation (28) here gives[8]

$$Q''(t)=Q_t+A(t), \quad (32)$$

so that

$$G'=\exp[A(t'')Q_{t''}^*]\times\exp\left[i(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\Gamma(t)e^{-i\omega t}\{Q_t+A(t)\}dt\right].$$

In the last factor Q_t can be replaced by $Q_{t'}$ since t is in any case less than t'', so that all the Q_t's come before the $Q_{t''}^*$, and Q_t need not be ordered relative to itself as it is a constant operator. Hence, we may, with a slight rewriting (for example, $i\beta^*$ for $A(t'')$), write

$$G'=\exp(i\beta^*Q_{t''}^*)\exp(i\beta Q_{t'})G_{00} \quad (33)$$

[7] For, $Q'(t)=\exp(itH_{\mathrm{osc}})Q\exp(-itH_{\mathrm{osc}})$ implies $dQ'/dt=iS^{-1}(t)\times(H_{\mathrm{osc}}Q-QH_{\mathrm{osc}})S(t)=-i\omega Q'(t)$, since $H_{\mathrm{osc}}Q-QH_{\mathrm{osc}}=-\omega Q$ by Eq. (28). Thus, since $Q'(0)=Q$, one obtains $Q'(t)=e^{-i\omega t}Q$.

[8] For, differentiation of the expression for $Q''(t)$ gives $dQ''(t)/dt=-A'(t)\exp(-A(t)Q^*)(Q^*Q-QQ^*)\exp(+A(t)Q^*)=A'(t)$ by Eq. (28), so integration gives $Q''(t)=A(t)+Q$, since $Q''(t')=Q$, inasmuch as $A(t')=0$.

with G_{00} equal to $\exp i(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\Gamma(t)e^{-i\omega t}A(t)dt$ and therefore identical to Eq. (27), and with

$$\beta=(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\Gamma(t)e^{-i\omega t}dt,$$
$$\beta^*=(2\omega)^{-\frac{1}{2}}\int_{t'}^{t''}\Gamma(t)e^{+i\omega t}dt, \tag{34}$$

just as in III, Eq. (58). The operator G' is now completely disentangled. Its matrix element between n and m we call G_{mn}. The matrix element may be evaluated by ordinary methods, since the t' and t'' in Q and Q^*, respectively, in Eq. (33) are unnecessary if the positional notation is used. That the element for $n=0$, $m=0$ is just what we call G_{00} is evident, for if $\exp(i\beta Q)$ be expanded as $1+\beta Q+\beta^2Q^2\cdots$ and the result applied to ϕ_0, all the terms beyond the first give zero for $Q\phi_0=0$. Thus, this exponential may effectively be replaced by unity. Likewise, the second can be replaced by unity for $\phi_0{}^*Q^*=0$.

The case of more general values of m, n, may be worked out by writing

$$\phi_n=(n!)^{-\frac{1}{2}}Q^{*n}\phi_0, \tag{35}$$

so that

$$G_{mn}=\langle\phi_0|\,(m!)^{-\frac{1}{2}}(n!)^{-\frac{1}{2}}Q^me^{i\beta^*Q^*}e^{i\beta Q}Q^{*n}\,|\,\phi_0\rangle G_{00}. \tag{36}$$

Then, since $e^{i\beta Q}Q^*=(Q^*+i\beta)e^{i\beta Q}$ (as in Eqs. (31), (32)), repetition n times gives $e^{i\beta Q}Q^{*n}=(Q^*+i\beta)^ne^{i\beta Q}$, and likewise $Q^me^{i\beta^*Q^*}=e^{i\beta^*Q^*}(Q+i\beta^*)^m$. We find

$$G_{mn}=\langle\phi_0|\,(m!)^{-\frac{1}{2}}(n!)^{-\frac{1}{2}}e^{i\beta^*Q^*}(Q+i\beta^*)^m \times(Q^*+i\beta)^ne^{i\beta Q}\,|\,\phi_0\rangle G_{00}. \tag{37}$$

The exponentials may now be replaced by unity as previously discussed. The other factors expanded by the binomial theorem give

$$G_{mn}=(m!)^{-\frac{1}{2}}(n!)^{-\frac{1}{2}}\sum_r\binom{m}{r}\sum_s\binom{n}{s}(i\beta^*)^{m-r} \times(i\beta)^{n-s}\langle\phi_0|\,Q^rQ^{*s}\,|\,\phi_0\rangle G_{00}.$$

The next to last factor by Eq. (35) is

$$(s!)^{\frac{1}{2}}(r!)^{\frac{1}{2}}\langle\phi_r|\,\phi_s\rangle=(s!)^{\frac{1}{2}}(r!)^{\frac{1}{2}}\delta_{rs},$$

so that finally

$$G_{mn}=(m!n!)^{-\frac{1}{2}}\sum_r\binom{m}{r}\binom{n}{r}r!(i\beta^*)^{m-r}(i\beta)^{n-r}G_{00} \tag{38}$$

as in III, Eq. (57).

Having this form for the behavior of a system of particles interacting with a single oscillator, we could go on and discuss the quantum electromagnetic field as a set of such oscillators. It is evident that to do so would be simply to repeat the steps described in III, Sec. 4, using Γ for γ and reinterpreting the symbols as ordered operators rather than as amplitudes associated with a path. The result in general is Eq. (43) below [in agreement with III, Eq. (48)], and there is no need to go into the details again of summing the effects of all the oscillators to obtain this result. We will pass directly to a discussion of the complete electromagnetic field.

6. QUANTUM ELECTRODYNAMICS

There are available several equivalent formulations of quantum electrodynamics.[1,2,9] We shall give a very brief outline of their interrelationships using the ordered operator notation. We can start with the usual formalism of Heisenberg, Pauli, and Dirac.[9] The wave function of the system, consisting of the electron-positron field and of the electromagnetic field in interaction, satisfies a wave equation $i\partial\psi/\partial t=H\psi$, where the hamiltonian for the system may be written $H=H_m+H_f+H_i$, where H_m is that of the electron-positron field free of potentials, H_f is that of the electromagnetic field in empty space, and H_i represents the interaction of the two fields. The problem is to obtain the wave function at time t_2 in terms of its value at a previous time t_1. It is therefore a study of the operator

$$\exp\left\{-i\int_{t_1}^{t_2}[H_m(t)+H_f(t)+H_i(t)]dt\right\}.$$

We can simplify this by first disentangling the exponential factor

$$\exp\left\{-i\int_{t_1}^{t_2}[H_m(t)+H_f(t)]dt\right\}.$$

That is, we go directly to the interaction representation, and find that we must analyze

$$\exp\left\{-i\int_{t_1}^{t_2}H_i'(t)dt\right\}.$$

We shall always use the interaction representation and shall omit the prime here for simplicity of notation. Furthermore, it will be sufficient for our purpose to consider only the case $t_1\to-\infty$ and $t_2\to+\infty$, so that quantum electrodynamics is a study of the operator[10]

$$S=\exp\left[-i\int_{-\infty}^{\infty}j_\mu(\mathbf{x},t)A_\mu(\mathbf{x},t)d^3\mathbf{x}dt\right], \tag{39}$$

where the $A_\mu(1)$ is the operator potential of the electromagnetic field and $j_\mu(1)$ is the operator current of the

[9] See, for example, P. A. M. Dirac, *The Principles of Quantum Mechanics*, (The Clarendon Press, Oxford, 1947), third edition, Chapter 12.

[10] For systems with particles of spin zero or one, S may be written in this same form by use of the Kemmer-Duffin matrices β_μ, as is shown by C. N. Yang and D. Feldman, Phys. Rev. 79, 972 (1950), for example, Eq. (33). Thus, all of these results given here for the Dirac field are equally correct for spin zero or one if γ_μ is replaced by β_μ. See also M. Neumann and W. H. Furry, Phys. Rev. 76, 1677 (1949), and R. Moorhouse, Phys. Rev. 76, 1691 (1949).

Dirac electron-positron field. They, of course, commute with each other, since they refer to different systems. Further,[11] $A_\mu(1)$, $A_\mu(2)$ commute if 1 and 2 are separated by a spacelike interval, as do $j_\mu(1)$, $j_\mu(2)$. In the expression for S the operators are ordered in accordance with the time t.

We may thus define the problem of quantum electrodynamics as a study of the operator S. Let us imagine for purposes of discussion that we disregard the derivation of S given in the preceding paragraph. We imagine the problem is given directly as the analysis of the operator S defined in Eq. (39) (assuming the commutation rules, reference 11). Let us see how the various formalisms are simply different ways of expressing or analyzing S.

First, we might try to define S in some way which would not require the use of the ordering notation. Suppose we split the range of integration of t into two regions $-\infty$ to τ and τ to ∞. Then the integral may be split into two parts. We can write the factors, as

$$S=\exp\left[-i\int_\tau^\infty j_\mu(\mathbf{x},t)A_\mu(\mathbf{x},t)d^3\mathbf{x}dt\right]\times\exp\left[-i\int_{-\infty}^\tau j_\mu(\mathbf{x}',t')A_\mu(\mathbf{x}',t')d^3\mathbf{x}'dt'\right].$$

Now, since $t'<t$, all the operators on the last factor act before those of the first factor, so they are disentangled relative to the first factor. Hence, we are led to define an operator function of τ,

$$\Omega(\tau)=\exp\left[-i\int_{-\infty}^\tau j_\mu(\mathbf{x},t)A_\mu(\mathbf{x},t)d^3\mathbf{x}dt\right].$$

If τ is changed to $\tau+d\tau$, an additional factor appears operating in front of all other $t\lessgtr\tau$, namely, $\exp[-i\Delta\tau\int j_\mu(\mathbf{x},\tau)A_\mu(\mathbf{x},\tau)d^3\mathbf{x}]$. Hence, $\Omega(\tau)$ satisfies the differential equation

$$id\Omega/d\tau=\left[\int j_\mu(\mathbf{x},\tau)A_\mu(\mathbf{x},\tau)d^3\mathbf{x}\right]\Omega(\tau),\qquad(40)$$

the operators operating in positional order.

Thus, we are lead to a differential equation, the solution of which can be used to define S (for S is $\Omega(\tau)$ as $\tau\to+\infty$ when $\Omega(\tau)$ is that solution of Eq. (40) which $\to I$ as $\tau\to-\infty$). If we define $\psi(-\infty)$ as an initial state wave function, clearly, $\psi(\tau)=\Omega(\tau)\psi(-\infty)$ satisfies the same equation as Ω. This is the Schrödinger equation in the usual formulation if written in interaction representation. (We probably would not be led to go back to the ordinary representation as this is an unnecessary increase in complexity.)

The apparent lack of covariance implied by using time to define the differential equation can be remedied by analyzing S in a slightly different manner, suggested by Tomanaga and by Schwinger.[1]

The variables $\mathbf{x}$, t over which one integrates in Eq. (39) may be divided into two groups in another way; those previous to and those following an arbitrary spacelike surface σ:

$$S=\exp\left[-i\int_b j_\mu(\mathbf{x},t)A_\mu(\mathbf{x},t)d^3\mathbf{x}dt\right]\times\exp\left[-i\int_a j_\nu(\mathbf{x}',t')A_\nu(\mathbf{x}',t')d^3\mathbf{x}'dt'\right],$$

where the region a of integration of the second factor are those points of space-time previous to σ, while b are those following σ. Now again the factors are disentangled. It might at first be argued that since there are some values of t' greater than t, the corresponding operators in $\mathcal{J}_a$ should follow, not precede, those in $\mathcal{J}_b$. But for those t' which exceed t, the points $\mathbf{x}$, $\mathbf{x}'$ are separated by a spacelike interval (as σ is a spacelike surface); hence the order of the $A_\mu(\mathbf{x},t)$ and $A_\nu(\mathbf{x}',t')$ as well as of $j_\mu(\mathbf{x},t)$ and $j_\nu(\mathbf{x}',t')$ is irrelevant, as these commute. Hence, the operators are, in fact, disentangled; and we can define

$$\Omega(\sigma)=\exp\left[-i\int_a j_\mu(\mathbf{x},t)A_\mu(\mathbf{x},t)d^3\mathbf{x}dt\right]$$

as an operator defined as a functional of the surface σ. A small change in surface at $\mathbf{x}$, t changes the operator by

$$\delta\Omega(\sigma)/\delta\sigma(\mathbf{x},t)=-ij_\mu(\mathbf{x},t)A_\mu(\mathbf{x},t)\Omega(\sigma),\qquad(41)$$

the equation of Schwinger[12] for $\Omega(\sigma)$ (and also for $\psi(\sigma)$ defined by $\Omega(\sigma)\psi(-\infty)$). Again, S is $\Omega(\sigma)$ as the surface σ is removed to $+\infty$.

These differential equations (40) or (41) are therefore needed to define the operator S if one is limited to conventional notation. The form (41) has the advantage of putting the relativistic invariance more into evidence. However, the solution (39) is common to both and is more easily used. It is likewise evidently invariant if we write it

$$S=\exp\left[-i\int j_\mu(1)A_\mu(1)d\tau_1\right]\qquad(42)$$

(with the point 1 representing $\mathbf{x}_1$, t_1 and $d\tau_1=d^3\mathbf{x}_1dt_1$) and assume the convention here that *if two operators in Eq. (42) correspond to points separated by either a timelike or a zero interval, that operates first which corresponds*

[11] See, for example, J. Schwinger, Phys. Rev. **74**, 1439 (1948). In his notation (except that we put a factor e in A_μ rather than j_μ), the commutation relations are [his Eqs. (2.28)and (2.29)]

$$[A_\mu(x),A_\nu(x')]=4\pi e^2i\delta_{\mu\nu}D(x-x')$$

and $j_\mu(x)=\bar\psi(x)\gamma_\mu\psi(x)$ with $\{\psi_\alpha(x),\bar\psi_\beta(x')\}=-iS_{\alpha\beta}(x-x')$ if no external potential is acting. Other combinations commute.

[12] I. Schwinger. Phys. Rev. **74**, 1439 (1948).

to the earlier time. If they are separated by a spacelike interval, no definition is necessary, for they commute.

The other developments consist in methods of actually evaluating Eq. (42), given the commutation relations[11] of the $A_\mu(1)$. The method explained by Dyson[3] consists of making a power series expansion of S and disentangling it term by term. For example, the second-order term is

$$-\tfrac{1}{2}\int j_\mu(1)A_\mu(1)d\tau_1\int j_\nu(2)A_\nu(2)d\tau_2.$$

This term may then be analyzed into the conventional notation by reordering the operators. In this example it is necessary merely to break the region of integration in t_2 up into two, $t_2<t_1$ and $t_2>t_1$. Actually, because of the symmetry they give equal contributions, so that the result is

$$-\int\int j_\mu(1)A_\mu(1)d^3\mathbf{x}_1dt_1\int^{t_1}\int j_\nu(2)A_\nu(2)d^3\mathbf{x}_2dt_2,$$

the ordering now being conventional. From here the matrix elements are computed between given states by use of the commutation relations (46) below. For further details we refer to Dyson's papers.[3] The result is that given by the rules of II.

Another method is to notice that the entire dependence of S on A_μ can be directly evaluated. As far as the states of the field are concerned, the evaluation of matrix elements of S is exactly the same as though $j_\mu(1)$ were a numerical function (since it commutes with all $A_\mu(1)$). Hence, these may be worked out by first obtaining the result for a field interacting with a given unquantized current distribution $j_\mu(1)$. This can be done, for example, by using the lagrangian methods described in III. For example, the matrix taken between states in which the field is empty of photons initially and finally is

$$S_{00}=\exp\left[-\tfrac{1}{2}ie^2\int\int j_\mu(1)j_\mu(2)\delta_+(s_{12}^2)d\tau_1d\tau_2\right] \quad (43)$$

as is shown in III (for j_μ a numerical function). This may now be interpreted as follows: The matrix element of S for a transition in which at $t=-\infty$ there are no real photons and the matter is in state χ_-, to the state at $+\infty$ also empty of photons with the matter in state χ_+, is the matrix element of S_{00} between χ_- and χ_+, where S_{00}, given in Eq. (43), operates now only on matter variables, the order of operators $j_\mu(1)$, $j_\mu(2)$ being determined just as in Eq. (42). This expression forms the basis for the author's treatment of virtual photon processes (II).

If an additional unquantized potential $B_\mu(1)$ is present, the expression (42) for S is altered just by the replacement of $A_\mu(1)$ by $A_\mu(1)+B_\mu(1)$.

The matrix corresponding to Eq. (43) would be a functional of $B_\mu(1)$ and a function of e^2:

$$S_{e^2}[B]=\exp\left[-\tfrac{1}{2}ie^2\int\int j_\mu(1)j_\mu(2)\delta_+(s_{12}^2)d\tau_1d\tau_2\right]$$

$$\times\exp\left[-i\int j_\nu(1)B_\nu(1)d\tau_1\right]. \quad (44)$$

It is evident by direct substitution, that $S_{e^2}[B]$ satisfies

$$dS_{e^2}/d(e^2)=\tfrac{1}{2}i\int\int\delta^2S_{e^2}/\delta B_\mu(1)\delta B_\mu(2)\delta_+(s_{12}^2)d\tau_1d\tau_2. \quad (45)$$

Since the equation is linear, any matrix element of S_{e^2}, say $T_{e^2}[B]$, between two states of the matter satisfies same equation. This is Eq. (45) of III, which is shown in III to be a general statement of the rules given in II for solving electrodynamic problems. Evidently, the case of real photons in initial or final state can be carried through in parallel to the discussion in III, with $j_\mu(1)$ now as an operator.

This completes our discussion in a general way of the relations between the various representations of electrodynamics. However, we wish to add a word concerning the derivation of Eq. (43). We have indicated how this may be done using the lagrangian method. However, we have seen from our example with the single forced oscillator that the same results may be obtained directly with the operator method, in just as simple a manner. Of course, by considering the field as a set of such oscillators we will arrive at Eq. (43), thus completely avoiding the lagrangian formulation. However, since the relation between Eqs. (42) and (43) is so fundamental, we should like to show how the operator method permits a simple direct passage from Eq. (42) to Eq. (43).[13] (We are simply following the steps leading from Eq. (26′) to Eq. (33) for the single harmonic oscillator, but are using $A_\mu(1)$ to replace q.)

The field operator $A_\mu(1)$ can be split into two parts $A_\mu(1)=A_\mu^+(1)+A_\mu^-(1)$, where the first $A_\mu^+(1)$ annihilates photons, and the second $A_\mu^-(1)$ creates them.[14] They satisfy the commutation relations (positional

[13] We omit the usual extra complications in all such demonstrations concerned with showing that disregard of the supplementary conditions on $\partial A_\mu/\partial x_\mu$ is legitimate.

[14] Ordinarily, the field operator $A_\mu(1)$ is expanded into modes $A_\mu(1)=\Sigma_iA_{\mu i}(x)(Q_i^*e^{i\omega_it}+Q_ie^{-i\omega_it})$, where $A_{\mu i}$ is the numerical function [for example, cosines or sines, III Eq. (1)] describing the classical mode i of frequency ω_i and Q_i^*, Q_i are the creation and annihilation operators into which q_i, the coordinate of the oscillator of this mode, has been split (29). The factors $e^{\pm i\omega t}$ result from use of interaction representation (30). Then, we have $A_\mu^-(1)=\Sigma_iA_{\mu i}(x)Q_i^*e^{i\omega_it}$ and $A_\mu^+(1)=\Sigma_iA_{\mu i}(x)Q_ie^{-i\omega_it}$. The commutation rule (46) then results from that of the Q and Q^* (28). Using the representation of III Eq. (1), the right-hand side of Eq. (46) comes from Eq. (28) directly in the form

$$(2\pi)^{-2}e^2\delta_{\mu\nu}\int\exp[-ik(t_2-t_1)]\cos(\mathbf{K}\cdot\mathbf{x}_1-\mathbf{K}\cdot\mathbf{x}_2)d^3\mathbf{K}/k,$$

which on integration for $t_2>t_1$ [see III Eq. (22)] is Eq. (46). The separation has been accomplished directly in coordinate space by J. Schwinger, Phys. Rev. 75, 651 (1949).

ordering)

$$A_\mu^-(1)A_\nu^+(2)-A_\nu^+(2)A_\mu^-(1)=-ie^2\delta_{\mu\nu}\delta_+(s_{12}^2) \quad (46)$$

for[15] $t_1<t_2$. We have set $s_{12}^2=(x_{1\mu}-x_{2\mu})(x_{1\mu}-x_{2\mu})$.

On the basis of this commutation rule, we are to disentangle the operator

$$S=\exp\left[-i\int j_\mu(1)A_{\mu t_1}^-(1)d\tau_1\right]\times\exp\left[-i\int j_\nu(2)A_{\nu t_2}^+(2)d\tau_2\right], \quad (42')$$

where for definiteness we indicate the time of operation by the subscript. This is already of the exponential form of theorem (18), using $-i\int j_\mu(1)A_{\mu t_2}^-(1)d^3\mathbf{x}_1$ as $P(s)$, $s=t_2$. Hence, the result is

$$S=\exp\left[-i\int j_\mu(1)A_{\mu\infty}^-(1)d\tau_1\right]\times\exp\left[-i\int j_\nu(2)A_{\nu t_2}^{\prime\prime+}(2)d\tau_2\right],$$

where

$$A_{\nu t_2}^{\prime\prime+}(2)=\exp\left[+i\int_{-\infty}^{t_2}j_\mu(1)A_\mu^-(1)d\tau_1\right]A_{\nu t_2}^+(2)\times\exp\left[-i\int_{-\infty}^{t_2}j_\mu(1')A_\mu^-(1')d\tau_{1'}\right], \quad (47)$$

where in Eq. (47) we suppress the ordering rules for A^+, A^- and use instead positional notation (but maintain the rules for j_μ).

The commutation rule (46) permits Eq. (47) to be written[16]

$$A_{\nu t_2}^{\prime\prime+}(2)=A_{\nu t_2}^+(2)+\int_{-\infty}^{t_2}e^2\delta_+(s_{12}^2)j_\mu(1)d\tau_1. \quad (48)$$

Hence, we have

$$S=\exp\left[-i\int j_\mu(1)A_{\mu\infty}^-(1)d\tau_1\right]\times\exp\left[-\tfrac{1}{2}ie^2\int\int j_\mu(1)j_\mu(2)\delta_+(s_{12}^2)d\tau_1d\tau_2\right]\times\exp\left[-i\int j_\mu(2)A_{\mu-\infty}^+(2)d\tau_2\right], \quad (49)$$

the A operators being entirely disentangled (the j's are still entangled). The ordering index t_2 on $A_{\mu t_2}^+$ has been changed to $-\infty$ in Eq. (49), since all the A^+ commute and act before $A_{\mu\infty}^-$, so that no ordering is necessary.

Taken between states empty of photons the result is just S_{00} of Eq. (43), for the annihilation operation A^+ on the state of zero photons is zero, and creation operation of A^- has zero amplitude of leaving a state without photons. If there is one photon present initially and we ask that no photons remain, we shall have to annihilate it and create none, so that if the A^- and A^+ exponentials are expanded in power series, we must take only the term linear in A^+ and independent of A^-. This is equivalent to a first-order action of the potential B_μ in Eq. (44) in perturbation. The corresponding rules for higher numbers of real photons are readily derived from Eq. (49). In this way we have completed an independent deduction of all the main formal results in quantum electrodynamics, by use of the operator notation.

7. THE DIRAC EQUATION

Up to now we have discussed the matter system using the description of second quantization. It was pointed out in I in the case of the electron-positron field where a small number of charges is involved, another simple interpretation is available. In this section we should like to discuss this from an operator point of view and to give in the following sections the formulas in this picture for electrons interacting through the agency of the electromagnetic field.

We begin by discussing the behavior of a single charge (plus the virtual pairs produced from it) in an unquantized potential $\boldsymbol{B}=\gamma_\mu B_\mu$, omitting the contributions from closed loop diagrams. This section will therefore constitute a brief summary of I using operator notation.

The behavior of a single charge is obtained by solving the Dirac equation

$$(i\nabla-\boldsymbol{B}-m)\psi=0 \quad (50)$$

with suitable boundary conditions and interpreting the solution as described in I. For convenience we shall always solve, instead,

$$(i\nabla-\boldsymbol{B}-m)\psi=iF, \quad (51)$$

where F is a source function, by writing

$$\psi=(i\nabla-\boldsymbol{B}-m)^{-1}iF \quad (52)$$

and interpreting the reciprocal operator in the definite sense implied by the limit of the operator when m has a vanishingly small negative imaginary part. If, for example, we wish the ordinary solution for $t>t_0$ which at $t=t_0$ has the form $f(\mathbf{x})$ representing an electron (i.e., $f(\mathbf{x})$ has only positive energy components), that solution

[15] This restriction at first sight looks unrelativistic. For $t_1>t_2$ we would have the complex conjugate of $-ie^2\delta_{\mu\nu}\delta_+(s_{12}^2)$, but $-i\delta_+(s_{12}^2)$ is real in spacelike regions (as $\delta_+(x)=\delta(x)-i\pi x^{-1}$).

[16] For, if $A_{\nu t_2}^{\prime\prime+}(2)$ of (47) is considered as a functional of j_μ its first variation with respect to $j_\mu(3)$ is $(t_3<t_2)$.

$$i\exp\left(+i\int_{-\infty}^{t_2}j_\mu A_\mu^-d\tau\right)[A_\mu^-(3)A_\nu^+(2)-A_\nu^+(2)A_\mu^-(3)]\times\exp\left(-i\int_{-\infty}^{t_2}j_\mu A_\mu^-d\tau\right)=e^2\delta_{\mu\nu}\delta_+(s_{12})$$

by Eq. (46). The first variation of expression (48) gives the same result, so that Eq. (48) is correct for all j_μ, since it obviously is correct for $j_\mu=0$.

is the ψ obtained from Eq. (52) by setting[17] $F(1)=\gamma_4\delta(t_1-t_0)f(\mathbf{x}_1)$. If f contains negative energy components, Eq. (52) gives the desired solution for these components for $t<t_0$.

From the definition of $K_+^{(B)}(2,1)$ in I, Eq. (15), we can write

$$K_+^{(B)}(2,1)=(i\nabla-\boldsymbol{B}-m)^{-1}i\delta(2,1), \tag{53}$$

so that $-iK_+^{(B)}(2,1)$ is the space-time representation of the operator $(i\nabla-\boldsymbol{B}-m)^{-1}$ needed for Eq. (52). Thus, the solution (52) is

$$\psi(2)=\int K_+^{(B)}(2,1)F(1)d\tau_1. \tag{54}$$

The perturbation theory, considering $\boldsymbol{B}$ as a perturbation on the free particle, arises from Eq. (53) from a power series expansion in $\boldsymbol{B}$. For any pair of operators A, B we have

$$(A+B)^{-1}=A^{-1}-A^{-1}BA^{-1}+A^{-1}BA^{-1}BA^{-1}\cdots, \tag{55}$$

so that with $A=(i\nabla-m)$, $B=-\boldsymbol{B}$ we have

$$(i\nabla-\boldsymbol{B}-m)^{-1}=(i\nabla-m)^{-1}+(i\nabla-m)^{-1}\boldsymbol{B}(i\nabla-m)^{-1}$$
$$+(i\nabla-m)^{-1}\boldsymbol{B}(i\nabla-m)^{-1}\boldsymbol{B}(i\nabla-m)^{-1}+\cdots \tag{56}$$

or in space representation (putting

$$K_+(2,1)=i(i\nabla-m)^{-1}\delta(2,1))$$

$$K_+^{(B)}(2,1)=K_+(2,1)-i\int K_+(2,3)\boldsymbol{B}(3)K_+(3,1)d\tau_3$$

$$-\int\int K_+(2,4)\boldsymbol{B}(4)K_+(4,3)\boldsymbol{B}(3)K_+(3,1)d\tau_4 d\tau_3, \tag{57}$$

as in I, Eqs. (13) and (14). The corresponding momentum representation is evident directly from Eq. (56), for $(i\nabla-m)^{-1}$ is $(p-m)^{-1}$.

If F is to represent an initial state, it is also convenient to use the free particle solution $f(1)=(i\nabla-m)^{-1}iF(1)$ to represent the state. We are often interested in the amplitude that the system is in a final state $g(\mathbf{x})$. In this case, we can define a sink function G and a corresponding free particle solution $\bar{g}=i\bar{G}(i\nabla-m)^{-1}$ (where we write the adjoint so that it will correspond to the solution $g(1)=-i(i\nabla-m')^{-1}G(1)$ corresponding to m' having the opposite sign of the imaginary part to m). The matrix element to go from f to g is then the space-time integral of

$$i\bar{G}(i\nabla-\boldsymbol{B}-m)^{-1}F. \tag{58}$$

The expansion (56) gives for this element

$$-i\bar{g}\boldsymbol{B}f-i\bar{g}\boldsymbol{B}(i\nabla-m)^{-1}\boldsymbol{B}f$$
$$-i\bar{g}\boldsymbol{B}(i\nabla-m)^{-1}\boldsymbol{B}(i\nabla-m)^{-1}\boldsymbol{B}f-\cdots \tag{59}$$

(assuming $\bar{g}$, f are orthogonal states, $\int\bar{g}(1)F(1)d\tau_1=0$ so the leading term vanishes). In space representation the first two terms of this are I, Eq. (22), and I, Eq. (23); in momentum representation the second term is I, Eq. (35).

If more than one real charge is present without interaction, there is an operator $(i\nabla-\boldsymbol{B}-m)^{-1}$ for each charge, operating exclusively on the space and spinor coordinates of that charge. Operators corresponding to distinct charges commute. Matrix elements are taken in the antisymmetric way described in I, Sec. 4, for accord with the exclusion principle.

The contribution from closed loops is a factor $C_v=\exp(-L)$, where L is not very easily defined directly in operators. But the first-order change on changing the potential from $\boldsymbol{B}$ to $\boldsymbol{B}+\Delta\boldsymbol{B}$ is

$$\Delta L=\mathrm{trace}[\{(i\nabla-\boldsymbol{B}-m)^{-1}-(i\nabla-m)^{-1}\}\Delta\boldsymbol{B}], \tag{60}$$

where the "trace" means the diagonal integral in coordinates and the "Sp" on the spinor indices, in space-time representation just I, Eq. (29).

This completes our summary in terms of operators of the results given in paper I on the theory of positrons. The main point is that aside from the problems of closed loops, one is merely analyzing by various techniques the consequences of Eq. (52) and, therefore, in general, the properties of the operator

$$(i\nabla-\boldsymbol{B}-m)^{-1}. \tag{61}$$

We may now turn to the quantum electrodynamics of such a particle, or system of particles. For simplicity we may restrict ourselves to the case of all virtual photons. The real photon case can, of course, as always be obtained by considering also the effects of external potentials. For simplicity further assume, at first, zero external potential. Our central problem, then, is the calculation of the matrix element

$$R={}_0\langle(i\nabla-\boldsymbol{A}-m)^{-1}\rangle_0 \tag{62}$$

of $(i\nabla-\boldsymbol{A}-m)^{-1}$ between states of the field empty of photons initially and finally. Here $\boldsymbol{A}=\gamma_\mu A_\mu$ and $A_\mu(1)$ is the operator $A_\mu^+(1)+A_\mu^-(1)$ acting on the field coordinates and satisfying commutation rules (46). This problem is relatively hard to solve directly. We do have the matrix element of any exponential form in A_μ in (43); but with A_μ involved in a reciprocal, it is another matter. We shall represent the reciprocal as a superposition of exponentials in the next section.

From Eq. (62) we can derive in a simple direct manner the perturbation series results of II. For we know[18] that

$${}_0\langle A_\mu(1)\rangle_0=0,\quad {}_0\langle A_\mu(1)A_\nu(2)\rangle_0=ie^2\delta_{\mu\nu}\delta_+(s_{12}^2),\quad \text{etc.} \tag{63}$$

[17] For $t<t_0$, the ψ from Eq. (52) would be zero and would not be the solution desired; it can be obtained only from F with a different definition of the poles of the reciprocal operator. We assume we are only interested in the solutions in regions of space-time later than the time the "initial" electron wave functions are specified and earlier than the "initial" positron function is given. Since we will take matrix elements between two states, this represents no real limitation (see I).

[18] The order of the A_μ operators in Eq. (63) is according to the time convention. If we put $A=A^+-A^-$ and use Eq. (46) to rearrange factors, they are evident, since A_μ^+ on the initial (and A_μ^- into the final) photon-free state vanishes. They are identical to the lagrangian relations III, Eq. (52), and form the basis of Dyson's description.

Hence, if we expand the reciprocal in power series as Eq. (56), in coordinate representation (57), with $\boldsymbol{B}$ replaced by $\boldsymbol{A}$ (or by $\boldsymbol{A}+\boldsymbol{B}$ if an unquantized potential $\boldsymbol{B}$ is present along with $\boldsymbol{A}$), we may readily write down the zero-zero matrix element of each term. For example, third term of Eq. (57) gives the self-energy contribution in accord with II, Eq. (6). The other results of II are just further consequences of this series expansion as we have emphasized in III.

8. USE OF A FIFTH PARAMETER IN DIRAC'S EQUATION

In this section we discuss the representation of the reciprocal operator in exponential form. Since $\int_0^\infty \exp(iWx)dW = i/x$ (or rather $\lim_{\epsilon\to 0} i/(x+i\epsilon)$), we we may write

$$i(i\nabla - \boldsymbol{B} - m)^{-1} = \int_0^\infty \exp[i(i\nabla - \boldsymbol{B} - m)W]dW, \quad (64)$$

the definition of the singularity (as the limit with m having an infinitesimal negative imaginary part) being automatically represented. (See, however, the remarks at the end of this section.) We can also write this in the ordered operator form,

$$\int_0^\infty \exp\left[i\int_0^W \{i\nabla(w) - \boldsymbol{B}(w)\}dw\right]\exp(-imW)dW, \quad (65)$$

where we have written $\boldsymbol{B}(w)$ for $\gamma_\mu(w)B_\mu(x_\nu(w))$, where $x_\nu(w)$ are the four ($\nu=1$ to 4) coordinate operators, of which B_μ is a function, ordered by the ordering parameter w and $\gamma_\mu(w)$ are the four Dirac matrices similarly ordered. Likewise, we have $i\nabla(w) = \gamma_\mu(w) i\partial_\mu(w)$, where $i\partial_\mu(w)$ are the four ordered momentum operators conjugate to x_μ.

In this form $\boldsymbol{B}$ can be replaced by $\boldsymbol{A}$ and the expectation value for virtual transitions can be taken. This is done in the next section. We continue here with a more complete discussion of Eq. (65).

The perturbation expansion in $\boldsymbol{B}$ of Eq. (65) should lead, of course, to Eq. (56). For example, the term first order in $\boldsymbol{B}$ in Eq. (65) is evidently

$$-i\int_0^\infty \exp(imW) \times \exp\left[i\int_0^W i\nabla(w')dw'\right]\int_0^W \boldsymbol{B}(w)dwdW. \quad (66)$$

Now the range of the w' may be divided into two regions and the quantities reordered (just as in Eqs. (5) and (6)) to

$$-i\int_0^\infty \int_0^W dWdw \exp[i(W-w)(i\nabla - m)]\boldsymbol{B} \times \exp[iw(i\nabla - m)],$$

so that changing the order of integration, one finds immediately from

$$\int_0^\infty \exp[iv(i\nabla - m)]dv = i(i\nabla - m)^{-1}$$

the result $-i(i\nabla - m)^{-1}\boldsymbol{B}(i\nabla - m)^{-1}$ as required.

The contribution L of a closed loop can be written directly in exponential form. It is easily shown from Eqs. (60) and (65) that

$$L = \int_0^\infty \text{trace}\left[\exp\left\{i\int_0^W (i\nabla(w) - \boldsymbol{B}(w))\right\}dw\right]\frac{dW}{W} \times \exp(-imW). \quad (67)$$

(The second term in Eq. (60) actually has zero trace and was added only to make convergence problems appear less difficult. It has been omitted in writing Eq. (67). Also, the value of Eq. (67) when $\boldsymbol{B}=0$ may be subtracted away, if desired, for a constant addition on L changes only the normalization of all probabilties.)

Incidentally, the method of rendering this expression convergent (see II, Sec. 7) for further calculations, is to call its value for mass m, $L(m^2)$ and then to calculate

$$L^P = \int_0^\infty [L(m^2) - L(m^2+\lambda^2)]G(\lambda)d\lambda,$$

where

$$\int_0^\infty G(\lambda)d\lambda = 1 \quad \text{and} \quad \int_0^\infty \lambda^2 G(\lambda)d\lambda = 0$$

and to assume that L^P is to be used as the correct value of L in place of $L(m^2)$. This is equivalent to replacing the factor $\exp(-imW)$ in the integrand of Eq. (67) by another function $F(W)$, where

$$F(W) = \int_0^\infty [\exp(-imW) - \exp\{-i(m^2+\lambda^2)^{\frac{1}{2}}W\}] \times G(\lambda)d\lambda.$$

For large W this approaches $\exp(-imW)$, but for small W it falls off, the real part of it, at least, varying as W^4. This renders Eq. (67) convergent. (The imaginary part of $F(W)$ does not seem to lead to momentum space integrals whose convergence would be in question.) This suggests a general method of maintaining convergence; by keeping processes corresponding to small intervals of w from occurring with large amplitude. This is briefly discussed in III in reference 22. What is said there applies qualitatively as well to the Dirac case analyzed here, with u replaced by w.

If there are several charges in the system, we must associate a separate w, for each, say w_n for the nth. Each must have its own set of matrices $\gamma_\mu^{(n)}$ and coordinates $x^{(n)}$ (γ's for different charges commute). If we call

$$\boldsymbol{B}^{(n)} = \gamma_\mu^{(n)}(w_n)B_\mu(x^{(n)}),$$

the total matrix for all N particles is

$$\prod_{n=1}^{N}\int_0^\infty \exp\left\{i\int_0^{W_n}[i\nabla^{(n)}(w_n)-\boldsymbol{B}^{(n)}(w_n)]dw_n\right\}$$
$$\times\exp(-imW_n)dW_n. \quad (68)$$

If in addition there are present a number of closed loops, the corresponding number of factors L must be multiplied in.

One might try to give a kind of physical or, rather, mathematical view by which the form of Eq. (65) can be appreciated, in the following manner:

We may deal with the Dirac equation somewhat in analogy to the method used in the discussion of the Klein-Gordon equation in the Appendix A of III. Consider a fifth variable w in addition to the four x, and that we have a wave function $\phi(x, w)$, which is to satisfy

$$-i\partial\phi/\partial w=(i\nabla-\boldsymbol{B})\phi. \quad (69)$$

Then since the potentials $B_\mu(x)$ are independent of w, the equation is separable in w, so that $\phi(x, w)=\exp(imw)\psi(x)$ is a solution of Eq. (69), if $\psi(x)$ is a solution of the Dirac Eq. (50). Also, if we have any special solution of Eq. (69), $\phi(x, w)$, we may obtain a solution of Eq. (50) by finding

$$\psi(x)=\int_{-\infty}^{\infty}\phi(x, w)\exp(-imw)dw. \quad (70)$$

Hence, by studying Eq. (69), we are at the same time studying the Dirac equation.

Given the wave function $\phi(x, 0)$ for $w=0$, the wave function at $w=W$ is given by

$$\phi(x, W)=\theta(W)\phi(x, 0), \quad (71)$$

where the operator $\theta(W)$ is

$$\theta(W)=\exp i\int_0^W[i\nabla(w)-\boldsymbol{B}(w)]dw \quad (72)$$

for $W>0$ and, for convenience,[17] we take $\theta=0$ for $W<0$. The important operator for the Dirac equation, in view of Eqs. (70) and (71), is

$$\int_{-\infty}^{\infty}\theta(W)\exp(-imW)dW,$$

which is just Eq. (65)

This interpretation suffers from a difficulty, however. For a free particle the operator $\theta(W)$ in momentum space is $\theta(W)=\exp(iW\not{p})=\cos(Wp)+i(\not{p}/p)\sin(pW)$, where $p=(p^2)^{\frac{1}{2}}$. The integral of this times $\exp(-imW)$ is really not always defined, even if m has a small negative imaginary part, for in intermediate states p^2 may be negative and p imaginary, so that θ contains positive exponentials in W and the integrand is oscillating with ever increasing amplitude. We therefore look at Eq. (64) as a *formal* definition of the value of the integral in all cases. Although this is satisfactory in a formal way for operators, it means that our interpretation cannot be taken literally. For example, we cannot obtain an unambiguous integral representation of $\theta(W)$ in coordinate space, for the requisite integral $\int\exp(-ip W)\exp(-ip\cdot x)d^4p$ is undefined. This is because it is probably not possible to obtain the wave function (71) at any value of W from that at $W=0$ from Eq. (69) without further definitions. At least, the corresponding second-order equation $(\partial^2\phi/\partial t^2)-\nabla^2\phi-\partial^2\phi/\partial W^2=0$ is apparently not of the kind for which this type of Huygens principle applies.

An alternative method of parametrizing the equation which does not seem to suffer from this interpretational difficulty is given in Appendix D. It leads, however, to more complicated (although algebraically equivalent) expressions for matrix elements than does Eq. (64).

9. DIRAC ELECTRONS IN QUANTUM ELECTRODYNAMICS

Returning now to quantum electrodynamics, for a single charge we want the expectation between photon free states of R in Eq. (62). This by Eq. (65) is the integral over all positive W of $\exp(-imW)$ times

$$\left\langle \exp i\int_0^W[i\nabla(w)-\boldsymbol{A}(w)]dw\right\rangle_0.$$

This is just $\exp[i\int_0^W i\nabla(w)dw]$ times (the $_0\langle\ \rangle_0$ refers to the photon states, that is, affects A_μ only)

$$\left\langle \exp -i\int_0^W\gamma_\mu(w)A_\mu(x(w))dw\right\rangle_0,$$

which is of the form

$$\left\langle \exp -i\int A_\mu(1)j_\mu(1)d\tau_1\right\rangle_0$$

of Eq. (42) with

$$j_\mu(1)=\int_0^W\gamma_\mu(w)\delta^4(x_\nu(w)-x_{\nu 1})dw, \quad (73)$$

where $x_{\nu 1}$ is the field point at which j_μ is calculated and $\delta^4(x_{\nu 2}-x_{\nu 1})$ means $\delta(2, 1)$. Thus, we may find the expectation value with the relation (43). With this value of j_μ substituted on the right side of Eq. (43), the δ^4 functions are immediately integrable, and we find finally

$$R=\int_0^\infty\exp\left[-\int_0^W\nabla(w)dw\right]$$
$$\times\exp\left[-\tfrac{1}{2}ie^2\int_0^W\int_0^W\gamma_\mu(w')\gamma_\mu(w'')\right.$$
$$\left.\times\delta_+(s^2_{w'w''})dw'dw''\right]\exp(-imW)dW, \quad (74)$$

in which we have written $s^2_{w'w''}$ for $[x_\mu(w')-x_\mu(w'')]\times[x_\mu(w')-x_\mu(w'')]$.

This expression then contains a description of a Dirac electron interacting with itself. If an extra factor $\exp(-i\int_0^W B(w)dw)$ is included, it describes such an electron also in an external potential. The terms may be expanded in powers of B and e^2, and each term may then be simplified in the way we have described many times before, for example, in connection with Eq. (66).

When several charges are present, the result from Eq. (68) is the integral[19]

$$\int_0^\infty\int_0^\infty\cdots\int_0^\infty \exp\left\{i\sum_n\int_0^{W_n}[i\nabla^{(n)}(w_n)-m]dw_n\right\}$$
$$\times\exp\left[-\tfrac{1}{2}i\sum_n\sum_m e^2\int_0^{W_n}\int_0^{W_m}\gamma_\mu{}^{(n)}(w_n)\gamma_\mu{}^{(m)}(w_m)\right.$$
$$\left.\times\delta_+(s^2{}_{w_nw_m})dw_ndw_m\right]dW_1dW_2\cdots dW_m. \quad (75)$$

The contributions from closed loops may be obtained from this by choosing some value of n, say, $n=i$, to represent matrices applying to a loop, dividing under the integral sign by W_i, and taking the trace with respect to the variables i.

The various present-day meson theories of nuclear interaction may be set up in quite analogous ways. For example, a nucleon interacting with itself through the agency of neutral pseudoscalar mesons with pseudoscalar coupling is evidently described by Eq. (74); but with m replaced by the proton mass, and the interaction term altered by the replacement of e^2 by g^2, γ_μ by γ_5, and $\delta_+(s^2)$ by $4\pi I_+(s^2)$, the appropriate propagation function for mesons of mass μ (I_+ is defined in I, Eq. (32), but $m=\mu$). Charged mesons may be represented by the use of isotopic spin operators also ordered by w.

10. SUMMARY OF NUMERICAL FACTORS FOR TRANSITION PROBABILITIES

The exact values of the numerical factors appearing in the rules of II for computing transition probabilities are not clearly stated there, so we give a brief summary here.[20]

The probability of transition per second from an initial state of energy E to a final state of the same total energy (assumed to be in a continuum) is given by ($\hbar=c=1$),

$$\text{Prob. trans/sec}=2\pi N^{-1}|\mathfrak{M}|^2\rho(E),$$

where $\rho(E)$ is the density of final states per unit energy range at energy E and $|\mathfrak{M}|^2$ is the square of the matrix element taken between the initial and final state of the transition matrix $\mathfrak{M}$ appropriate to the problem. N is a normalizing constant. For bound states conventionally normalized it is 1. For free particle states it is a product of a factor N_i for each particle in the initial and for each in the final energy state. N_i depends on the normalization of the wave functions of the particles (photons are considered as particles) which is used in computing the matrix element of $\mathfrak{M}$. The simplest rule (which does not destroy the apparent covariance of $\mathfrak{M}$), is[21] $N_i=2\epsilon_i$, where ϵ_i is the energy of the particle. This corresponds to choosing in momentum space, plane waves for photons of unit vector potential, $e^2=-1$. For electrons it corresponds to using $(\bar{u}u)=2m$ (so that, for example, if an electron is deviated from initial p_1 to final p_2, the sum over all initial and final spin states of $|\mathfrak{M}|^2$ is $\mathrm{Sp}[(p_2+m)\mathfrak{M}(p_1+m)\overline{\mathfrak{M}}]$). Choice of normalization $(\bar{u}\gamma_t u)=1$ results in $N_i=1$ for electrons. The matrix $\mathfrak{M}$ is evaluated by making the diagrams and following the rules of II, but with the following definition of numerical factors. (We give them here for the special case that the initial, final, and intermediate states consist of free particles. The momentum space representation is then most convenient.)

First, write down the matrix directly without numerical factors. Thus, electron propagation factor is $(p-m)^{-1}$, virtual photon factor is k^{-2} with couplings $\gamma_\mu\cdots\gamma_\mu$. A real photon of polarization vector e_μ contributes factor e. A potential (times the electron charge, e) $A_\mu(x)$ contributes momentum q with amplitude $a(q)$, where $a_\mu(q)=\int A_\mu(1)\exp(iq\cdot x_1)d^4x_1$. (Note: On this point we deviate from the definition of a in I which is there $(2\pi)^{-2}$ times as large.) A spur is taken on the matrices of a closed loop. Because of the Pauli principle the sign is altered on contributions corresponding to an exchange of electron identity, and for each closed loop. One multiplies by $(2\pi)^{-4}d^4p=(2\pi)^{-4}dp_tdp_xdp_ydp_z$ and integrates over all values of any undetermined momentum variable p. (Note: On this point we again differ.[20])

The correct numerical value of $\mathfrak{M}$ is then obtained by multiplication by the following factors. (1) A factor $(4\pi)^{\frac{1}{2}}e$ for each coupling of an electron to a photon. Thus, a virtual photon, having two such couplings, contributes $4\pi e^2$. (In the units here, $e^2=1/137$ approximately and $(4\pi)^{\frac{1}{2}}e$ is just the charge on an electron in heaviside units.) (2) A further factor $-i$ for each virtual photon.

For meson theories the changes discussed in II, Sec. 10 are made in writing $\mathfrak{M}$, then further factors are

[19] This equation with its interpretation was proposed as a formulation of the laws of quantum electrodynamics (for virtual photons) by the author at the Pocono Conference of Theoretical Physics (1948). The notation for ordering operators was explained there. However, at this time, the author had no complete formal derivation of Eq. (75) from the conventional electrodynamics, nor did he know of a satisfactory method of dealing with the closed loop divergences.

[20] In I and II the unfortunate convention was made that d^4k means $dk_tdk_1dk_2dk_3(2\pi)^{-2}$ for momentum space integrals. The confusing factor $(2\pi)^{-2}$ here serves no useful purpose, so the convention will be abandoned. In this section d^4k has its usual meaning, $dk_tdk_1dk_2dk_3$.

[21] In general, N_i is the particle density. It is $N_i=(\bar{u}\gamma_t u)$ for spin one-half fields and $i[(\phi^*\partial\phi/\partial t)-\phi\partial\phi^*/\partial t]$ for scalar fields. The latter is $2\epsilon_i$ if the field amplitude ϕ is taken as unity.

(1) $(4\pi)^{\frac{1}{2}}g$ for each meson-nucleon coupling and (2) a factor $-i$ for each virtual spin one meson, but $+i$ for each virtual spin zero meson.

This suffices for transition probabilities, in which only the absolute square of $\mathfrak{M}$ is required. To get $\mathfrak{M}$ to be the actual phase shift per unit volume and time, additional factors of i for each virtual electron propagation, and $-i$ for each potential or photon interaction, are necessary. Then, for energy perturbation problems the energy shift is the expected value of $i\mathfrak{M}$ for the unperturbed state in question divided by the normalization constant N_i belonging to each particle comprising the unperturbed state.

The author has profited from discussions with M. Peshkin and L. Brown.

APPENDIX

In this Appendix (A, B, C) an attempt will be made to discuss some of the properties of ordered operators and of functionals in a somewhat more general way.

Almost certainly many of the equations will be incorrect in their general form. This is especially true of those involving fourier transforms in function space. However, it is expected that they are correct in the special cases in which the formulas have been applied in the main part of the paper. Therefore, at least at first, when new results using these methods are derived, care should be taken to check the final result in some independent way. It is analogous to using power series expansions, or fourier transforms, in a calculation in a situation in which the conditions for the validity of the power expansions or of the transform have not been checked, or are not known to be satisfied. The physicist is very familiar with such a situation and usually satisfied with it, especially since he is confident that he can tell if the answer is physically reasonable. But mathematicians may be completely repelled by the liberties taken here. The liberties are taken not because the mathematical problems are considered unimportant. On the contrary, this appendix is written to encourage the study of these forms from a mathematical standpoint. In the meantime, just as a poet often has license from the rules of grammar and pronunciation, we should like to ask for "physicists' license" from the rules of mathematics in order to express what we wish to say in as simple a manner as possible. (These remarks do not apply to Appendix D.)

A. Relation to Theory of Functionals

In this section we would like to suggest how a general theory of ordered operators might be built up, and in particular, to point out certain relations to the theory of functionals. For clarity of exposition in this Sec. A, only, we represent all operators by bold-faced letters $\mathbf{M}$ and ordinary functions in regular type M. We have mentioned that with every functional $F[M(s), N(s)\cdots]$ of the argument functions $M(s)$, $N(s)$ we wish to associate an operator (by identifying $M(s)$ with an operator $\mathbf{M}(s)$ interpreting s as an ordering parameter with the operators $\mathbf{M}(s)$, $\mathbf{N}(s)$ supposedly known and with known commutation relations). The general theory of these associations might instead have begun by defining the meaning for the special case of the exponential functional $\exp\int_0^1 M(s)ds$ (we assume throughout this section, for convenience, that the range of s is 0 to 1). The corresponding operator

$$\mathbf{R}=\exp\int_0^1 \mathbf{M}(s)ds \tag{1-a}$$

is defined as the value $\mathbf{G}(1)$ at $s=1$ of that solution of the operator differential equation

$$d\mathbf{G}(s)/ds=\mathbf{M}(s)\mathbf{G}(s), \tag{2-a}$$

which is the identity operator at $s=0$, i.e., $\mathbf{G}(0)=\mathbf{I}$. We have thereby defined the operators corresponding to more complex functionals such as $F=\exp\int_0^1[\mu(s)M(s)+\nu(s)N(s)+\cdots]ds$, where $\mu(s)$, $\nu(s)\cdots$ are numerical functions and $\mathbf{M}$, $\mathbf{N}$ arbitrary operators (which need not commute) as the $\mathbf{G}(1)$ from

$$d\mathbf{G}(s)/ds=[\mu(s)\mathbf{M}(s)+\nu(s)\mathbf{N}(s)+\cdots]\mathbf{G}(s) \tag{3-a}$$

with $\mathbf{G}(0)=\mathbf{I}$. For clearly $\mu(s)\mathbf{M}(s)+\nu(s)\mathbf{N}(s)+\cdots$ can be considered as a single operator function of s, the $\mathbf{M}(s)$ in Eqs. (1-a) and (2-a).

Next we make the general definition that the operator to be associated with the sum of two or more functions $F_1[M(s), N(s)\cdots]+F_2[M(s), N(s)\cdots]$ is the sum of the operators $F_1[\mathbf{M}(s), \mathbf{N}(s)\cdots]+F_2[\mathbf{M}(s), \mathbf{N}(s)\cdots]$ corresponding to each separately.

Considering a derivative as the limit of a difference, we can use this idea of superposition to further extend the range of functionals for which operators are defined. As an example, in virtue of the fact that $\int_0^1 M(s)ds\int_0^1 N(s)ds$ is the first derivative with respect to both μ, ν of $\exp\int_0^1[\mu M(s)+\nu N(s)]ds$ evaluated at $\mu, \nu=0$ we may define the operator corresponding to $\int_0^1 \mathbf{M}(s)ds\int_0^1 \mathbf{N}(s)ds$ as the corresponding derivative of the operator $\exp\int_0^1[\mu\mathbf{M}(s)+\nu\mathbf{N}(s)]ds$. Then from a study of the properties of the solution of Eq. (3-a) expanded in powers of μ, ν we may readily verify that $\int_0^1 \mathbf{M}(s)ds\int_0^1 \mathbf{N}(s)ds$ could also be evaluated directly by considering s as an ordering index on the operators.

Thus, the superposition rule permits a wide increase in the class of functionals for which we have defined operators. In fact, with some mathematical license, we have defined the operator for any functional. We wish to imagine that any functional can be represented as a superposition of exponential ones in a manner analogous to the representation of an arbitrary function as a superposition of exponential functions. Thus, we expect to be able to write for any functional $F[M(s)]$ (the true mathematical restrictions are completely unknown to me)

$$F[M(s)]=\int \exp\left[i\int_0^1 \mu(s)M(s)ds\right]\mathfrak{F}[\mu(s)]\mathfrak{D}\mu(s), \tag{4-a}$$

where $\mathfrak{F}[\mu(s)]$ is a new (complex) functional, the functional transform of $F[M(s)]$, and $\int\cdots\mathfrak{D}\mu(s)$ represents (some kind of an) integration over the space of functions $\mu(s)$. For simplicity we take the case of just one argument function $M(s)$. If $F[M(s)]$ is given, $\mathfrak{F}$ can be determined perhaps from

$$\mathfrak{F}[\mu(s)]=\int \exp\left[-i\int_0^1 \mu(s)M(s)ds\right]F[M(s)]\mathfrak{D}M(s) \tag{5-a}$$

with suitable normalization. Then, if $\mathfrak{F}$ is known, we define the operator $F[\mathbf{M}(s)]$ as

$$F[\mathbf{M}(s)]=\int \exp\left[i\int_0^1 \mu(s)\mathbf{M}(s)ds\right]\mathfrak{F}[\mu(s)]\mathfrak{D}\mu(s), \tag{6-a}$$

where $\mu(s)$ is a numerical function. Since we have already defined the operator $\exp[i\int_0^1 \mu(s)\mathbf{M}(s)ds]$ (by Eq. (3-a) with μ replaced by $i\mu$), we now simply require superposition of such operators for various $\mu(s)$. The extension to functionals of several variables is evident.

With these definitions of operators in terms of exponential functionals, the various theorems are easily proved. For example, the theorem (18) of Sec. 3 is first readily demonstrated for the special case that F is an exponential (1-a). Thus, to calculate

$$\mathbf{R}=\exp\int_0^1 \mathbf{P}(s)ds\exp\int_0^1 \mathbf{M}(s)ds$$

we must solve $d\mathbf{G}(s)/ds=[\mathbf{P}(s)+\mathbf{M}(s)]\mathbf{G}(s)$. We try a solution $\mathbf{G}(s)=\mathbf{U}(s)\mathbf{X}(s)$, so that $d\mathbf{G}/ds=(d\mathbf{U}/ds)\mathbf{X}+\mathbf{U}d\mathbf{X}/ds=\mathbf{PG}+\mathbf{U}d\mathbf{X}/ds$ in virtue of Eq. (15). Thus, we have a solution if $d\mathbf{X}/ds=\mathbf{U}^{-1}\mathbf{MG}=\mathbf{U}^{-1}\mathbf{MUX}=\mathbf{M}'\mathbf{X}$ with $\mathbf{M}'$ as in Eq. (16). Since $\mathbf{G}(0)=1$, if $\mathbf{U}(0)=1$, we must have $\mathbf{X}(0)=1$ so the solution of the $\mathbf{X}$ equation is $\exp\int_0^1 \mathbf{M}'(s)ds$ in accordance with the definition (1-a), (2-a). (If $\mathbf{U}(0)\neq 1$, replace $\mathbf{X}$ by $\mathbf{XU}^{-1}(0)$ throughout.) Hence, Eq. (18) is established for exponential functionals. And since the theorem involves F linearly, it is therefore true for any

superposition of exponentials and hence for any functional which can be defined by means of such superposition.

B. Momentum and Coordinate Operators

In nonrelativistic quantum mechanics, without spin, all operators can be made up of coordinate operators and their conjugate momentum operators. We show in this section how, at least in principle, all such operator functions can be disentangled.

We can consider the case of one degree of freedom Q, and its momentum P. (When more variables are present, they present no new problem as variables corresponding to different independent coordinates commute.) Thus, we are to disentangle the general operator $F[P(s), Q(s)]$ subject to the condition

$$PQ-QP=-i. \tag{7-a}$$

This is the problem solved in this section. We can satisfy the commutation relation by putting $P=-id/dQ$ (so that our solution may have applications outside quantum mechanics for the combination of operators X, d/dX are of frequent occurrence). Then the operator F can be defined by giving the function g of Q resulting from

$$g(Q)=F[P(s), Q(s)]f(Q) \tag{8-a}$$

for arbitrary functions $f(Q)$, where $Q(s)$ and $P(s)$ are interpreted as multiplication by Q, and $-i$ times differentiation with respect to Q in the order defined by F.

To obtain the relation of g, f suppose the P-dependence of F can be expanded as a functional transform, ($p(s)$, $q(s)$, $v(s)$ are numerical functions)

$$F[p(s), q(s)]=\int \exp\left[-i\int_0^1 p(s)v(s)ds\right]\mathfrak{F}[q(s), v(s)]\mathfrak{D}v(s), \tag{9-a}$$

where $\mathfrak{F}$ is a functional of $v(s)$ and of $q(s)$. Now to evaluate the operator

$$\exp\left[-i\int_0^1 P(s)v(s)ds\right]\mathfrak{F}[Q(s), v(s)] \tag{10-a}$$

we use our theorem (20) to disentangle the $P(s)$ operator. We use $\alpha(s)=-iv(s)$ in Eq. (20), calling $y(s)=\int_0^s v(s')ds'$, so that Eq. (10-a) is

$$e^{-iy(1)P(1)}\mathfrak{F}[Q'(s), v(s)], \tag{11-a}$$

where $Q'(s)=e^{+iy(s)P}Q_s e^{-iy(s)P}$. As is well known from Taylor's theorem, the operator $e^{hd/dx}$ displaces x by h so that[22]

$$Q'(s)=Q_s+y(s). \tag{12-a}$$

Substitution into Eq. (11-a) finds all the Q_s preceding the $P(1)$ so the operators are disentangled and Q_s may be written simply Q_0, whence we have

$$F[P(s), Q(s)]=\int \exp\left[-iP_1\int_0^1 v(s)ds\right]\times\mathfrak{F}\left[Q_0+\int_0^s v(s')ds', v(s)\right]\mathfrak{D}v(s), \tag{13-a}$$

which in principle, at least, solves the problem.

We can go a bit further and assume Eq. (9-a) can be inverted as

$$\mathfrak{F}[q(s), v(s)]=\int \exp\left[i\int_0^1 p(s)v(s)ds\right]F[p(s), q(s)]\mathfrak{D}p(s), \tag{14-a}$$

where $p(s)$ is a numerical function for transforming F. Also, $y(s)=\int_0^s v(s)ds$ is as good a function as $v(s)$ for purposes of integration,[23] and we may write, substituting Eq. (14-a) into Eq. (13-a),

$$F[P(s), Q(s)]=\int\int \exp\left[-iP_1y(1)+i\int_0^1 p(s)\dot{y}(s)ds\right]\times F[p(s), Q_0+y(s)]\mathfrak{D}p(s)\mathfrak{D}y(s), \tag{15-a}$$

the integral extending over all $p(s)$, and all $y(s)$ subject to $y(0)=0$.

Considering P as $-id/dQ$, the operator $F[P(s), Q(s)]$ may be considered to operate on a function $f(Q)$ to produce another function of Q. In particular, we are often interested in quantum mechanics in the projection of this final function into a given "final state" function g; that is, F is often defined through its matrix element

$$(g^*Ff)=\int g^*(Q)F[P(s), Q(s)]f(Q)dQ.$$

If we substitute into this expression (15-a), the P_1 can be considered to act entirely on $g^*(Q)$ and since $\exp(+iyP)g(Q)=g(Q+y)$, we find

$$(g^*Ff)=\int g^*(Q_0+y(1))\exp\left[i\int p(s)\dot{y}(s)ds\right]F[p(s), Q_0+y(s)]\times\mathfrak{D}p(s)\mathfrak{D}y(s)\cdot f(Q_0)dQ_0.$$

Define $q(s)$ as the numerical function $q(s)=Q_0+y(s)$ and write finally ($q_0=Q_0$)

$$(g^*Ff)=\int\int g^*(q_1)\exp\left[i\int p(s)\dot{q}(s)ds\right]F[p(s), q(s)]\times\mathfrak{D}p(s)\mathfrak{D}q(s)f(q_0)dq_0dq_1, \tag{16-a}$$

where the integral $\mathfrak{D}p(s)$ is over all $p(s)$ and the integral $\mathfrak{D}q(s)$ is over all trajectories $q(s)$ which go between the initial position q_0 and the final one q_1, the final integration on dq_1 being represented explicitly. This represents a complete reduction of an ordered operator $F[P(s), Q(s)]$ involving conjugate operators P, Q (7-a)to a property of the corresponding numerical functional $F[p(s), q(s)]$, for in Eq. (16-a) $p(s)$, $q(s)$ are numerical functions so that all the operators have been eliminated.

This is obviously related to the lagrangian form of quantum mechanics of C. In fact, for transitions, we are interested in the operator $S=\exp[-i\int_0^T H(t)dt]$, where, for example, $H=(1/2m)P^2+V(Q, t)$. The matrix elements of this, according to Eq. (16-a) are (use t for s in range 0 to T)

$$(g^*Sf)=\int\int g^*(q_T)\exp\left[i\int_0^T p(t)\dot{q}(t)dt-i\int_0^T (1/2m)p(t)^2dt-i\int_0^T V(q(t), t)dt\right]f(q_0)\mathfrak{D}p(t)\mathfrak{D}q(t)dq_0dq_T.$$

The integral on $p(t)$ is easily done. (See Appendix C for a more general discussion of gaussian integrals.) Substitute $p(t)=m\dot{q}+p'(t)$, so that

$$\int_0^T [(1/2m)p(t)^2-p(t)\dot{q}(t)]dt=(1/2m)\int_0^T p'(t)^2dt-\tfrac{1}{2}m\int_0^T \dot{q}(t)^2dt.$$

Also, $\mathfrak{D}p(t)=\mathfrak{D}p'(t)$, since p and p' differ by a constant at each t (keeping $q(t)$ integration until later). The $\mathfrak{D}p'(t)$ integral then separates out and integrates to some constant. Hence, within such a normalizing constant, the matrix element is

$$(g^*Sf)=\int g^*(q_T)\exp\left\{+i\int_0^T [\tfrac{1}{2}m\dot{q}(t)^2-V(q(t), t)]dt\right\}\times f(q_0)\mathfrak{D}q(t)dq_0dq_T. \tag{17-a}$$

That is, the transition amplitude from point q_0 at $t=0$ to q_T at $t=T$ is the integral over all trajectories connecting these points of $\exp i\int_0^T L[\dot{q}(t), q(t)]dt$, L being the lagrangian for this problem. This is the fundamental theorem on which the interpretation of C is based.

The fact that the nonrelativistic quantum mechanical operators (other than spin) can be expressed in terms of an integral over

[22] Or, differentiating $Q'(s)$ with respect to s, find $dQ'(s)/ds=ie^{iy(s)P}(PQ-QP)e^{-iy(s)P}dy/ds$. If we use the commutation relation, this is dy/ds, whence $Q'(s)$ differs from $y(s)$ by the constant operator Q, the evident value of Q' for $y(s)=0$, establishing Eq. (12-a).

[23] For, if $\mathfrak{D}v(s)$ be considered as the limit as $\Delta\to 0$ of an integration over all the variables $v_i=v(s_i)$ with $s_{i+1}-s_i=\Delta$, then the change is from the variables $v_i=(y_{i+1}-y_i)\Delta^{-1}$. Integration over y_i for all $i>0$ is equivalent to integration on all v_i. (Since $dv_i=\Delta^{-1}dy_{i+1}$, the jacobian of the transformation is $\Delta^{-(1/\Delta)}$, which is only a change of normalization, and we are disregarding normalization factors.)

trajectories is based on the fact that the operators involved satisfy Eq. (7-a). If other operators are involved, such as Pauli's spin operators $\boldsymbol{\sigma}$, or Dirac matrices γ_μ, which satisfy different commutation rules, a complete reduction eliminating all the operators is not nearly so easily affected. It is possible to eliminate the p operators in the Dirac or Pauli equation and get forms like Eq. (17-a), but the amplitude for a single trajectory is then a hypercomplex quantity in the algebra of the γ_μ or $\boldsymbol{\sigma}$. We give an example of this.

Without disentangling the γ_μ operators we shall disentangle the momentum operators $p_\mu = i\partial/\partial x_\mu$ from the operator $\theta(W)$ of Eq. (72), which is a key operator in the analyses of the Dirac equation.

If we write

$$\theta(W) = \exp\left[i\int_0^W \gamma_\mu(w)p_\mu(w)dw\right]\exp\left[-i\int_0^W B\{x(w), w\}dw\right],$$

the p_μ operators are already in exponential form and no fourier transforms are necessary. We may disentangle the p_μ in the first integral by a direct use of the theorem (20) with $P(s) = p_\mu(w)$ (p_μ for each value μ is disentangled separately) and $\alpha(s) = i\gamma_\mu(w)$. The resulting $x_\mu'(w)$ operator is $x_\mu(w) + \int_0^w \gamma_\mu(w')dw'$ just as in Eq. (12-a) so that we obtain

$$\theta(W) = \exp\left[ip_\mu(W)\int_0^W \gamma_\mu(w)dw\right] \times \exp\left[-i\int_0^W B\left\{x_\mu(0) + \int_0^W \gamma_\mu(w')dw', w\right\}dw\right]. \quad (18\text{-a})$$

Here the x_μ and p_μ are completely separated, but the γ_μ are thoroughly entangled. The w in B keeps track of the fact that the γ_μ in its definition acts at the order w; thus, it is $\gamma_\mu(w)B_\mu[x_\mu(0) + \int_0^w \gamma_\mu(w')dw']$. A similar separation may be made in the operator for self-action (74) which now is

$$R = \int_0^\infty \exp[ip_{\mu,\infty}C_\mu(W)]\exp\left\{-\tfrac{1}{2}ie^2\int_0^W\int_0^W \dot{C}_\mu(w')\dot{C}_\mu(w'') \times \delta_+[\{C_\mu(w') - C_\mu(w'')\}^2]dw'dw''\right\}\exp(-imW)dW, \quad (19\text{-a})$$

where one must substitute $C_\mu(w) = \int_0^w \gamma_\mu(w')dw'$, $\dot{C}_\mu(w) = \gamma_\mu(w)$ ($p_{\mu,\infty}$ refers to the momentum operator operating on the final state). All reference to space coordinates have disappeared. The problem of self-energy of an electron is reduced to the algebraic one of disentangling a combination of γ_μ in an expression in which, however, they are almost hopelessly tangled up. Not much has been done with this expression. (It is suggestive that perhaps coordinates and the space-time they represent may in some future theory be replaced completely by an analysis of ordered quantities in some hypercomplex algebra).

Since the spin operators are so simple and fundamental to quantum mechanics, they present some interesting unsolved problems. For example, if $F[x(s), y(s), z(s)]$ is a known functional of a three-space trajectory $x(s)$, $y(s)$, $z(s)$, evaluate in terms of this functional, the operator $F[\sigma_x(s), \sigma_y(s), \sigma_z(s)]$, where the σ_x, σ_y, σ_z are the anticommuting Pauli operators of unit square satisfying $\sigma_x\sigma_y\sigma_z = i$. The corresponding problem with Dirac operators is a kind of four-dimensional generalization of this. Alternatively, since the Dirac operators can be represented as the outer product of two commuting sets of Pauli operators, the solution of the problem with Pauli operators could be directly extended to the Dirac case. The Pauli matrices (times i) are the basis for the algebra of quaternions so that the solution of such problems might open up the possibility of a true infinitesimal calculus of quantities in the field of hypercomplex numbers.

C. Gaussian Functionals

In a large number of problems the operators appear in exponentials only up to the second degree. For this reason it is handy to have available a formula for the integration of gaussian functionals. We can define a gaussian functional $G[y(s)]$, of one function $y(s)$, as one of the form, $G[y(s)] = \exp iE[y(s)]$ with $E[y(s)]$ quadratic. Thus, we have

$$E[y(s)] = \tfrac{1}{2}\int_0^1\int_0^1 A(t, s)y(t)y(s)dtds + \int_0^1 B(s)y(s)ds, \quad (20\text{-a})$$

where $A(t, s)$, $B(s)$ are functions independent of y (that is, G is gaussian if the second functional derivative of $\ln G$ is independent of y). Gaussian functionals of several variables are of frequent occurrence. All the quantum field theory hamiltonians and lagrangians are of this form in the field variables. A formula for the integral of $G[y(s)]$ over all paths $y(s)$ has been found useful. It will be developed here. Consider the integral (we suppose $A(t, s)$ real, or at least has a positive definite imaginary part)

$$I[A, B] = \int \exp\{iE[y(s)]\}\mathfrak{D}y(s). \quad (21\text{-a})$$

It is a functional of $A(t, s)$ and $B(s)$. First, the dependence on B may be determined, as follows.

Let $\bar{y}(s)$ be that trajectory which makes the exponent $E[y(s)]$ an extremum. That is, $\bar{y}$ is a solution of (assuming A symmetric)

$$\int_0^1 A(t, s)\bar{y}(s)ds = -B(t). \quad (22\text{-a})$$

Or, if N be the reciprocal kernel to A (which can often most easily be found merely by solving Eq. (22-a)),

$$\bar{y}(s) = -\int^1 N(t, s)B(t)dt. \quad (23\text{-a})$$

Then put $y(s) = \bar{y}(s) + x(s)$. (Note, $\mathfrak{D}y(s) = \mathfrak{D}x(s)$.) In virtue of Eq. (22-a), one finds $E[y] = E[\bar{y}] + \frac{1}{2}\int_0^1\int_0^1 A(t, s)x(t)x(s)dtds$. Here, $E[\bar{y}(s)]$ can also be written explicitly as $-\frac{1}{2}\int_0^1\int_0^1 B(s) \times N(t, s)B(t)dtds$, using Eq. (23-a). Substituting this into Eq. (21-a), we see a factor $\exp\{iE[\bar{y}(s)]\} = G[\bar{y}(s)]$ may be taken outside the integral, as it is independent of $x(s)$. Hence, we have

$$I[A, B] = G[\bar{y}(s)]J[A], \quad (24\text{-a})$$

where

$$J[A] = \exp\left[\tfrac{1}{2}i\int_0^1\int_0^1 A(t, s)x(t)x(s)dtds\right]\mathfrak{D}x(s) = I[A, 0] \quad (25\text{-a})$$

does not depend on B, and

$$G[\bar{y}(s)] = \exp\{iE[\bar{y}(s)]\} = \exp\left[-\tfrac{1}{2}i\int_0^1\int_0^1 B(s)N(t, s)B(t)dtds\right]. \quad (26\text{-a})$$

Often this is as far as it is necessary to go, as the dependence of I on B may have been all that is necessary to know, $J[A]$ being a kind of normalizing factor that is not of importance or that can be obtained in some other manner.

Having this form for I, we may obtain other integrals. For example, since

$$\delta I[A, B]/\delta B(t) = i\int G[y(s)]y(t)\mathfrak{D}y(t),$$

this integral can be immediately evaluated by differentiation of the expression (24-a) for I with respect to $B(t)$. Since

$$\delta G[\bar{y}(s)]/\delta B(t) = iG[\bar{y}(s)]\delta E(\bar{y}(s))/\delta B(t) = -i\int_0^1 B(s)N(t, s)ds \cdot G[\bar{y}(s)],$$

we find

$$\int G[y(s)]y(t)\mathfrak{D}y(s) = \bar{y}(t)I[A, B]. \quad (27\text{-a})$$

Differentiating a second time, since $\delta\bar{y}(t)/\delta B(t') = -N(t, t')$, one finds

$$\int G[y(s)]y(t)y(t')\mathfrak{D}y(s) = [\bar{y}(t)\bar{y}(t') + iN(t, t')]I[A, B], \quad (28\text{-a})$$

etc., for higher powers of y.

Incidentally, this permits us to obtain the properties of J. For the left-hand side of (28-a) is also $-2i\delta I[A, B]/\delta A(t, t')$. In the special case $B = 0$, we have $\bar{y} = 0$ from (23-a), and since $J(A) = I[A, 0]$, we find

$$\delta J/\delta A(t, t') = -\tfrac{1}{2}N(t, t')J. \quad (29\text{-a})$$

This property of J determines it to within a numerical factor independent of A.

We have used these theorems, or something like them, on various occasions. One example was the passage from Eq. (16-a) to Eq. (17-a). In more generality, put

$$F=\exp\left[-i\int H(p(s),q(s))ds\right],$$

where H is quadratic in p. The integrations on $p(s)$ in Eq. (16-a) now represent an example of our theorem with $y=p$ and $E[p(s)]=\int[p\dot{q}-H(p,q)]ds$. The extremum requires $\partial H/\partial p=\dot{q}$. If the solution of this is $\bar{p}$, considered as a function of $\dot{q}$, q, then the integral on p produces within unimportant factors an exponential of $\int[\bar{p}\dot{q}-H(\bar{p},q)]ds$, that is, $i\int Lds$, where L is the lagrangian. This example shows that in our discussion, we have not been sufficiently rigorous mathematically, for the important problem of the order of noncommuting operators p, q, in the original definition of H does not seem to have arisen.

A second example is the integration of $\exp(i\int Ldt)$ when L is quadratic in $\dot{q}$, q. For the forced harmonic oscillator where $L=\frac{1}{2}(\dot{q}^2-\omega^2q^2)+\gamma(t)q(t)$, the integral was carried out in III, footnote 7. The operator $A(t,s)$ is $-[(d^2/dt^2)+\omega^2]\delta(t-s)$, the inverse N of which involves sines and cosines but is not unique. However, in this case boundary conditions exist at the end points $q(0)$, $q(T)$, and these boundary conditions determine N and also restrict the range of y integration. The footnote serves a model of what to do under circumstances and will not be discussed further here.

The problem of integrating

$$\exp\left[-i\int j_\mu(1)A_\mu(1)d\tau_1\right]\cdot\exp\left[i\int(\partial A_\mu/\partial x_\nu)^2d\tau\right](8\pi e^2)^{-1}$$

over all distributions of field $A_\mu(1)$ required in III, Sec. VIII, serves as a further example. Here $y(s)$ is replaced by $A_\mu(1)$, and $B(s)$ by $j_\mu(1)$. The operator $A(t,s)$ becomes $\Box_1^2\delta(2,1)$, the inverse of which is again not unique. The inverse $N(t,s)$ required in III is $\delta_+(s_{12}^2)$. (The boundary conditions required to define this particular inverse are probably related to the condition that no photons are supplied in the past and none are wanted in the future, so that the inverse must have no positive frequencies for $t\to-\infty$ and no negative ones to $t\to+\infty$.) Thus, $E[\bar{y}(s)]$ becomes the important quantity $-\frac{1}{2}\int\int j_\mu(1)j_\mu(2)\delta_+(s_{12}^2)d\tau_1d\tau_2$, so that Eq. (24-a) gives Eq. (43) or III, Eq. (48), which we had taken such pains in III to derive in a more rigorous manner. In none of these examples do we require J.

A more complicated example is that of the analysis of the operators corresponding to the electron-positron field given in I, Appendix. If electrons obeyed Bose statistics, the commutation rules would have been altered, the net effect being just to change a few signs in the final expressions. Analyzed as an Einstein-Bose field, however, the operators Ψ can be considered as ordinary functions, and the lagrangian technique may be used. The problem then requires gaussian integrals (actually, integrals of exponentials of bilinear expressions, but these are as easy to work out). The y corresponds to Ψ (or Ψ^*), the $A(t,s)$ is related to the Dirac hamiltonian, and its inverse $K_+^{(A)}(2,1)$ replaces N. The problem of determining C_v corresponds to that of finding $J[A]$. The problem is complicated somewhat by the necessity of keeping the order of the γ_μ-operators correctly.

The relation of problems with operators obeying Fermi-Dirac statistics and those with the same operators obeying Einstein-Bose commutation rules is very close. The results of the former in practical cases,[24] at least, may be obtained from the latter by simply altering some signs. The Einstein-Bose case is very easily analyzed by ordered operator algebra (as in Sec. 7) or by the lagrangian integral methods. The anticommuting operators seem at first sight more complicated; but this they cannot be, as the results are just as simple. It would seem worth while to develop the analysis of anticommuting operators in much more detail than has been given here. Presumably, good use can be made of the similarity to the Einstein-Bose case. The theorems developed in analysis of this problem may conceivably have application in the problem of disentangling Pauli spin operators.

[24] The only known practical case, of course, is the electron-positron field. Here the problem has been completely worked out. I seem to be affected by the disease so prevalent today in theoretical physics, to delight in seeing a very general method of solving a problem, when actually in physics only one example of the type of problem exists and this has already been worked out.

D. Fock's Parameterization of the Dirac Equation

We wish to call attention to an interesting alternative method of parameterizing the Dirac equation, suggested by Fock.[4] It, like that of Sec. 8, would also have permitted us to pass directly to the formulation of electrodynamic problems. It is more readily interpreted than that of Sec. 8.[25]

As a consequence of the Dirac equation $(i\nabla-B)\psi=m\psi$, ψ also satisfies $(i\nabla-B)(i\nabla-B)\psi=m^2\psi$. Expanding the operator, this is equivalent to

$$[(i\partial/\partial x_\mu)-B_\mu]^2\psi-\tfrac{1}{2}\sigma_{\mu\nu}F_{\mu\nu}\psi=m^2\psi, \tag{30-a}$$

with $\sigma_{\mu\nu}=\frac{1}{2}i(\gamma_\mu\gamma_\nu-\gamma_\nu\gamma_\mu)$. This differs from the Klein-Gordon equation only through the addition of the term $-\frac{1}{2}\sigma_{\mu\nu}F_{\mu\nu}$, where $F_{\mu\nu}=(\partial B_\nu/\partial x_\mu)-\partial B_\mu/\partial x_\nu$ is the field tensor.

Just as in the Klein-Gordon case, III, Appendix A, this can be converted by the aid of a fifth parameter to Fock's equation,

$$i\partial\phi/\partial u=\tfrac{1}{2}[(i\partial/\partial x_\mu)-B_\mu]^2\phi-\tfrac{1}{4}\sigma_{\mu\nu}F_{\mu\nu}\phi, \tag{31-a}$$

for which the special solution $\phi=\exp(-\frac{1}{2}im^2u)\psi$ leads back to Eq. (30-a). It can then be analyzed by the lagrangian method. The final result is that the amplitude to go from one point to another [see III Eq. (5-a)] is the sum over all trajectories $x_\mu(u)$ of the hyper complex amplitude

$$\exp\left\{-i\int_0^{u_0}[\tfrac{1}{2}(dx_\mu/du)^2+(dx_\mu/du)B_\mu(x)-\tfrac{1}{4}\sigma_{\mu\nu}(u)F_{\mu\nu}(x(u))]du\right\}, \tag{32-a}$$

the order of operation of the $\sigma_{\mu\nu}$ being determined by the parameter u. This, in fact, is the lagrangian formulation of the Dirac equation suggested in C, Sec. XIV.

A rotation (and Lorentz transformation) by angle $\omega_{\mu\nu}$ in the $\mu\nu$-plane, is represented in Dirac theory by the operator $\exp(\frac{1}{4}i\omega_{\mu\nu}\sigma_{\mu\nu})$. (The summation on both μ and ν accounts for a factor 2.) Hence, we can say that Eq. (32-a) means that the amplitude for arrival is $\exp(iS)$, where S is the classical action $-\int[\frac{1}{2}(dx_\mu/du)^2+B_\mu dx_\mu/du]du$, but the orientation represented by the hypercomplex amplitude has rotated at each point in its path at an angular velocity (per du) equal to the field strength at that point. (Angular velocity in four dimensions is an antisymmetric tensor of second rank, as is the field strength.)

Since the potentials appear in exponential form, this may be directly connected to the form representing the action of virtual photons. The result is a set of rules like that for the Klein-Gordon case, Sec. IX, but with an additional coupling $F_{\mu\nu}\gamma_\mu\gamma_\nu$. They may be shown to be algebraically equivalent to the rules usually given for the Dirac equation, but are somewhat more complicated and not very interesting. There are some properties of the Dirac electron, however, which are more obvious in this formulation than in the usual one, and these we will discuss.

It is apparent from Eq. (32-a) that in the classical limit the trajectory is that of minimum S and therefore satisfies

$$d^2x_\mu/du^2=(dx_\nu/du)F_{\nu\mu}. \tag{33-a}$$

(Hence, $(dx_\mu/du)^2=(ds/du)^2$ is a constant of the motion where s is the proper time, and the minimum action S is $-\frac{1}{2}(ds/du)^2u$ plus a term independent of u. Since this is to vary as $-\frac{1}{2}m^2u$, we find $ds=mdu$.) As $\hbar\to0$, the magnetic moment approaches zero and does not affect the trajectory. But since the intrinsic spin angular momentum also goes to zero, the rate of precession of spin has a classical limit. For completeness we should also give the equation

[25] Y. Nambu, Prog. Theor. Phys. (Japan) 5, 82, 1950.

of motion of the spin axis (just as the spot on a billiard cue ball has a motion, although it does not affect the trajectory of the ball). From Eq. (32-a) we see the spin axis precesses at angular velocity $F_{\mu\nu}$ (per du). (These are well-known results of the WKB approximation method when applied to the Dirac equation.) Since Eq. (33-a) says only that dx_μ/du precesses at the same angular velocity, we can summarize the classical equations of motion, and of spin precession for a Dirac electron as: *The velocity vector and spin plane are fixed in a four-dimensional coordinate system turning at each instant at an angular velocity per unit proper time equal to e/m times the field strength acting on the electron at that instant.* (For example, for a slowly moving electron in a magnetic field **B** the velocity vector revolves about the magnetic field as an axis at angular velocity $\omega=(e/m)\mathbf{B}$, the cyclotron frequency. The spin does likewise precessing therefore at the same frequency, which is twice the Larmor frequency.)

I have expended considerable effort to obtain an equally simple word description of the quantum mechanics of the Dirac equation. Very many modes of description have been found, but none are thoroughly satisfactory. For example, that of Eq. (32-a) is incomplete, even aside from the geometrical mysteries involved in the superposition of hypercomplex numbers. For in (32-a) the field enters in two *apparently unrelated* ways, once into defining S and again in the rotation rate. In the classical limit both effects of the field can be neatly stated in one principle. What makes things particularly simple in quantum mechanics if, for a diffusing wave, a rotation at rate $F_{\mu\nu}$ is accompanied by a phase shift equal to the line integral of A_μ?[26]

In the case that the fields $F_{\mu\nu}$ are constant in space and time, the operator factor $\exp(\frac{1}{4}iu\sigma_{\mu\nu}F_{\mu\nu})$ is independent of the trajectory and factors out of Eq. (32-a). The remaining path integral is gaussian and can be carried out exactly (Appendix C), giving the results of Fock[4] and Nambu.[25]

If the operator on the right-hand side of Eq. (31-a) is considered as a type of hamiltonian, the rate of change with u of all the relevant physical quantities (given by the commutator with this operator) are very easily interpreted by classical analogy.

There are, of course, twice as many solutions of Eq. (30-a) as solutions of the Dirac equation (50). (The others correspond to Eq. (50) with negative m.) If χ is a solution of Eq. (30-a), the projected part $\psi=(2m)^{-1}(i\nabla-B+m)\chi$ solves the Dirac equation (50). Projection operators must still be used, therefore, in calculating matrix elements if Eq. (30-a) in perturbation is used instead of the Dirac equation.[27]

[26] If the $\frac{1}{4}\sigma_{\mu\nu}F_{\mu\nu}$ term is considered to have a coefficient α analogous to a kind of anomalous magnetic moment, difficulties arise in the resulting theory unless $\alpha=1$ or $\alpha=0$. Thus, the real part of L, in the amplitude for a vacuum to remain a vacuum, $C_v=\exp(-L)$, should always be positive if the theory is to be easily interpreted (see I, Sec. V). For general α, it seems that the real part of L is positive for some processes (or potentials), negative for others. It is always positive only if $\alpha=1$. But for $\alpha=0$ it is always negative, so we can reinterpret the theory in this case as referring to Bose particles, in which case C_v should be $\exp(+L)$ (I, Sec. V). For $\alpha=0$, Eq. (30-a) becomes the Klein-Gordon equation, of course.

[27] A convenient way to make the correspondence of solutions χ of Eq. (30-a) and ψ of the Dirac equation unique is to assume χ is also an eigenfunction of γ_5, that is, $\gamma_5\chi=i\chi$. This is possible, as γ_5 commutes with the operators of Eq. (30-a). Then, for each ψ, the corresponding χ is $\chi=(1-i\gamma_5)\psi$.

PHYSICAL REVIEW VOLUME 97, NUMBER 3 FEBRUARY 1, 1955

Slow Electrons in a Polar Crystal

R. P. FEYNMAN
California Institute of Technology, Pasadena, California
(Received October 19, 1954)

A variational principle is developed for the lowest energy of a system described by a path integral. It is applied to the problem of the interaction of an electron with a polarizable lattice, as idealized by Fröhlich. The motion of the electron, after the phonons of the lattice field are eliminated, is described as a path integral. The variational method applied to this gives an energy for all values of the coupling constant. It is at least as accurate as previously known results. The effective mass of the electron is also calculated, but the accuracy here is difficult to judge.

AN electron in an ionic crystal polarizes the lattice in its neighborhood. This interaction changes the energy of the electron. Furthermore, when the electron moves the polarization state must move with it. An electron moving with its accompanying distortion of the lattice has sometimes been called a polaron. It has an effective mass higher than that of the electron. We wish to compute the energy and effective mass of such an electron. A summary giving the present state of this problem has been given by Fröhlich.[1] He makes simplifying assumptions, such that the crystal lattice acts much like a dielectric medium, and that all the important phonon waves have the same frequency. We will not discuss the validity of these assumptions here, but will consider the problem described by Fröhlich as simply a mathematical problem. Aside from its intrinsic interest, the problem is a much simplified analog of those which occur in the conventional meson theory when perturbation theory is inadequate. The method we shall use to solve the polaron problem is new, but the pseudoscalar symmetric meson field problems involve so many further complications that it cannot be directly applied there without further development.

We shall show how the variational technique which is so successful in ordinary quantum mechanics can be extended to integrals over trajectories.

STATEMENT OF THE PROBLEM

With Fröhlich's assumptions, the problem is reduced to that of finding the properties of the following Hamiltonian:

$$H=\tfrac{1}{2}\mathbf{P}^2+\sum_K a_K^+a_K+i(\sqrt{2}\pi\alpha/V)^{\frac{1}{2}}\sum_K\frac{1}{K}\times[a_K^+\exp(-i\mathbf{K}\cdot\mathbf{X})-a_K\exp(i\mathbf{K}\cdot\mathbf{X})]. \quad (1)$$

Here $\mathbf{X}$ is the vector position of the electron, $\mathbf{P}$ its conjugate momentum, a_K^+, a_K the creation and annihilation operators of a phonon (of momentum $\mathbf{K}$). The frequency of a phonon is taken to be independent of $\mathbf{K}$. Our units are such that $\hbar$, this frequency, and the electron mass are unity. The quantity α acts as a coupling constant, which may be large or small. In conventional units it is given by

$$\alpha=\frac{1}{2}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right)\frac{e^2}{\hbar\omega}\left(\frac{2m\omega}{\hbar}\right)^{\frac{1}{2}},$$

where ϵ, ϵ_∞ are the static and high frequency dielectric constant, respectively. In a typical case, such as NaCl, α may be about 5. The wave function of the system satisfies ($\hbar=1$)

$$i\partial\psi/\partial t=H\psi, \quad (2)$$

so that if φ_n and E_n are the eigenfunctions and eigenvalues of H,

$$H\varphi_n=E_n\varphi_n, \quad (3)$$

then any solution of (2) is of the form

$$\psi=\sum_n C_n\varphi_n e^{-iE_nt}.$$

Now we can cast (1) and (2) into the Lagrangian form of quantum mechanics and then eliminate the field oscillators (specializing to the case that all phonons are virtual). Doing this in exact analogy to quantum electrodynamics,[2] we find that we must study the sum over all trajectories $\mathbf{X}(t)$ of $\exp(iS')$, where

$$S'=\frac{1}{2}\int\left(\frac{d\mathbf{X}}{dt}\right)^2dt+2^{-\frac{1}{2}}\alpha i\int\int|\mathbf{X}_t-\mathbf{X}_s|^{-1}e^{-i|t-s|}dtds. \quad (4)$$

This sum will depend on the initial and final conditions and on the time interval T. Since it is a solution of the Schrödinger Eq. (2), considered as a function of T it will contain frequencies E_n, the lowest of which we seek. It is difficult to isolate the lowest frequency, however.

For that reason, consider the mathematical problem of solving

$$\partial\psi/\partial t=-H\psi, \quad (5)$$

without question as to the meaning of t. This has the same eigenvalues and eigenfunctions as (3), but a

[1] H. Fröhlich, Advances in Physics 3, 325 (1954). References to other work is given here.

[2] R. P. Feynman, Phys. Rev. 80, 440 (1950).

solution will have the form

$$\psi=\sum_n C_n e^{-E_n t}\varphi_n.$$

For large t any solution therefore asymptotically dies out exponentially, the last exponent surviving being that of the lowest E, say E_0.

An equation such as (5) can be converted to a path integral just as easily as (2) is, and the integral over the oscillator coordinates can again be done in an analogous way. The Lagrangian form corresponding to (5) turns out to be

$$K=\int \exp S\mathfrak{D}\mathbf{X}(t), \tag{6}$$

with

$$S=-\frac{1}{2}\int\left(\frac{d\mathbf{X}}{dt}\right)^2 dt + 2^{-\frac{1}{2}}\alpha\int\int |\mathbf{X}_t-\mathbf{X}_s|^{-1}e^{-|t-s|}dtds. \tag{7}$$

This is just as one might expect from replacing t in (4) by $-it$. Now, since K is a solution of (5), its asymptotic form for a large t interval, 0 to T is

$$K\sim e^{-E_0 T} \tag{8}$$

as $T\to\infty$. Therefore, we must estimate the path integral (6) for large T.

VARIATIONAL PRINCIPLE

The method we shall use is a type of variational method. Choose any S_1 which is simple and purports to be some sort of approximation to S. Then write

$$\int \exp S\mathfrak{D}\mathbf{X}(t)=\int \exp(S-S_1)\exp S_1\mathfrak{D}\mathbf{X}(t). \tag{9}$$

Now this last expression can be looked upon as the average of $\exp(S-S_1)$, the average being taken with positive weight $\exp S_1$. But for any set of real quantities f the average of $\exp f$ exceeds the exponential of the average,

$$\langle\exp f\rangle\geq\exp\langle f\rangle. \tag{10}$$

Hence if in (9) we replace $S-S_1$ by its average,

$$\langle S-S_1\rangle = \int (S-S_1)\exp S_1\mathfrak{D}\mathbf{X}(t)\Big/\int \exp S_1\mathfrak{D}\mathbf{X}(t), \tag{11}$$

we will underestimate the value of (9). Therefore, if E is computed from

$$\int \exp(\langle S-S_1\rangle)\exp S_1\mathfrak{D}\mathbf{X}(t)\sim\exp -ET, \tag{12}$$

then we know that E exceeds the true E_0,

$$E\geq E_0. \tag{13}$$

If there are any free parameters in S_1 we can choose as the "best" values those which minimize E.

Since $\langle S-S_1\rangle$ defined in (11) is proportional to T, let us write

$$\langle S-S_1\rangle=sT. \tag{14}$$

Furthermore, the factor $\exp\langle S-S_1\rangle$ in (12) is constant, of course, and may be taken outside the integral. Finally, suppose the lowest energy E_1 for the action S_1 is known,

$$\int \exp S_1\mathfrak{D}\mathbf{X}(t)\sim\exp(-E_1T), \tag{15}$$

then we have

$$E=E_1-s \tag{16}$$

from (12), with s given by (11) and (14). (In the case that S and S_1 are both simple actions [of the form of (18) below] this can readily be shown to be equivalent to the usual variational principle.)

POSSIBLE TRIAL ACTIONS

Some of the methods which have been applied to this problem, so far, correspond to various choices for S_1. The perturbation method corresponds to $S_1=-\frac{1}{2}\int(d\mathbf{X}/dt)^2dt$ and gives

$$E=-\alpha. \tag{17}$$

We see immediately that the perturbation result is an upper limit to E_0, a result proven only with much greater effort by more usual methods, by Gurari[3] and Lee and Pines.[4] Another suggestion is

$$S_1=-\tfrac{1}{2}\int(d\mathbf{X}/dt)^2dt+\int V(\mathbf{X}_t)dt, \tag{18}$$

where V is a potential to be chosen. If a Coulomb potential is chosen, $V(R)=Z/R$, and the parameter Z varied, one finds

$$E=-(25/256)\alpha^2=-0.098\alpha^2$$

asymptotically for the case that α is very large. For large α this corresponds to Landau's method[1] with a trial function of the form $e^{-\beta r}$. If a harmonic potential $V(R)=kR^2$ is used (corresponding to a Gaussian trial function in Landau's method) the value is somewhat improved:

$$E=-(1/3\pi)\alpha^2=-0.106\alpha^2. \tag{19}$$

If α is not so large, the form (18) can still be used in (16). The evaluation of s requires knowledge of the eigenfunctions and eigenvalues for the potential V

[3] M. Gurari, Phil. Mag. 44, 329 (1953).
[4] T. Lee and D. Pines, Phys. Rev. 88, 960 (1952). Lee, Low, and Pines, Phys. Rev. 90, 297 (1953).

The result is somewhat difficult to evaluate for the Coulomb potential, but fairly simple for the harmonic case [see (34) below]. However, it is readily shown that for any α less than about 6 no choice of V can improve the result (17) for $V=0$. Fröhlich has asked for a method which works uniformly over the entire range of α. He points out that the artificial binding to a special origin, which (18) implies, is a disadvantage. It is this which presumably makes any potential V give a poorer result than $V=0$ for small α.

To remedy this, I thought a good idea would be to use for S_1 the action for a particle bound by a potential $V(\mathbf{X}-\mathbf{Y})$ to another particle of coordinate $\mathbf{Y}$. This latter could have finite mass, so no permanent origin would be assumed. Of course the action for such a system would contain both $\mathbf{X}(t)$ and $\mathbf{Y}(t)$. But the variables $\mathbf{Y}(t)$ could be integrated out, at least in principle, leaving an effective S_1 depending only on $\mathbf{X}$. At first I tried a Coulomb interaction for $V(\mathbf{X}-\mathbf{Y})$ but it was rather complicated. The technique may be useful in more difficult problems. But here we have already seen that an harmonic binding should be as good, if not better. Further, an extra particle bound harmonically has its variables $\mathbf{Y}(t)$ appearing quadratically in the action. It may therefore be easily eliminated explicitly. The result we know from studies of similar problems in electrodynamics. We are, in this way, led to consider the choice

$$S_1=-\tfrac{1}{2}\int\left(\frac{d\mathbf{X}}{dt}\right)^2dt-\tfrac{1}{2}C\int\int[\mathbf{X}_t-\mathbf{X}_s]^2\times\exp(-w|t-s|)dtds, \quad (20)$$

where C and w are parameters, to be chosen later to minimize E.

EVALUATION OF THE ENERGY

Since S_1 contains $\mathbf{X}$ only quadratically, all the necessary path integrals are easily done.[5] Because the method may not be familiar we outline it briefly here. Define the symbol $\langle\ \rangle$ as

$$\langle F\rangle=\int F\exp S_1\mathfrak{D}\mathbf{X}(t)\Big/\int\exp S_1\mathfrak{D}\mathbf{X}(t).$$

Then comparison of S_1 and S shows that

$$s=\frac{1}{T}\langle S-S_1\rangle=2^{-\frac{1}{2}}\alpha\int\langle|\mathbf{X}_t-\mathbf{X}_s|^{-1}\rangle e^{-|t-s|}ds+\tfrac{1}{2}C\int\langle(\mathbf{X}_t-\mathbf{X}_s)^2\rangle e^{-w|t-s|}ds=A+B. \quad (21)$$

[5] R. P. Feynman, Phys. Rev. 84, 108 (1951), Appendix C.

We concentrate first on the first term A of (21). In it we may express $|\mathbf{X}_t-\mathbf{X}_s|^{-1}$ by a Fourier transform,

$$|\mathbf{X}_t-\mathbf{X}_s|^{-1}=\int d^3\mathbf{K}\exp[i\mathbf{K}\cdot(\mathbf{X}_t-\mathbf{X}_s)](2\pi^2K^2)^{-1}. \quad (23)$$

For this reason we need to study

$$\langle\exp[i\mathbf{K}\cdot(\mathbf{X}_\tau-\mathbf{X}_\sigma)]\rangle=\int\exp S_1\exp[i\mathbf{K}\cdot(\mathbf{X}_\tau-\mathbf{X}_\sigma)]\mathfrak{D}\mathbf{X}(t)\Big/\int\exp S_1\mathfrak{D}\mathbf{X}(t). \quad (23)$$

The integral in the numerator is of the form

$$I=\int\exp\left[-\tfrac{1}{2}\int\left(\frac{d\mathbf{X}}{dt}\right)^2dt-\tfrac{1}{2}C\int\int(\mathbf{X}_t-\mathbf{X}_s)^2\times e^{-w|t-s|}dtds+\int\mathbf{f}(t)\cdot\mathbf{X}(t)dt\right]\mathfrak{D}\mathbf{X}(t), \quad (24)$$

where specifically

$$\mathbf{f}(t)=i\mathbf{K}\delta(t-\tau)-i\mathbf{K}\delta(t-\sigma). \quad (25)$$

Now we shall find (24) insofar as it depends on $\mathbf{f}$ or $\mathbf{K}$ aside from a normalization factor which drops out in (23). Incidentally let us notice that the three rectangular components separate in (24) and we need only consider a scalar case. The method of integration is to substitute $X(t)=X'(t)+Y(t)$, where $X'(t)$ is that special function for which the exponent is maximum. The variable of integration is now $Y(t)$. Since the exponent is quadratic in $X(t)$ and X' renders it an extremum, it can contain $Y(t)$ only quadratically. Evidently Y then separates off as a factor not containing f, which may be integrated to give an unimportant constant (depends on T only). Therefore within such a constant

$$I=\exp\left[-\tfrac{1}{2}\int\dot{X}_t'^2dt-\tfrac{1}{2}C\int\int(X_t'-X_s')^2\times e^{-w|t-s|}dtds+\int f(t)X_t'dt\right], \quad (26)$$

where X' is that function which minimizes the expression [subject for convenience, to $X'(0)=X'(T)=0$ if the time interval is 0 to T]. The variation problem gives the integral equation

$$d^2X'(t)/dt^2=2C\int(X_t'-X_s')e^{-w|t-s|}ds-f(t). \quad (27)$$

Using (27), (26) can be simplified to

$$I=\exp\left[\tfrac{1}{2}\int f(t)X'(t)dt\right]. \quad (28)$$

We need merely solve (27) and substitute into (28). To do this we define

$$Z(t)=\frac{w}{2}\int e^{-w|t-s|}X_s'ds,$$

so that

$$d^2Z(t)/dt^2=w^2[Z(t)-X'(t)],$$

while (27) is

$$d^2X'(t)/dt^2=\frac{4C}{w}[X'(t)-Z(t)]-f(t).$$

The equations are readily separated and solved. The solution for $X'(t)$ substituted into (28) gives, for the case (25),

$$I=\langle\exp[i\mathbf{K}\cdot(\mathbf{X}_\tau-\mathbf{X}_\sigma)]\rangle$$
$$=\exp\left[-\frac{2CK^2}{v^3w}(1-e^{-v|\tau-\sigma|})-\frac{w^2}{2v^2}K^2|\tau-\sigma|\right], \quad (29)$$

where we have made the substitution

$$v^2=w^2+(4C/w). \quad (30)$$

The result is correctly normalized since it is valid for $\mathbf{K}=0$. The integral on K in (22) is a simple Gaussian, so that substitution into A gives

$$A=\pi^{-\frac{1}{2}}\alpha v\int_0^\infty\left[w^2\tau+\frac{v^2-w^2}{v}(1-e^{-v\tau})\right]^{-\frac{1}{2}}e^{-\tau}d\tau. \quad (31)$$

To find B we need $\langle(\mathbf{X}_t-\mathbf{X}_s)^2\rangle$. This can be obtained by expanding both sides of (29) with respect to $\mathbf{K}$ up to order K^2. Therefore

$$\tfrac{1}{3}\langle(\mathbf{X}_\tau-\mathbf{X}_\sigma)^2\rangle=\frac{4C}{v^3w}(1-e^{-|\tau-\sigma|})+\frac{w^2}{v^2}|\tau-\sigma|.$$

The integral in B is now easily performed and the expression simplifies to

$$B=3C/vw. \quad (32)$$

Finally we need E_1, the energy belonging to our action S_1. This is most easily obtained by differentiating both sides of (15) with respect to C. One finds immediately

$$CdE_1/dC=B,$$

so that, in view of (32) and (30), integration gives

$$E_1=\tfrac{3}{2}(v-w),$$

since $E_1=0$ for $C=0$. Since $E_1-B=(3/4v)(v-w)^2$ we obtain finally for our energy expression:

$$E=\frac{3}{4v}(v-w)^2-A, \quad (33)$$

with A given in (31). The quantities v, w can be considered as two parameters which may be varied separately to obtain a minimum.

The integral in A unfortunately cannot be performed in closed form, so that a complete determination of E requires numerical integration. It is, however, possible to obtain approximate expressions in various limiting cases. The case of large α corresponds to large v. The choice $w=0$ leads to an integral

$$A=\pi^{-\frac{1}{2}}\alpha v^{\frac{1}{2}}\int_0^\infty e^{-\tau}d\tau[1-e^{-v\tau}]^{-\frac{1}{2}}=\frac{\alpha\Gamma(1/v)}{v^{\frac{1}{2}}\Gamma(\frac{1}{2}+1/v)}, \quad (34)$$

and $E_1=3v/4$. It corresponds to the use of a fixed harmonic binding potential in (18). For large v, $e^{-v\tau}$ can be neglected, so that $A=\pi^{-\frac{1}{2}}\alpha v^{\frac{1}{2}}$. This corresponds to using a Gaussian trial function in Landau's method. For α less than 5.8 and $w=0$, (33) does not give a minimum unless $v=0$, so that the $w=0$ case does not give a single expression for all ranges of α. In spite of this disadvantage the result with (34) is relatively simple and fairly accurate. For $\alpha>6$, only fairly large v are important, and the asymptotic formula (good to 1 percent for $v>4$),

$$A=\alpha(v/\pi)^{\frac{1}{2}}[1+(2\ln 2)/v],$$

is convenient. Fröhlich, however, considers the discontinuity at $\alpha=6$ as a serious disadvantage, which it is the purpose of this paper to avoid. This we do by choosing w different from zero.

Let us study (33), just for small α, in case w is not zero. The minimum will occur for v near w. Therefore write $v=(1+\epsilon)w$, consider ϵ small, and expand the root in (31). This gives

$$A=\alpha(v/w)\left[1-\epsilon\int_0^\infty\tau^{-\frac{1}{2}}e^{-\tau}(1-e^{-w\tau})d\tau/w\pi^{\frac{1}{2}}+\cdots\right].$$

The integral is

$$2w^{-1}[(1+w)^{\frac{1}{2}}-1]=P. \quad (35)$$

The problem (33) then corresponds, in this order, to minimizing

$$E=\tfrac{3}{4}w\epsilon^2-\alpha-\alpha\epsilon(1-P).$$

That is,

$$\epsilon=2\alpha(1-P)/3w,$$

which is valid for small α only, as ϵ was assumed small. The resulting energy is

$$E=-\alpha-\alpha^2(1-P)^2/3w.$$

Our method therefore gives a correction even for small α. It is least for $w=3$, in which case it gives

$$E=-\alpha-\alpha^2/81=-\alpha-1.23(\alpha/10)^2. \quad (36)$$

It is not sensitive to the choice of w. For example, for $w=1$ the 1.23 falls only to 0.98. The method of Lee and Pines[6] gives exactly this result (36) to this order. The perturbation expansion has been carried to

[6] T. Lee and D. Pines, Phys. Rev. 92, 883 (1953).

second order by Haga[7] who shows that the exact coefficient of the $(\alpha/10)^2$ term should be 1.26, so that our variational method is remarkably accurate for small α.

The opposite extreme of large α corresponds to large v, and, as we shall see, w near 1. Since $v \gg w$ the integral (31) reduces in the first approximation to (34), which we can use in its asymptotic form. The next approximation in w can be obtained by expanding the radical in (31), considering $w/v \ll 1$. Furthermore, e^{-vr} is negligible. In this way we get

$$E=\frac{3}{4v}-(v-w)^2-\alpha(v/\pi)^{\frac{1}{2}}\left(1+\frac{2\ln 2}{v}-\frac{w^2}{2v}\right). \quad (37)$$

This is minimum, within our approximation of large v, when $w=1$, and $v=(4\alpha^2/9\pi)-(4\ln 2-1)$:

$$E=-\alpha^2/3\pi-3\ln 2-\tfrac{3}{4}=-0.106\alpha^2-2.83. \quad (38)$$

The approximations do not keep E as an upper limit as, unfortunately, the further terms, of order $1/\alpha^2$ are probably positive.

For further numerical work it is probably sufficiently accurate to take $w=1$ for all α, rather than do the extra work needed to minimize this extra variable. This value of w means that the trial S_1 has the same time exponential in the interaction term as does S. For small α, that is, v near 1, the integral can be expanded in a power series in $(v-1)$. The resulting energy is $(w=1)$:

$$E=-\alpha-0.98(\alpha/10)^2-0.60(\alpha/10)^3 -0.14(\alpha/10)^4\cdots. \quad (39)$$

The two expressions (38), (39) fit fairly well near $\alpha=5$. For practical purposes it may suffice to use (39) below $\alpha=5$ and (38) above. If more accuracy than 3 percent is needed near $\alpha=5$ numerical integration of A must be performed. The value of v which gives (39) is

$$v=1+1.14(\alpha/10)+1.35(\alpha/10)^2+1.88(\alpha/10)^3.$$

This may help to choose an appropriate v. For $w=3$ the results are

$$E=-\alpha-1.23(\alpha/10)^2-0.64(\alpha/10)^3\cdots,$$
$$v=3+2.22(\alpha/10)+1.97(\alpha/10)^2\cdots.$$

EFFECTIVE MASS

Another quantity of interest is the effective mass. If the particle moves with a mean group velocity V, its energy should be greater. For small V the energy goes as V^2, and writing it as $mV^2/2$ we call m the effective mass. Since there is an operator analogous to the momentum which commutes with the Hamiltonian, it would be expected that there is a variational principle which minimizes the energy for each momentum. That is, we ought to be able to extend our method to yield an upper limit to the energy for each value of V, or better, of momentum Q. We have not found the expected extension.

If we limit ourselves just to finding the effective mass for low velocities, however, we may proceed in this manner: For a free particle of mass m whose initial coordinate is 0 and final coordinate is $\mathbf{X}_T$ the sum on trajectories is

$$\exp(-m\mathbf{X}_T^2/2T). \quad (40)$$

Hence we can study the effective mass for our system by studying the asymptotic form of (6) in the case $\mathbf{X}_T \neq 0$. The asymptotic form should vary for small $\mathbf{X}_T$ as $\exp(-E_0T-m\mathbf{X}_T^2/2T)$, its dependence on $\mathbf{X}_T$ determining m. This only requires that (27) be solved for the boundary conditions $\mathbf{X}'=0$ at $t=0$ and $\mathbf{X}'=\mathbf{X}_T$ at $t=T$. There are some confusing complications at the end points so it is easier to proceed as follows. We will put $\mathbf{X}_T=\mathbf{U}T$ so that the propagation (40) is $\exp(-\frac{1}{2}mU^2T)$. [Note that $\mathbf{U}$ is not a physical velocity because t is an artificial parameter in Eq. (5), and is not the time.] That is, we seek the total energy and equate it to $E_0+\frac{1}{2}mU^2$. But if we substitute $\mathbf{X}'=\mathbf{X}''+\mathbf{U}t$ into (27), we see that it is a solution if $\mathbf{X}''$ is. This $\mathbf{X}''$ goes from 0 at $t=0$ to 0 at $t=T$, and is therefore our previous solution. Such a substitution into (26) means that the term involving $\mathbf{f}$ adds a term $\exp(\int t\mathbf{U}\cdot\mathbf{f}dt)$ so that this is the factor by which I is multiplied, aside from normalization. For the $\mathbf{f}$ given in (25) this is $\exp[i\mathbf{K}\cdot\mathbf{U}(\tau-\sigma)]$ so that we now have

$$\langle\exp[i\mathbf{K}\cdot(\mathbf{X}_\tau-\mathbf{X}_\sigma)]\rangle = \exp\left[-\frac{K^2}{2v^2}F(|\tau-\sigma|)+i\mathbf{K}\cdot\mathbf{U}(\tau-\sigma)\right], \quad (41)$$

where

$$F(\tau)=w^2\tau+\frac{v^2-w^2}{v}(1-e^{-v\tau}). \quad (42)$$

Substitution into (22) and (21) gives for A the value

$$A(\mathbf{U})=2^{-\frac{1}{2}}\alpha\int_0^\infty\int(2\pi^2K^2)^{-1}e^{-\tau} \times\exp\left[-\frac{K^2}{2v^2}F(\tau)+i\mathbf{K}\cdot\mathbf{U}\tau\right]d^3\mathbf{K}d\tau. \quad (43)$$

Second differentiation of (41) with respect to $\mathbf{K}$ shows that

$$\langle(\mathbf{X}_t-\mathbf{X}_s)^2\rangle=3F(t-s)v^{-2}+U^2(t-s)^2,$$

so that one obtains for B the value

$$B=\frac{3C}{vw}+\frac{2C}{w^3}U^2.$$

We again find E_1 from $dE_1/dC=B/C$ and $E_1=\frac{1}{2}U^2$ for $C=0$. Thus

$$E_1=\tfrac{3}{2}(v-w)+\tfrac{1}{2}U^2(1+4Cw^{-3}),$$

[7] E. Haga, Progr. Theoret. Phys. (Japan) 11, 449 (1954).

and our final expression is

$$E=\tfrac{1}{2}U^2+(3/4v)(v-w)^2-A(\mathbf{U}). \tag{44}$$

We next expand $A(\mathbf{U})$ to order U^2 and write the kinetic energy as $mU^2/2$ to find, finally,

$$m=1+\tfrac{1}{3}\pi^{-\frac{1}{2}}\alpha v^3\int_0^\infty [F(\tau)]^{-\frac{3}{2}}e^{-v\tau}\tau^2 d\tau. \tag{45}$$

The values of the parameters to use in (45) are those which were previously found to minimize E when $\mathbf{U}=0$.

For small α this gives

$$m=1+\tfrac{1}{6}\alpha+0.025\alpha^2+\cdots \tag{46}$$

for $w=3$, while for $w=1$ the 0.025 becomes 0.023. For large α it becomes

$$m=16\alpha^4/81\pi^4=202(\alpha/10)^4. \tag{47}$$

Our energy values, coming from a minimum principle, are much more accurate than the mass values, whose precision, especially for large α, is hard to judge. Since (46) and (47) do not match well, intermediate values of α require numerical integration of (45).

Lee and Pines[6] have worked with a different type of variational principle. It seems to be nearly as good as ours for α less than about 5, but is poor for larger α (for example, at $\alpha=15$, Lee and Pines find $E_0<-17.6$, while we find $E_0<-26.8$). This appears to contradict their statement that their method is exact for large α. They are referring to a different problem, however, in which the upper momenta are cut off. This means that in S in (7) the function $|\mathbf{X}_t-\mathbf{X}_s|^{-1}$ is replaced by some other function $V(|\mathbf{X}_t-\mathbf{X}_s|)$ which differs for small $|\mathbf{X}_t-\mathbf{X}_s|$. It is evident, for large α, that the best trajectory will be the one that wanders only slightly and the energy will be $2^{-\frac{1}{2}}\alpha V(0)$ in the limit. Their method gives this result in the limit, as ours would also. For the case where V is singular, so $V(0)$ does not exist their method is not exact, and it is inaccurate for for intermediate values of α even if $V(0)$ exists, if V has steep walls.

The method is readily extended to cases in which the photon frequencies are not constant, and the coupling is not just proportional to K^{-1}. The same trial action S_1 can be used, but the integral for A becomes more complicated. For the Hamiltonian

$$H=\tfrac{1}{2}P^2+\textstyle\sum_K\omega_K a_K^+a_K+V^{-\frac{1}{2}}\sum_K[C_K^*a_K^+\exp(-i\mathbf{K}\cdot\mathbf{X})+C_Ka_K\exp(+i\mathbf{K}\cdot\mathbf{X})],$$

Eq. (33) still holds; the only change is that the integral for A becomes

$$A=\int\int_0^\infty \exp\left[-\omega_K\tau-\frac{K^2}{2v^2}F(\tau)\right]|C_K|^2 d\tau d^3\mathbf{K}(2\pi)^{-3},$$

where $F(\tau)$ is given in (42).

An attempt has been made to apply this method to meson problems. The case of scalar nucleons interacting by scalar mesons seems tractable, but the greater complexity of the more realistic problems shows the need for further development.

We are limited in our choice of S_1 to quadratic functionals, for those are the only ones we can evaluate directly as path integrals. It would be desirable to find out how this method may be expressed in conventional notation, for a wider class of trial functionals might thereby become available.

I am indebted to H. Fröhlich for bringing the problem to my attention, and for his comments on it, and to G. Speisman for emphasizing the importance of the general inequality (10).

Note added in proof.—Professor Frölich and Professor Pines have kindly informed me that S. I. Pekar [Zhur. Eksptl. i Teort. Fiz. **19**, 796 (1949)] has calculated the limiting values of energy and mass for large α, by an adiabatic approximation. The energy is $-0.1088\alpha^2$ and the mass is $232(\alpha/10)^4$. Therefore our variational method gives an error of only 3 percent in the energy and 15 percent in the mass for large α, and presumably smaller errors for smaller α.

PHYSICAL REVIEW VOLUME 127, NUMBER 4 AUGUST 15, 1962

Mobility of Slow Electrons in a Polar Crystal

R. P. FEYNMAN
California Institute of Technology, Pasadena, California

R. W. HELLWARTH
Hughes Research Laboratories, Malibu, California

C. K. IDDINGS
Enrico Fermi Institute for Nuclear Studies, University of Chicago, Chicago, Illinois

AND

P. M. PLATZMAN
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received March 26, 1962)

We have obtained an approximate expression for the impedance function at all frequencies, temperatures, and coupling strengths of an electron coupled to a polar lattice (a system commonly called a polaron). The starting point for the calculation is the quantum mechanical expression for the expected current. The phonon coordinates are eliminated from this expression by well-known field-theory techniques. The resulting exact "influence functional" is then approximated by a corresponding quadratic "influence functional" which, it is hoped, imitates the real polaron. Correction terms are computed to account for the difference between the approximate impedance and the exact polaron impedance in a manner closely analogous to Feynman's treatment of the polaron self-energy. In fact, the analytic evaluation of the expression for the impedance obtained here is carried out using the approximate "influence functional" that was successfully employed in minimizing the binding (and free) energy of the polaron in earlier calculations. However, the accuracy obtained using this approximation, for the impedance calculation, is less satisfactory and its limitations are discussed. Nevertheless, beginning at intermediate coupling strengths, the approximate impedance produces a level structure of increasing complexity and narrowing resonances as the coupling strengthens. This suggests that further refinements may be fruitful. Methods for finding a better quadratic influence functional for use in our impedance expression as well as ways of improving the expression further are suggested. A comparison of our results with those of the Boltzmann equation points up interesting differences which arise from reversing the order of taking limits of zero frequency and coupling.

I. INTRODUCTION

AN electron in a polar crystal interacts with the surrounding crystal. The effect of this interaction is to surround the electron with a distorted lattice: a cloud of phonons. The nature of this system, "the polaron," has been extensively studied.[1–4] It is interesting as a phenomenon in solids, but it has an extended interest since it is one of the simplest examples of the interaction of a particle and a field. It is in many ways analogous to the problem of a nucleon interacting with a meson field. (The extra complications of spin and isotopic spin do not, however, permit direct use of the methods to be described here, without some extension of their power.) In cases of practical interest, the coupling between the electron and the longitudinal optical modes of vibration of the crystal is sufficiently strong that simple perturbation methods do not apply. It is the strong-coupling aspect of the problem which has aroused so much interest. For this reason, the "polaron problem" has generally been studied in a considerably idealized form.

It is assumed that in the undistorted lattice the electron would move as a free particle (with possibly an altered mass), that only the optical modes interact with the electron, that they do so in a very simple way, and that they all have the same frequency. These are quite drastic simplifications; however, sufficient data are not available to improve these assumptions so as to represent any actual crystal. The methods given here do not require these simplifications (except perhaps that the electron's kinetic energy is a quadratic function of its momentum); the same techniques can be readily applied to include variations of frequency and coupling of the optical modes with wave number, influences of other modes, etc., although some of the integrals done analytically here might have to be done numerically.

In discussing losses and mobility such idealization may alter completely the true behavior, because some essential loss mechanism such as lattice defects or interaction with acoustic phonons has been idealized away. It is important to appreciate, therefore, that in all the remaining analysis and discussion we shall be talking only about a strictly idealized problem.

One of us[1] has shown that the ground-state energy and effective mass of the polaron could be calculated with considerable accuracy from a variational principal obeyed by path integrals. Of more interest, experimentally, is the mobility of the polaron and, moro generally, its response to weak, spatially uniform, time varying electric fields. This is a more complicated problem involving the rate at which a drifting electron loses momentum by phonon interactions, through emissions of phonons or collisions with phonons already present. In the practical situation at temperatures not too near the melting point of the crystal the density of

[1] R. P. Feynman, Phys. Rev. 97, 660 (1955), hereafter to be called I.
[2] H. Frohlich, *Advances in Physics*, edited by N. F. Mott (Taylor and Francis, Ltd., London, 1954), Vol. 3, p. 325.
[3] S. I. Pekar, Zhur. Eksp. i Teoret. Fiz. 19, 796 (1949).
[4] T. D. Schultz, Phys. Rev. 116, 526 (1960).

optical phonons is quite low as a result of the high energy required to excite them. In our idealized model losses can occur only through collisions with optical phonons, so that these collisions could be analyzed by first finding the collision cross section and then using the Boltzmann equation (or equivalently the usual formulas for transport cross section) to get the mobility.[5] This is the technique generally employed in transport problems. Yet there exists a class of transport problems in which this cannot be done. If many phonons are colliding simultaneously with an electron most of the time, and if there are possibly quantum interferences among these collisions (such that the cross section for scattering from one phonon depends on the presence and behavior of others), the collisions cannot be separated in time as required for the validity of the Boltzmann transport equation. What we need to calculate (the average position of an electron at time t, if at $t=0$ a pulsed electric field was applied) can be easily written formally, but little has been done with such a form unless the coupling is weak or the collisions are well separated.[6,7]

A secondary interest which we had in this problem was to see if we could compute transport problems in cases when not only the perturbation theory, but also the Boltzmann equation is inadequate. Therefore, in spite of its lack of reality, we have analyzed the problem of the impedance of a polaron of arbitrary coupling strength in an oscillating electric field, for arbitrary temperatures (temperatures so high perhaps that the Boltzmann factor $e^{-\hbar\omega/kT}$ for the energy $\hbar\omega$ of the optical modes is not necessarily small).

In any specific range of conditions, such as low temperature, high temperature, or high frequency of external electric field, etc., special approximations might be made to obtain a better answer than is given by our general formula. However, it was of interest to see how well one could do in a general way for arbitrary values of the parameters.

II. FORMULATION OF THE MOBILITY PROBLEM IN TERMS OF THE ELECTRON COORDINATES ALONE

If a weak alternating electric field $E=E_0e^{i\nu t}$ is applied to the crystal in the x direction, the current induced (by motion of the electron) may be written as

$$j(\nu)=[z(\nu)]^{-1}E_0e^{i\nu t}. \tag{1}$$

This defines the impedance function $z(\nu)$ which we wish to calculate. We will assume that the crystal is isotropic so that $j=\langle\dot{x}\rangle$, where $\langle x\rangle$ is the expectation of the electron displacement in the x direction (taking the electric charge as unity). The displacement $\langle x\rangle$ is $E/i\nu z(\nu)$. Transformed to time variables, this implies that

$$\langle x(\tau)\rangle=-\int_{-\infty}^{\infty}iG(\tau-\sigma)E(\sigma)d\sigma, \tag{2}$$

where $G(\tau)$, i times the electron displacement at time τ induced by a pulsed electric field at time zero, has the inverse transform

$$\int_{-\infty}^{\infty}G(\tau)e^{-i\nu\tau}d\tau=G(\nu)=[\nu z(\nu)]^{-1}. \tag{3}$$

We take $G(\tau)=0$ for $\tau<0$.[8]

The effect of a perturbing field $\mathbf{E}(t)$ in the x direction is to add to the complete Hamiltonian of the system H, the term $-xE(t)\equiv-\mathbf{E}\cdot\mathbf{X}$ (where x is the component of the vector position of the electron $\mathbf{X}$ in the direction of the field). If at some time (a), long before the field is turned on [i.e., $\mathbf{E}(t)=0$ for $t<a$] the state of the system is represented by the density matrix ρ_a, then the density matrix at time τ is $U(\tau,a)\rho_aU'^{-1}(\tau,a)$. Thus, the expected position at time τ is

$$\langle x(\tau)\rangle=\mathrm{Tr}[xU(\tau,a)\rho_aU'^{-1}(\tau,a)], \tag{4}$$

where

$$U(\tau,a)=\exp\left\{-i\int_a^{\tau}[H_s-\mathbf{X}_s\cdot\mathbf{E}(s)]ds\right\} \tag{5}$$

is the unitary operator for the development of a state in time with the complete Hamiltonian $H-\mathbf{X}\cdot\mathbf{E}$.

We use a time-ordered operator notation; all unprimed operators are placed to the left, latest times farthest to the left, then the matrix ρ at the right and finally all primed operators on the right of ρ, with latest times farthest to the right.[9] Thus, primed operators are ordered oppositely to unprimed. We can, therefore, write

$$U'^{-1}(\tau,a)=\exp\left\{i\int_a^{\tau}[H_s'-\mathbf{X}_s'\cdot\mathbf{E}'(s)]ds\right\}. \tag{6}$$

The quantity E is not an operator but simply a function of s so that in (4), $E'(s)=E(s)$. However, as we shall see in a moment, it is convenient to handle a more general case where E and E' are different arbitrary functions of s.

For weak fields we expand (4) to first order in E and find an expression for $x(\tau)$ of the form (2). Evidently, $-iG(\tau-\sigma)$ is the response to a δ function E, so we may set $E(s)=\epsilon\delta(s-\sigma)=E'(s)$, substitute into (4), and expand the exponential to first order in ϵ. However, we note that (4) itself may be considered to be $-i$ times the first functional derivative with respect to $E(\tau)-E'(\tau)$ of

$$g=\mathrm{Tr}[U(b,a)\rho_aU'^{-1}(b,a)], \tag{7}$$

[5] J. Howarth and E. H. Sondheimer, Proc. Roy. Soc. (London) A219, 53 (1953).
[6] R. Kubo, J. Phys. Soc. Japan 12, 570, 1203 (1957).
[7] M. Lax, Phys. Rev. 109, 1921 (1958).
[8] In our idealized model X and E will be in the same direction, although in general (in the presence of magnetic field or anisotropic crystalline fields), G and z will be tensors.
[9] R. P. Feynman, Phys. Rev. 84, 108 (1951).

as $b\to+\infty$, and $a\to-\infty$. That is to say, we calculate g from (5), (6), and (7) with

$$E(s)=\epsilon\delta(s-\sigma)+\eta\delta(s-\tau) \tag{8a}$$

and

$$E'(s)=\epsilon\delta(s-\sigma)-\eta\delta(s-\tau). \tag{8b}$$

The quantity we require is

$$G(\tau-\sigma)=\tfrac{1}{2}(\partial^2 g/\partial\eta\partial\epsilon)_{\eta=\epsilon=0}. \tag{9}$$

If the initial state is one of a definite temperature T, then

$$\rho_a=\exp(-\beta H)/Q, \tag{10}$$

where $\beta=1/kT$ and Q is a normalizing constant, which we eliminate by calculating $(1/2g)(\partial^2 g/\partial\eta\partial\epsilon)$ evaluated at $\epsilon=\eta=0$.

The Hamiltonian representing an electron in interaction with the vibrational modes of a crystal is

$$H=\mathbf{P}^2/2m+\textstyle\sum_{\mathbf{K}}\omega_{\mathbf{K}}a_{\mathbf{K}}^{\dagger}a_{\mathbf{K}}+V^{-1/2}\sum_{\mathbf{K}}[C_{\mathbf{K}}^{*}a_{\mathbf{K}}^{\dagger}\exp(-i\mathbf{K}\cdot\mathbf{X})+C_{\mathbf{K}}a_{\mathbf{K}}\exp(i\mathbf{K}\cdot\mathbf{X})]. \tag{11}$$

In this expression, $a_{\mathbf{K}}$, $a_{\mathbf{K}}^{\dagger}$ are the annihilation and creation operation of phonons of momentum $\mathbf{K}$, frequency $\omega_{\mathbf{K}}$, coupled to the electron via the coupling coefficient $C_{\mathbf{K}}$; $\mathbf{P}$ is the momentum of the electron; $\mathbf{X}$ is its coordinate; m is its effective mass calculateed in a fixed lattice; V is the crystal volume. We take $\hbar=1$, $m=1$.

As a specific example we shall take the simplified model of Fröhlich[2] in which $\omega_{\mathbf{K}}=1$ independent of $\mathbf{K}$, and $C_{\mathbf{K}}=i2^{3/4}\pi^{1/2}\alpha^{1/2}/|\mathbf{K}|$, where α is a constant related to the dielectric constant; intermediate coupling corresponds to $\alpha\approx 6$.

The quantity ρ_a, the initial distribution, should be $e^{-\beta H}$ for the full Hamiltonian H. If the time (a) is sufficiently far in the past we can just as well take $\rho_a=\text{const}\times\exp(-\beta\sum_{\mathbf{K}}\omega_{\mathbf{K}}a_{\mathbf{K}}^{\dagger}a_{\mathbf{K}})$. That is, we may assume that in the past only the oscillators were in thermal equilibrium at temperature β^{-1}. As a result of the coupling, the entire system will come very quickly to thermal equilibrium at the same temperature. The energy of the single electron and its coupling are infinitesimal (of the order $1/V$) relative to the heat bath of the system of phonon oscillators, so that the exchange of energy between the electron and the lattice will bring everything to thermal equilibrium at the original lattice temperature.

With this choice of ρ_a the dependence of U, U', and ρ_a in (7) on the phonon oscillator coordinates is sufficiently simple so that the oscillator coordinates may be eliminated and the entire expression reduced to a double path integral involving the electrons coordinates only. This reduction, explained in Appendix A, is carried out by methods analogous to those used before by one of the authors on problems in electrodynamics.[9]

The result is (taking $a\to-\infty$, $b\to+\infty$)

$$g=\int\int e^{i\Phi}\mathfrak{D}\mathbf{X}(t)D\mathbf{X}'(t), \tag{12}$$

where

$$\Phi=\int_{-\infty}^{+\infty}\left[\frac{1}{2}\left(\frac{d\mathbf{X}(t)}{dt}\right)^2-\frac{1}{2}\left(\frac{d\mathbf{X}'(t)}{dt}\right)^2\right]dt-\int_{-\infty}^{+\infty}[\mathbf{E}(t)\cdot\mathbf{X}(t)-\mathbf{E}'(t)\cdot\mathbf{X}'(t)]dt$$
$$+\frac{i}{2}\int\frac{d^3K}{(2\pi)^3}|C_{\mathbf{K}}|^2\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\{\exp[i\mathbf{K}\cdot\mathbf{X}(t)]-\exp[i\mathbf{K}\cdot\mathbf{X}'(t)]\}\times[\gamma(\omega_{\mathbf{K}},t-s)$$
$$\times\{\exp[-i\mathbf{K}\cdot\mathbf{X}(s)]+\exp[-i\mathbf{K}\cdot\mathbf{X}'(s)\}+i\alpha(\omega_{\mathbf{K}},t-s)\{\exp[-i\mathbf{K}\cdot\mathbf{X}(s)]-\exp[-i\mathbf{K}\cdot\mathbf{X}'(s)]\}]dsdt. \tag{13}$$

The functions $\gamma(\omega,\tau)$ and $\alpha(\omega,\tau)$ are given in Appendix A. In the special case of Fröhlich's Hamiltonian the integral on $\mathbf{K}$ can be performed to give

$$\Phi_F=\int_{-\infty}^{+\infty}\left[\frac{1}{2}\left(\frac{d\mathbf{X}(t)}{dt}\right)^2-\frac{1}{2}\left(\frac{d\mathbf{X}'(t)}{dt}\right)^2\right]dt+\int_{-\infty}^{+\infty}[\mathbf{E}(t)\cdot\mathbf{X}(t)-\mathbf{E}'(t)\cdot\mathbf{X}'(t)]dt$$
$$+i\alpha 2^{-3/2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[\frac{e^{-i|t-s|}+2P(\beta)\cos(t-s)}{|\mathbf{X}(t)-\mathbf{X}(s)|}+\frac{e^{+i|t-s|}+2P(\beta)\cos(t-s)}{|\mathbf{X}'(t)-\mathbf{X}'(s)|}-\frac{2[e^{-i(t-s)}+2P(\beta)\cos(t-s)]}{|\mathbf{X}'(t)-\mathbf{X}(s)|}\right]dtds, \tag{14}$$

where $P(\beta)=[e^{\beta}-1]^{-1}$.

The double integral $\mathfrak{D}\mathbf{X}(t)\mathfrak{D}\mathbf{X}'(t)$ is only over those paths which satisfy the boundary condition $\mathbf{X}(t)-\mathbf{X}'(t)=0$ at times t approaching $\pm\infty$. The boundary conditions on the paths at large positive or negative times, reflects the arbitrariness of the initial electron state.

Thus, we have reduced the problem of find G via (7), (8), and (9) to that of finding the dependence of a path integral (12) on the forcing functions E and E'. This expression is exact [for the Hamiltonian (11)] but quite complicated. In the next section we discuss approximate methods of evaluation.

III. A METHOD OF APPROXIMATION

In I, a path integral, similar to (14), had to be evaluated. It was argued there that in some rough approximation the "interaction of the charge with itself" represented there by a term in the action function S, $2^{-3/2}\alpha e^{-|t-s|}|\mathbf{X}(t)-\mathbf{X}(s)|^{-1}$, might be imitated by a function S_0 in which this term is replaced by $1/2Ce^{-w|t-s|}[\mathbf{X}(t)-\mathbf{X}(s)]^2$. One may think of the interaction term in S as indicating that at time t the particle acts as though it were in a potential $2^{-1/2}\alpha\int_{-\infty}^{t}e^{-(t-s)}\times|\mathbf{X}(t)-\mathbf{X}(s)|^{-1}ds$ resulting from the electrostatic interaction of the electron with its mean charge density of its previous positions [the weight for different times being $e^{-(t-s)}$]. The assumption then is that such a potential may be roughly replaced by a parabolic potential centered at the mean position of the electron in the past [the weight for different times being $e^{-w(t-s)}$]. In fact, the extra parameter w can be adjusted to compensate partly for the error of using a parabolic potential in place of the true potential form. This argument strongly suggests that the dynamical behavior of the electron, (its motion under an applied electric field) might be described approximately if we replace Φ by a Φ_0, where

$$\Phi_0=\int_{-\infty}^{+\infty}\left[\frac{1}{2}\left(\frac{d\mathbf{X}(t)}{dt}\right)^2-\frac{1}{2}\left(\frac{d\mathbf{X}'(t)}{dt}\right)^2\right]dt-\int_{-\infty}^{+\infty}[\mathbf{E}(t)\cdot\mathbf{X}(t)-\mathbf{E}'(t)\cdot\mathbf{X}'(t)]dt$$
$$-\frac{iC}{2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\{[\mathbf{X}(t)-\mathbf{X}(s)]^2[e^{-iw|t-s|}+2P(\beta w)\cos w(t-s)]$$
$$+[\mathbf{X}'(t)-\mathbf{X}'(s)]^2[e^{+iw|t-s|}+2P(\beta w)\cos w(t-s)]$$
$$-2[\mathbf{X}'(t)-\mathbf{X}(s)]^2[e^{-iw(t-s)}+2P(\beta w)\cos w(t-s)]\}dtds. \quad (15)$$

The parameters C and w are to be determined so as to approximate Φ as closely as possible. At zero temperature ($P=0$), we shall fix C and w at the values given in I. The assumption that Φ_0 is a good approximation to Φ for computing the mobility at low temperatures is based on the supposition that the comparison Lagrangian, which gives a good fit to the ground-state energy at zero temperature, will also give the dynamical behavior of the system. In finding the ground-state energy, the parameters can be chosen by a variational principle but we know of no such principle for the mobility. At finite temperatures the parameters C and w can be determined from a variational principal for the free energy which is a direct extension of the method used in I for the ground-state energy, and reduces to it in the zero-temperature limit. Others[10,10a] have derived in detail the expressions from which the best C and w may be determined for finite β. Thus, C and w can be considered as known functions of α and β even though, unfortunately, no closed analytic form exists, and in any specific calculation they would have to be evaluated numerically.

Actually, we shall not be satisfied merely to replace Φ by Φ_0, but we shall obtain a first correction to $z(\nu)$ by studying, in the next section, the first term in an expansion of $\exp[i(\Phi-\Phi_0)]$:

$$g=\int\int e^{i\Phi}\mathfrak{D}\mathbf{X}\mathfrak{D}\mathbf{X}'\approx\int\int e^{i\Phi_0}\mathfrak{D}\mathbf{X}\mathfrak{D}\mathbf{X}'$$
$$+\int e^{i\Phi_0}i(\Phi-\Phi_0)\mathfrak{D}\mathbf{X}\mathfrak{D}\mathbf{X}'. \quad (16)$$

In this section, however, we will consider only the first term,

$$g_0=\int e^{i\Phi_0}\mathfrak{D}\mathbf{X}\mathfrak{D}\mathbf{X}'. \quad (17)$$

We can expect to evaluate the integral (17) exactly, because the expression for Φ_0 is a quadratic form in $\mathbf{X}(t)\,\mathbf{X}'(t)$ and all such "Gaussian" path integrals can be evaluated exactly.[9] There are several ways to perform the integration in (17). One way is to observe that the expression (15) is obtained by eliminating the variable $\mathbf{Y}$ from a system in which an electron interacts with a single particle described by the Lagrangian

$$L_0=\tfrac{1}{2}(d\mathbf{X}/dt)^2+\tfrac{1}{2}(d\mathbf{Y}/dt)^2-\tfrac{1}{2}k(\mathbf{X}-\mathbf{Y})^2+\mathbf{E}\cdot\mathbf{X}. \quad (18)$$

If we calculate the g for such a system by integrating over all the $\mathbf{Y}$ variables first, then (17) results; provided we choose $k=(v^2-w^2)$ and $M=(v^2-w^2)/w^2$, where $v^2=w^2+4C/w$. However, L_0 can be re-analyzed as the

[10] Y. Osaka, Progr. Theoret. Phys. (Kyoto) 22, 437 (1959).

[10a] M. A. Krivologz and S. I. Pekar, Bull. Acad. Sci. U.S.S.R. 21, 1, 13, 29 (1957).

sum of two normal modes,

$$L_0=\tfrac{1}{2}(M+1)(d\mathbf{X}_2/dt)^2+\mathbf{E}\cdot\mathbf{X}_2+\frac{M}{2(M+1)}\times[(d\mathbf{X}_1/dt)^2-v^2\mathbf{X}_1^2]-\frac{M}{(M+1)}\mathbf{E}\cdot\mathbf{X}_1, \quad (19)$$

so that g_0 can be written as the product of two factors, one for each harmonic oscillator. For a single oscillator of mass m and frequency ω, coupled as $\mathbf{\Gamma}(t)\cdot\mathbf{X}$, the value of g is given in Appendix A. Here we have two oscillators, one of mass $m_1=M+1=v^2/w^2$ and frequency $\omega_1=0$ coupled with a $\mathbf{\Gamma}(t)=\mathbf{E}(t)$, and a second oscillator with $m_2=M/M+1$, $\omega_2=v$, and $\mathbf{\Gamma}(t)=-(M/M+1)\mathbf{E}(t)=-[(v^2-w^2)/v^2]\mathbf{E}(t)$. If the contributions of the two normal modes are combined, g_0 takes the form

$$g_0=\exp\left(\frac{i}{4\pi}\int_{-\infty}^{+\infty}[f(-\nu)-f'(-\nu)]\{[f(\nu)+f'(\nu)]Y_0(\nu)+i[f(\nu)-f'(\nu)]A_0(\nu)\}d\nu\right), \quad (20)$$

where[11]

$$Y_0(\nu)=-(\nu^2-w^2)/(\nu-i\varepsilon)^2[(\nu-i\varepsilon)^2-v^2] \quad (21)$$

and

$$A_0(\nu)=\frac{\pi}{2}\left\{\frac{2w^2}{v^2\beta\varepsilon^2}[\delta(\nu+\varepsilon)+\delta(\nu-\varepsilon)]+[1+2P(\beta v)]\times\frac{(v^2-w^2)}{v^3}[\delta(\nu+v)+\delta(\nu-v)]\right\}. \quad (22)$$

We have expressed $E(t)$ and $E'(t)$ by their Fourier transforms

$$f(\nu)=\int_{-\infty}^{+\infty}E(t)e^{-i\nu t}dt. \quad (23)$$

To obtain $G_0(\tau-\sigma)$ we must evaluate g_0 for E and E' given in (8), that is to say, we must substitute

$$f(\nu)=\epsilon e^{-i\nu\sigma}+\eta e^{i\nu\tau},\qquad f'(\nu)=\epsilon e^{-i\nu\sigma}-\eta e^{i\nu\tau}, \quad (24)$$

into (20), and find the term of order $\epsilon\eta$. Evidently

$$G_0(\tau-\sigma)=\frac{+i}{2\pi}\int_{-\infty}^{+\infty}Y_0(\nu)e^{+i\nu(\tau-\sigma)}d\nu, \quad (25)$$

where $G_0(\tau-\sigma)$ is the zeroth order approximation to $G(\tau-\sigma)$ [Eq. (2)]. Therefore,

$$G_0(\nu)=+iY_0(\nu). \quad (26)$$

Since Y_0 is the classical response function for the comparison system L_0, the result has the immediate interpretation that $G_0(\nu)$ is the response we would have predicted for the system L_0 had we treated it classically. In addition, there is no temperature dependence in G_0 (except through the variation of the parameters v, w with temperature). Both of these well-known results follow from the linearity of L_0.[12,13]

For a particle of mass m, for low frequencies, $G\approx -i/m\nu^2$ so that a comparison of this expression with (21) and (26) gives an effective electron mass $m=v^2/w^2$. This value for m is not the same as the more accurate value given in I, but as Shultz[4] has shown, it is numerically not very different over a wide range of α. Thus, the reactive part of G_0 may be satisfactory, but the dissipative (real part) appearing as it does all at the single frequency v, must be only a very crude approximation.

In the next section we shall compute the corrections implied by the additional expansion terms in (16). We shall find that the mass is now exactly that given in I, and that the dissipatiion has a much more realistic behavior.

IV. FIRST CORRECTION TERM

To evaluate the second term on the right-hand side of (16) we shall have to integrate $e^{i\Phi_0}(\Phi-\Phi_0)$. In order to see what is involved, consider only one of the terms arising from $\Phi e^{i\Phi_0}$:

$$T=\int e^{i\Phi_0}\left\{\int\frac{d^3K}{(2\pi)^3}|C_K|^2\times\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty}\exp\{i\mathbf{K}\cdot[\mathbf{X}(t)-\mathbf{X}(s)]\}\times[\gamma(\omega_K,t-s)+ia(\omega_K,t-s)]dsdt\right\}\times\mathfrak{D}\mathbf{X}(t)\mathfrak{D}\mathbf{X}'(t). \quad (27)$$

Other terms from Φ are similar to (27) with some replacements of $\mathbf{X}$ by $\mathbf{X}'$, while the terms from Φ_0 we will consider later. Evaluation of (27) requires a knowledge of the path integral

$$R(K,t,s)=\int e^{i\Phi_0}\exp\{i\mathbf{K}\cdot[\mathbf{X}(t)-\mathbf{X}(s)]\}\mathfrak{D}\mathbf{X}\mathfrak{D}\mathbf{X}'. \quad (28)$$

Once $R(K,t,s)$ has been evaluated, (27) becomes an ordinary multiple integral:

$$T=\int\frac{d^3K}{(2\pi)^3}|C_K|^2\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty}R(K,t,s)\times[\gamma(\omega_K,t-s)+ia(\omega_K,t-s)]dsdt. \quad (29)$$

[11] ε is a small positive quantity and the limit $\varepsilon\to 0$ is to be taken.

[12] F. L. Vernon, Jr., Ph.D. thesis, California Institute of Technology, 1959 (unpublished).

[13] R. W. Hellwarth, Hughes Research Laboratories, Fourth Quarterly Progress Report, September 15, 1958 (unpublished).

The function R is easily evaluated. It is clearly given by our general formula (20) with

$$\mathbf{f}(\nu)=(\epsilon e^{-i\nu\sigma}+\eta e^{-i\nu\tau})\hat{\imath}+\mathbf{K}(e^{-i\nu t}-e^{-i\nu s}) \tag{30}$$

and[14]

$$\mathbf{f}(\nu)=(\epsilon e^{-i\nu\sigma}-\eta e^{-i\nu\tau})\hat{\imath}. \tag{31}$$

Similarly, a term of the form (27) with $\mathbf{X}(t)$ replaced by $\mathbf{X}'(t)$ can be expressed in terms of our general path integral (20) by using the proper f's,

$$\mathbf{f}(\nu)=(\epsilon e^{-i\nu\sigma}+\eta e^{-i\nu\tau})\hat{\imath}-\mathbf{K}e^{-i\nu s} \tag{32}$$

and

$$\mathbf{f}'(\nu)=(\epsilon e^{-i\nu\sigma}-\eta e^{-i\nu\tau})\hat{\imath}-\mathbf{K}e^{-i\nu t}. \tag{33}$$

In this way, $\int e^{i\Phi_0}\Phi\mathfrak{D}\mathbf{X}\mathfrak{D}\mathbf{X}'$ can be evaluated. Similarly the term $\int e^{i\Phi_0}\Phi_0\mathfrak{D}\mathbf{X}\mathfrak{D}\mathbf{X}'$ may be obtained. To get $\int e^{i\Phi_0}[\mathbf{X}(t)-\mathbf{X}(s)]^2\mathfrak{D}\mathbf{X}\mathfrak{D}\mathbf{X}'$, one can differentiate R with respect to K twice and evaluate it at $K=0$. (Details are given in Appendix B.)

The final result for the first-order change in G is

$$G_1=-iV_0^2(\nu)[\chi(\nu)+(4C/w)\nu^2/(\nu^2-w^2)], \tag{34}$$

where

$$\chi(\nu)=\int_0^\infty [1-e^{i\nu u}]\,\mathrm{Im}S(u)du, \tag{35a}$$

and

$$S(u)=\int \frac{d^3K}{(2\pi)^3}|C_K|^2\frac{2K^2}{3}e^{-K^2D(u)/2}$$
$$[\exp(i\omega_K u)+2P(\beta\omega_K)\cos(\omega_K u)]. \tag{35b}$$

"Im" means the imaginary part and the function $D(u)$ is defined as

$$D(u)=\frac{w^2}{v^2}\left[\frac{v^2-w^2}{w^2v}[1-e^{+iv u}+4P(\beta v)\sin^2(vu/2)]-iu+u^2/\beta\right]. \tag{35c}$$

For Fröhlich's Hamiltonian, the integration over $\mathbf{K}$ may be done to give

$$S(u)=2\alpha/3\sqrt{\pi}\{[D(u)]^{-3/2}[e^{iu}+2P(\beta)\cos u]\}. \tag{36}$$

We have found an approximate form for $G(\nu)$:

$$G(\nu)=G_0(\nu)+G_1(\nu). \tag{37}$$

From it we may find the impedance to first order in $G_1(\nu)$:

$$\nu z(\nu)=1/G(\nu)\sim 1/G_0(\nu)-[1/G_0(\nu)^2]G_1(\nu). \tag{38}$$

The question arises as to whether it is more accurate to expand in this way or to leave the formula as $\nu z(\nu)=(G_0+G_1)^{-1}$. Of course, if G_1 were truly small it would not matter. However, there are excellent reasons to believe that the expanded form is far more accurate. This is best explained by considering a simple example of a free particle to which we add a harmonic binding as a perturbation. The resulting G's are [Φ_0 arising from $\frac{1}{2}m(d\mathbf{X}/dt)^2$ only],

$$G_0(\nu)=+iV_0(\nu)=-i/m\nu^2 \tag{39a}$$

and

$$G_1(\nu)=-i\omega_0^2/m\nu^4, \tag{39b}$$

where ω_0 is the natural frequency of the oscillator. In this case the expanded form of $z(\nu)$ is

$$i\nu z(\nu)=m(\omega_0^2-\nu^2). \tag{40}$$

The true $G(\nu)$ shows a structure (resonance at $\nu=\omega_0$) which is not reflected in an expanded form of G, but which is precisely duplicated (for this linear system) if one expands $z(\nu)$. Therefore, we substitute (34) into (38) to obtain the simple result,

$$-i\nu z(\nu)=\nu^2-\chi(\nu). \tag{41}$$

With $\chi(\nu)$ given by (35), this is our final expression for the impedance of the polaron. It is Eqs. (41) and (35) which we will evaluate in various limits and discuss in the following sections.

The first term on the right-hand side of (41) is a pure free-particle term, while $\chi(\nu)$ contains all of the corrections due to the interaction with phonons. The entire dependence of our results (41) on the trial action Φ_0 is in $D(u)$, Eq. (35c). $D(u)$ in turn appears only in the exponential term in Eq. (35b). This exponential is an effect due to recoil as can be seen by expanding the $\exp(i\mathbf{K}\cdot\mathbf{X})$ term in the Hamiltonian as $1+i\mathbf{K}\cdot\mathbf{X}$ (which we may call a dipole, or linear coupling approximation). If the expansion is made, then the exponential term $e^{-K^2D(u)/2}$ in (35b) will not appear. In other words, if we had any problem in which the field oscillators were coupled linearly to the electron's coordinate $\mathbf{X}$, then our method would give us the exact formula for the impedance irrespective of the choice made for the trial functional Φ_0.[15] This is fact in the best argument for treating the perturbation expansion (16) as an expansion for $z(\nu)$ in the manner of Eq. (38).

Therefore, insofar as the system of phonons behaves as though they were linearly coupled, so there were no recoil effects, (41) is exact. However, recoil effects *are* included in (41); it is only that they are not included precisely. They are approximated by finding their effect for the imitative functional Φ_0 rather than the true functional Φ. For this reason we expect (41) to be an excellent approximation to the true impedance of the polaron.

[14] Equation (20) is a one-dimensional formula. For the case of vector forces the product of two f's is to be interpreted as a dot product.

[15] Assuming that Φ_0 also implies a linear coupling and is a quadratic functional of $\mathbf{X}(t)$, $\mathbf{X}'(t)$.

V. BEHAVIOR OF THE IMPEDANCE

For purposes of further analysis in which we shall change the contour of integration on the variable u, we list some properties of $S(u)$. For real u, $S^*(u)=S(-u)$ and $S(iu)$ is purely real. In addition $S(u)=S(i\beta-u)$ for complex u. The real part of $D(u)$ must be positive in order for the K integral to converge. Therefore, the region of u where this happens, namely the strip parallel to the real axis between the lines $u=$real and $u=$real$+i\beta$, is the region free of singularities for $S(u)$. In the limit of zero temperatures ($\beta\to\infty$) this strip, over which $S(u)$ is analytic, widens to include the entire upper half-plane.

1. Zero Temperature, $\nu<1$; Effective Mass

For Fröhlich's case we consider first the case $\nu<1$, $\beta=\infty$. Then the path of integration [in the integral for $\chi(\nu)$] along the real axis may be rotated to the path along the positive or negative imaginary axis $u\approx 0$ to $\pm i\infty$ (depending on the sign of $e^{i\nu}$). The resulting expression is

$$-i\nu z(\nu)=\nu^2-\frac{2\alpha}{3\sqrt{\pi}}\int_0^\infty e^{-\tau}(1-\cosh\nu\tau)\times\left[\frac{v^2-w^2}{v^3}(1-e^{-v\tau})+\frac{w^2}{v^2}\tau\right]^{-3/2}d\tau. \quad (42)$$

Therefore, z is purely imaginary for $\nu<1$ and there is no dissipation at the absolute zero of temperature (a dc field will continue to accelerate the electrons indefinitely). The reason for this behavior is simply that there are no existing phonons for the electrons to scatter off and none can be created by the electrons until the frequency ν of the applied field is high enough to excite the electrons to a state of energy $\hbar\nu$ higher than the energy $\hbar\omega$ needed to create a phonon. If there is a range of frequencies ω down to zero (as for acoustic modes) then a resistance exists at any frequency of the applied field and at zero temperature. In the Fröhlich model it begins at $\nu=1$. Of course, a dc field will eventually speed the electrons up until they can radiate phonons and dissipate energy. However, this is a nonlinear effect in the applied field strength and is not described by a theory of the impedance. For extremely low frequencies ν we can put $(1-\cosh\nu\tau)=-\nu^2\tau^2/2$. The result is that

$$-i\nu z(\nu)=\nu^2\left\{1+\frac{\alpha}{3\sqrt{\pi}}\int_0^\infty u^2e^{-u}\times\left[\left(\frac{v^2-w^2}{v^3}\right)(1-e^{-vu})+\frac{w^2}{v^2}u\right]^{-3/2}du\right\}. \quad (43)$$

The polaron behaves like a free particle with an effective mass. This mass is the same as the one derived in I, by a modification of the variational ground-state energy calculation.

2. General Expression for Dissipation

The analytic properties of $S(u)$ outlined at the beginning of this section allows one to rewrite the expression for the $\mathrm{Im}\chi(\nu)$ in a form more convenient for computation. We may write (35a) as

$$\mathrm{Im}\chi(\nu)=\mathrm{Im}\int_0^\infty \sin(\nu u)S(u)du. \quad (44)$$

Using the fact that $S(u)$ is analytic between $u=$real and $u=$real$+i\beta$, we may change the contour of integration in (44) from along the real axis to one which goes first from 0 to $i\beta/2$ up the imaginary axis and then from $i\beta/2$ to $i\beta/2+\infty$ parallel to the real axis. (The closing piece of the contour required at infinity does not contribute.) Because $S(iu)$ and $\sin(i\nu u)d(iu)$ are real, the leg of the contour up the imaginary axis contributes nothing to (44), which requires the imaginary part. The contribution from the remaining part of the contour (from $i\beta/2$ to $i\beta/2+\infty$) gives

$$\mathrm{Im}\chi(\nu)=\sinh(\beta\nu/2)\int_0^\infty \cos(\nu u)\Sigma(u)du, \quad (45a)$$

where $\Sigma(u)=S(u+i\beta/2)$ is given by

$$\Sigma(u)=\int\frac{d^3K}{(2\pi)^3}|C_K|^2\times\tfrac{2}{3}K^2\frac{\cos(\omega_K u)\exp[-\frac{1}{2}K^2\Delta(u)]}{\sinh(\beta\omega_K/2)} \quad (45b)$$

and $\Delta(u)\equiv D(u+i\beta/2)$.

$$\Delta(u)=\frac{w^2}{v^2}\left[\left(\frac{v^2-w^2}{w^2v}\right)\frac{\cosh(\beta v/2)-\cos(vu)}{\sinh(\beta v/2)}+\frac{u^2}{\beta}+\frac{\beta}{4}\right]. \quad (45c)$$

The dc mobility μ for the polaron is given by

$$\mu^{-1}=\lim_{\nu\to 0}\mathrm{Im}\chi(\nu)/\nu.$$

Our results (44), therefore, gives

$$\mu^{-1}=\tfrac{1}{2}\beta\int_0^\infty\Sigma(u)du. \quad (46)$$

For the case of Fröhlich's Hamiltonian, we find that

$$\mathrm{Im}\chi(\nu)=\frac{2\alpha}{3\sqrt{\pi}}\frac{\beta^{3/2}\sinh(\beta\nu/2)}{\sinh(\beta/2)}\left(\frac{v}{w}\right)^3\times\int_0^\infty\frac{\cos(\nu u)\cos(u)du}{[u^2+a^2-b\cos(vu)]^{3/2}}, \quad (47a)$$

where

$$a^2\equiv\beta^2/4+R\beta\coth(\beta v/2),\qquad b\equiv R\beta/\sinh(\beta v/2), \quad (47b)$$

and

$$R\equiv(v^2-w^2)/w^2v. \tag{47c}$$

3. Dissipation at Low Temperatures

For low temperatures $e^{-\beta}$ and, therefore, $e^{-\beta v}$ are very much less than one, so that (47a) may be expanded as a power series in b.

$$\mathrm{Im}\chi(\nu)=\frac{2\alpha}{3\sqrt{\pi}}\frac{\sinh(\beta\nu/2)}{\sinh(\beta/2)}\left(\frac{v}{w}\right)^3\int_0^\infty\frac{\cos(\nu u)\cos u}{(u^2+a^2)^{3/2}}\times\left\{1+\frac{3\beta Re^{-\beta v/2}\cos(vu)}{(u^2+a^2)}+\frac{15}{4}\frac{\beta^2R^2e^{-\beta v}}{(u^2+a^2)^2}[1+\cos(2vu)]+\cdots\right\}du. \tag{48}$$

Now an integral like $\int\cos(\lambda u)du/(u^2+a^2)^{3/2}$ falls off exponentially like $(2\pi|\lambda|)^{1/2}(a)^{-3/2}e^{-|\lambda|a}$ as λ increases. Thus, the smallest values of λ count, and these count with the smallest power $e^{-\beta}$ in front. This permits us to select the important terms for each ν. For example, the last term in brackets contributes when $\nu\approx 2v+1$ for there is a contribution from $\cos(\nu u)\cos(u)\cos(2vu)$. The $e^{-\beta v}$ is compensated for by the $e^{\beta(\nu-1)/2}$ in front.

For $\nu<1+v$, and $|1-\nu|>\beta^{-1}$, only the first term in the expansion contributes and we obtain

$$\mathrm{Im}\chi(\nu)\approx\tfrac{2}{3}\alpha\left(\frac{v}{w}\right)^3(1-e^{-\beta\nu})e^{\beta(\nu-1)/2}\times[(|\nu-1|)^{1/2}e^{-\beta|\nu-1|/2}e^{-R|\nu-1|}+(\nu+1)^{1/2}e^{-\beta(\nu+1)/2}e^{-R(\nu+1)}]. \tag{49}$$

As $\beta\to\infty$ this shows a threshold at $\nu=1$. Below $\nu=1$ the result is nearly zero; above, it is $\sim\frac{2}{3}\alpha(v/w)^3(\nu-1)^{1/2}\times\exp[-R(\nu-1)]$. This is the threshold to create one optical phonon from the energy quantum $\hbar\nu$ supplied by the external field. If we have an excess energy $\nu-1$, the final electron has momentum proportional to $(\nu-1)^{1/2}$ and this appears as a factor because of the phase space available. The cross section depends in some way on the frequency above threshold; the factor $\exp[-R(\nu-1)]$ is a rough approximation to this, generated by our model Φ_0.

To study the dependence on β in a little more detail, in case $\nu<1$, (49) can be rewritten as

$$\mathrm{Im}\chi(\nu)\approx\tfrac{2}{3}\alpha(v/w)^3[(1-\nu)^{1/2}e^{-\beta(1-\nu)}e^{-R(1-\nu)}+(1+\nu)^{1/2}e^{-R(1+\nu)}e^{-\beta}-(1-\nu)^{1/2}e^{-R(1-\nu)}e^{-\beta}-(1+\nu)^{1/2}e^{-R(1+\nu)}e^{-\beta(1+\nu)}]. \tag{50}$$

The terms are easily understood. In the first, an electron absorbs a quantum of energy ν to emit a phonon with energy one. The chance that the electron has enough energy, $1-\nu$, to do this is $e^{-\beta(1-\nu)}$. In the second, the electron absorbs a quantum ν and also absorbs a phonon (they are present in number $e^{-\beta}$) so the outgoing electron momentum is $(1+\nu)^{1/2}$. The third corresponds to the electron absorbing a phonon and emitting a quantum $\hbar\nu$ to the electric field. This emission contributes negatively to the energy loss of the electric field. The fourth term results from a particularly energetic electron of energy exceeding $1+\nu$ radiating a phonon and emitting a quantum to the field.

Since the dc mobility is given by the $\mathrm{Im}\chi(\nu)/\nu$ as $\nu\to 0$ (50) gives an expression at low temperatures for the dc mobility of the Fröhlich model. Using our trial functional Φ_0, we find that

$$\mu=\left(\frac{w}{v}\right)^3(3/4\alpha\beta)e^{\beta}\,e^{(v^2-w^2)/w^2v}. \tag{51}$$

The dependence of (51) on the coupling strength α is as α^{-1} for small α (and high β) because the "best" $w\approx v$ for small α (see I). This dependence on α is of course the same as is derived by perturbation theory. As α becomes large ($\alpha\gg 1$) the best parameters satisfy the relation $v/w\sim\alpha^2$. Therefore, the mobility μ becomes proportional to $\alpha^{-7}e^{\alpha^2}$ at high coupling strengths. The result (51) *can not* be compared directly with the results of previous calculations[4,5,16,17] because its temperature dependence is e^{β}/β rather than the e^{β} dependence found in the other approaches. At high β (where these previous calculations are valid), the different dependence on temperature would be experimentally unobservable. However, the origin and significance of this (incorrect) temperature dependence is interesting and will be discussed at some length in a later section.

Returning to our general expression for the $\mathrm{Im}\chi(\nu)$ (48), we see that there are other thresholds at higher frequencies coming from higher terms in the sum (48). For $\beta=\infty$ the next threshold is at $\nu=v+1$, the contribution above threshold being $(2\alpha/3)(v/w)^3(\nu-v-1)^{3/2}\times\exp[-R(\nu-v-1)]$. These higher thresholds correspond to exciting the electron to an excited state of energy v and emitting a phonon. The position of this excited state (at v) and the higher ones along with the selection rule that says these cannot be excited without the emission of a p-state phonon is a fiction supplied by our imitating action Φ_0.

Of course, for strong couplings, there will be such complicated excited states, with partial selection rules leading to a complex curve for $\mathrm{Im}\chi(\nu)$. For strong electron-phonon coupling, the electron is in effect bound in a potential which it makes by distorting the lattice in its neighborhood. If the lattice were held fixed in this distorted state, we would expect the electron to have various excited states in this potential. In fact, the lattice moves so that "states" are unstable, but for large α the excitation energies are of order α^2 larger than the lattice frequency so it cannot follow quickly enough.

[16] F. Low and D. Pines, Phys. Rev. 98, 414 (1955).
[17] Y. Osaka, Progr. Theoret. Phys. (Kyoto) 25, 4, 517 (1961).

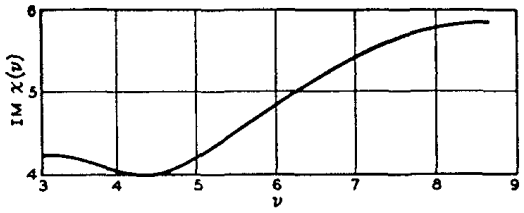


FIG. 1. Plot of Im $\chi(\nu)$ for $\alpha=3$, $w=2.5$, $v=3.4$ and $\beta=100$. ν is plotted in units of the optical mode frequency ω and Im $\chi(\nu)$ is plotted in units of (m/ω).

The Im$\chi(\nu)$ should have a maximum when ν is equal to a frequency which can be absorbed in going to such an excited "state." The widths of these maxima reflect the lifetimes of these "states" for phonon emission. Naturally, we cannot expect our approximate formula (35a) to give such detailed results correctly.[18] It is reassuring, however, that our method gives such a realistic looking behavior, and strongly suggests that it represents a long step forward toward the correct Im$\chi(\nu)$. In the last section we outline some ways of improving the Im$\chi(\nu)$. When the coupling is not too strong ($\alpha\approx 3$), these thresholds are weak and hard to see, and the curves will have a "washed out" appearance. In Figs. 1–3 are given curves of Im$\chi(\nu)$ vs ν for $\alpha=3$, $\alpha=5$, $\alpha=7$ and for low temperatures ($\beta=100$). These figures show the resonance effects very nicely. As a functon of frequency each curve consists of maxima of increasing width. As a function of α the curves show more maxima of decreasing width as α increases. All of the curves were computed numerically on an IBM 7090 computer, using an infinite power series expansion of (47a) in terms of K functions (Bessel functions of imaginary argument). The values of the parameters used were $w=2.5$, $v=3.4$ for $\alpha=3$, $w=2.1$, $v=4.0$ for $\alpha=5$, and $w=1.6$, $v=5.8$ for $\alpha=7$.

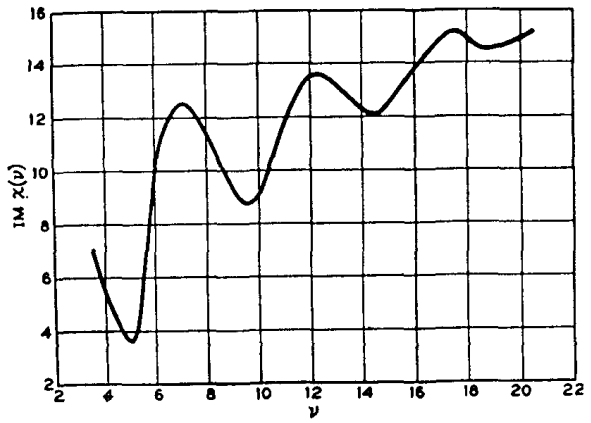


FIG. 2. Plot of Im $\chi(\nu)$ for $\alpha=5$, $w=2.1$, $v=4.0$, and $\beta=100$. ν is plotted in units of the optical mode frequency ω and Im $\chi(\nu)$ is plotted in units of (m/ω).

[18] Pekar, in the strong-coupling limit, using Gaussian trial wave functions for the electron, finds an excited state at precisely v.

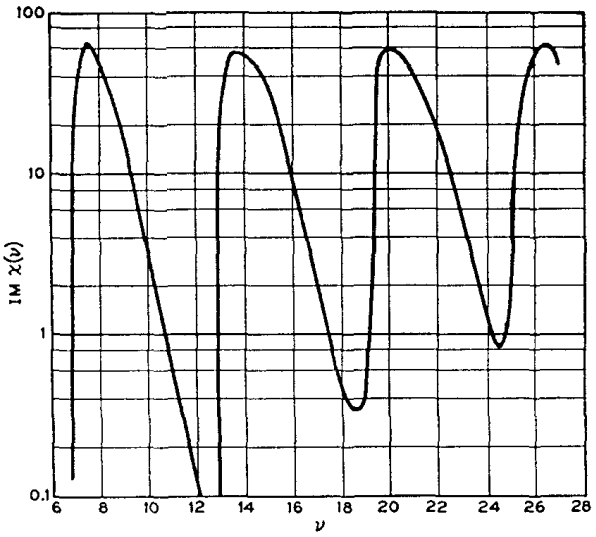


FIG. 3. Plot of Im $\chi(\nu)$ for $\alpha=7$, $w=1.6$, $v=5.8$, and $\beta=100$. ν is plotted in units of the optical mode frequency ω and Im $\chi(\nu)$ is plotted in units of (m/ω).

4. Behavior at High Temperatures

For high temperatures, β is small and (disregard for a moment the variation of w/v with temperature) $\Delta(u)$ varies like u^2/β. Therefore, only small u will be of importance in the exponent, and we can expand $\Delta(u)$ (45c) as

$$\Delta(u)=1/\beta[(u^2+\beta^2/4) -(v^2-w^2/12)(u^2+\beta^2/4)^2+\cdots]. \quad (52)$$

The leading term is, therefore, that of perturbation theory, and our formula is insensitive to the trial functional Φ_0. Actually this is even more accurate than it appears, because as T rises the parameters v and w change in just such a way that v^2-w^2 falls making the approximation (52) still better.[10,10a]

At very high temperatures the perturbation theory works because the electron has on the average, an energy high compared to the lattice frequency. In this case the effective polarization should fall to zero in the limit of infinite temperature. We, therefore, expect that the accuracy of our formula will increase as the temperature rises.

VI. WEAK-COUPLING LIMIT; THE BOLTZMANN EQUATION

In the limit of weak coupling (C_K small or α small) the "best" model parameters are $v\approx w$ or $C\approx 0$. That is, the model of (15) becomes the bare free electron. Perturbation theory is also simply the expansion (16) but with Φ_0 just the free-electron influence functional [i.e., $D(u)=-iu+u^2/\beta$]. Therefore, our series $G_0+G_1+\cdots$ for the admittance agrees with perturbation theory in the weak-coupling limit.

One would expect then, that in the limit of weak coupling the mobility in constant fields obtained here

would be exactly that obtained from the Boltzmann equation using perturbation theory for the electron-phonon scattering cross sections. This turns out to be not true; the question being whether the limit $\nu \to 0$ is taken before or after the limit of weak coupling $|C_{\mathbf{K}}|^2 \to 0$ is taken. A detailed comparison of our result to that of the Boltzmann equation for weak coupling is instructive.

The Boltzmann equation for a particle with momentum distribution $f(\mathbf{P})$ in an electric field $\mathbf{E}(t)$ in the x direction is

$$\partial f/\partial t + E\partial f/\partial P_x = -\int [\gamma(\mathbf{P}\to\mathbf{P}')f(\mathbf{P}) - \gamma(\mathbf{P}'\to\mathbf{P})f(\mathbf{P}')]d^3P'/(2\pi)^3. \quad (53)$$

In (53), $\gamma(\mathbf{P}\to\mathbf{P}')$ is the probability per second that an electron of momentum $\mathbf{P}$ is scattered to momentum $\mathbf{P}'$ by collisions with the phonon gas. This rate, using the usual perturbation theory, is

$$\gamma(\mathbf{P}\to\mathbf{P}') = 2\pi|C_{\mathbf{K}}|^2[(1-e^{-\beta\omega_{\mathbf{K}}})^{-1}\delta(\tfrac{1}{2}P'^2-\tfrac{1}{2}P^2+\omega_{\mathbf{K}}) + (e^{\beta\omega_{\mathbf{K}}}-1)^{-1}\delta(\tfrac{1}{2}P'^2-\tfrac{1}{2}P^2-\omega_{\mathbf{K}})]. \quad (54)$$

The first term in (54) is just the probability to emit a phonon of momentum $\mathbf{K}=\mathbf{P}'-\mathbf{P}$. The second term is the probability to absorb a phonon of momentum $\mathbf{K}$.

The Maxwell distribution, $f_0(P)=(2\pi/\beta)^{-3/2}e^{-\beta P^2/2}$, is a solution of (54) with no electric field since

$$\gamma(\mathbf{P}\to\mathbf{P}')f_0(P) = \gamma(\mathbf{P}'\to\mathbf{P})f_0(P'). \quad (55)$$

If E is a very weak, spatially uniform field varying as $E=E_0e^{i\nu t}$, the deviation of f from equilibrium can be written as $f=f_0[1+E_0e^{i\nu t}P_xh(\mathbf{P})]$, where $h(\mathbf{P})$ is a function of P^2 satisfying [using (54)]

$$(i\nu h+\beta)P_x = -\int \gamma(\mathbf{P}\to\mathbf{P}')[P_xh(\mathbf{P})-P_x'h(\mathbf{P}')]d^3P'/(2\pi)^3. \quad (56)$$

The current j is $-E_0e^{i\nu t}\int P_x^2h(\mathbf{P})f_0(P)d^3P$, so the impedance is given by $1/z(\nu)=-\int P_x^2h(\mathbf{P})f_0(P)d^3P$. Multiplying (56) by $P_xh(\mathbf{P})f_0(P)$ and integrating over all $\mathbf{P}$, we find another expression for the impedance:

$$z(\nu)=\frac{\frac{1}{2}\int\int f_0(P)\gamma(\mathbf{P}\to\mathbf{P}')[P_xh(\mathbf{P})-P_x'h(\mathbf{P}')]^2d^3P'd^3P+i\nu\int P_x^2h^2(\mathbf{P})f_0(P)d^3P\,(2\pi)^3}{(2\pi)^3\beta\left[\int P_x^2h(\mathbf{P})f_0(P)d^3P\right]^2} \quad (57)$$

The integral equation is quite difficult to solve in general. We will content ourselves with an approximate analysis. The expression (57) has been written so that it is stationary[19] for variation of h about the true solution (56). That is, errors in h will appear only in second order in $z(\nu)$.

The simplest approximate solution to (56) is that $h(\mathbf{P})$ is a constant. Then (57) gives

$$z(\nu)-i\nu=\Gamma, \quad (58)$$

where

$$\Gamma=\frac{\beta}{2}\int\int f_0(P)\gamma(\mathbf{P}\to\mathbf{P}')(P_x-P_x')^2d^3Pd^3P'/(2\pi)^3. \quad (59)$$

This represents a frequency-independent pure resistance Γ. Thus, $i\nu\Gamma$ should be compared to the $\chi(\nu)$ of (35a). To do so we substitute (54) into (59). One sees that the first term in brackets in (54) gives the same contribution as the second, so calling $\mathbf{P}'=\mathbf{P}\approx\mathbf{K}$ we get (replacing K_x^2 by $K^2/3$)

$$\Gamma=\frac{\beta}{3(2\pi)^2}\int\int K^2|C_{\mathbf{K}}|^2P(\omega_{\mathbf{K}})f_0(P) \times\delta[\tfrac{1}{2}(\mathbf{P}-\mathbf{K})^2-\tfrac{1}{2}P^2-\omega_{\mathbf{K}}]d^3Pd^3K. \quad (60)$$

Next we replace the δ function by $\delta(x)=\int_{-\infty}^{\infty}e^{iux}du/2\pi$. The $\mathbf{P}$ integral is then readily evaluated to give

$$\Gamma=\beta\int d^3K\,(2\pi)^{-3}|C_{\mathbf{K}}|^2\tfrac{2}{3}K^2\int_{-\infty}^{\infty}du\,P(\omega_{\mathbf{K}}) \times\exp[-iu\omega_{\mathbf{K}}-\tfrac{1}{2}K^2(-iu+u^2/\beta)]. \quad (61)$$

Replacement here of u by $u+i\beta/2$ shows that $\Gamma=\mu^{-1}$ from (45a) if $\Delta(u)$ takes on its free-particle value, $u^2/\beta+\beta/4$.

This result is what we expected but there are two points to be made. Firstly, we made no assumption that ν was small in solving the Boltzmann equation. Why then do we not find that $\Gamma=\mathrm{Im}\chi(\nu)/\nu$ as a function of frequency instead of a constant, the limit as $\nu\to 0$ of $\mathrm{Im}\chi(\nu)/\nu$? The answer does not lie in trying to solve (56) more exactly, for if we take a high-frequency case so

[19] If $\nu=0$ (57) is a minimum, giving the useful variational principle for μ^{-1} discussed by Wilson [A. Wilson, *Theory of Metals* (Cambridge University Press, New York, 1959), p. 300].

that the collision term in (56) is negligible compared to the $i\nu$ term, in first approximation $h=(\beta/i\nu)$ and is indeed constant as we assumed. That is, (58) is a *closer* approximation to the prediction of (56) the *higher* the value of ν relative to Γ. The answer is that the original formulation of the Boltzmann equation is faulty at high frequencies. It is assumed that the collisions are made and that between collisions the particle drifts in the electric field. But at higher frequencies new processes are possible in which, for example, the electron absorbs a quantum $\hbar\nu$ from the field and radiates a phonon. In the quantum theory for higher ν this cannot be analyzed as the succession of two independent events. Therefore, at the higher frequencies we may use our formulas (35). If the results deviate from that of the Boltzmann equation we must conclude the latter is inaccurate.

The second point to discuss is this. We did not solve the Boltzmann equation exactly; presumably, therefore, Γ is not exact. Why then is our result μ^{-1} from (51) not asymptotically exact as $\alpha\rightarrow 0$, in spite of our argument that our formulas should be correct in perturbation theory? The reason is easy to see from (57). For *finite* ν, as the coupling gets weaker the collision term falls below ν and Γ is in fact exact. Thus, for any ν *other than zero* in the limit of infinitesimal coupling our result (45) is *exact*. However, for $\nu=0$, to get the exact answer no matter how small the coupling, the full Boltzmann equation must be solved and our result for μ, (45), is only an approximation. [Mathematically, the lack of uniform convergence arises when we invert G and expand, because the resistive part of $\chi(\nu)$ exceeds the leading (reactive) term ν^2 no matter how small the coupling is if $\nu=0$.]

Although not exact, our result (45) for μ^{-1} is still a good approximation to the solution of the Boltzmann equation. The value of μ^{-1} obtained for Fröhlich's model at low temperature from (51) varies as $\beta^{-1}e^{-\beta}$, while from the Boltzmann equation we know that it should vary as a constant times $e^{-\beta}$.[5] But because of the rapid variation of the exponent these two are hard to distinguish (for example, the temperature at which the mobility reaches a given value is imperceptibly different in the two cases). At higher temperatures our results for μ^{-1} (46) no longer behaves as $e^{-\beta}/\beta$. In this case the values obtained from (46) and the Boltzamnn equation would come closer together. (At extremely low temperatures Fröhlich's model, of course, fails. Although acoustic phonons are not very effective, they cannot be disregarded for there are virtually no optical phonons excited.) As a test of our approximate solution of the Boltzmann equation we have also analyzed a system interacting at high temperature with acoustic phonons with $|C_{\mathbf{K}}|^2$ proportional to K^2, $\omega_{\mathbf{K}}=Kv_0$, and $\beta^{-1}>mv_0^2$. Such a coupling leads to a relaxation time for the electrons which varies inversely with their velocity. In this case μ is proportional to β in either theory, but Eq. (46) gives a result $32/9\pi$ or 13% higher than the more accurate solution of (56) given by (59).

VII. SUGGESTIONS FOR IMPROVING ACCURACY

The entire dependence of our result (35) for the polaron impedance on the imitating quadratic influence functional Φ_0 is contained in $D(u)$ which is expressed in terms of the model's response function, $Y_0(\nu)$. If in the expansion we used a more elaborate (but still quadratic) Φ_0, then all results would be the same but with a $D(u)$ coming from a different $Y_0(\nu)$ in (35). What is the best $Y_0(\nu)$ to take? We note that $Y_0(\nu)$ was also the first approximation to the desired function $Y(\nu)=[i\nu z(\nu)]^{-1}$. A natural suggestion, therefore, is that the best $Y_0(\nu)$ is the "true admittance function of the real polaron". Since the true $Y(\nu)$ is unknown, perhaps the next best alternative would be to use a $Y_0(\nu)$ in (35) such that the $z(\nu)$ itself equals $[i\nu Y_0(\nu)]^{-1}$; that is to use a $Y_0(\nu)$ which satisfies (35) self-consistently.

To find a self-consistent $Y_0(\nu)$ is not, in general, easy. However, one might use the results calculated here, for example, and re-insert them as a new $Y_0(\nu)$ in (35) and recalculate $z(\nu)$ to find a second iteration, which might provide an even better impedance with which to recalculate again if necessary. Aside from the great amount of work involved and questions of convergence of the procedure, we cannot even be sure if a substantial improvement would result; however, the following observations do suggest that a self-consistent solution of (35) could result in a considerable increase in accuracy.

In the variational principle of I, one can try a trial action functional S_0 which is more general than the two parameters one employed there and which describes an electron coupled to a general linear system,

$$S_0=\frac{1}{2}\int\left(\frac{d\mathbf{X}(t)}{dt}\right)^2dt+\frac{1}{2}\int h(t-s)x(t)x(s)dtds. \quad (62)$$

Then, putting $h(t)=\int\mathbf{h}(\nu)e^{i\nu t}d\nu/2\pi$, one finds that the function $\mathbf{h}(\nu)$ which gives the lowest energy in the variational principal at zero temperature, satisfies the integral equation

$$\mathbf{h}(\mu)=\frac{2}{3(2\pi)^3}\int\int K^2|C_{\mathbf{K}}|^2d^3K(1-\cos\mu\tau)\times e^{-\omega_{\mathbf{K}}\tau}e^{-\frac{1}{2}K^2\varphi(\tau)}d\tau, \quad (63)$$

where

$$\varphi(\tau)=\int^{\infty}2(1-\cos\mu\tau)\frac{1}{-\mu^2+\mathbf{h}(\mu)}\frac{d\mu}{(2\pi)}. \quad (64)$$

The Eqs. (35) treated as a self-consistent set with $D(u)$ generated from $Y(\nu)=[\nu^2+\chi(\nu)]^{-1}$ can be transformed for $\beta=0$ at least, to exactly this same pair of Eqs. (63) and (64). In this case $\chi(-i\mu)$ replaces $\mathbf{h}(\mu)$. We therefore, can conclude (for $\beta=0$) that if one tries as a trial action functional that for an electron coupled to a general linear system, no such system will produce a better result than one which has an impedance which is

the self-consistent solution of (35).[20] (Possibly the same is true for arbitrary β, but we have not checked this point.) These considerations substantiate further the interpretation of the expansion (38) in terms of an impedance rather than an admittance.

The above point, and the existence of a minimum principle for the Boltzmann equation, suggest that some minimum principle exists for the mobility $\chi(-i\mu)$ in quantum mechanics at arbitrary β.

Another way to improve accuracy is to try to include the next term $(\Phi-\Phi_0)^2$ in the expansion. The integrals to be performed still are of the form required for the evaluation of the $(\Phi-\Phi_0)$ term but with more complicated driving forces E and E'. Although the calculation would proceed straightforwardly it would, in general, be very laborious. However, for certain values of ν and β, one could do what amounts to the same thing in a somewhat easier way, For example, for $\beta^{-1}<\nu<1$ the various terms in (50) could be improved by calculating the appropriate cross sections more accurately. The cross section to absorb a quantum from the electric field and emit a single phonon requires matrix elements of quantities like $x_t \exp(i\mathbf{K}\cdot\mathbf{X}_s)$. Equation (50) corresponds to calculating these with the propagator $\exp(i\Phi_0)$, but an improvement can be made by calculating them with the propagator

$$[1+i(\Phi-\Phi_0)]\exp(i\Phi_0).$$

For zero frequency the Boltzmann equation can be used with the rates $\gamma(\mathbf{P}'\to\mathbf{P})$ calculated with the propagator $\exp(i\Phi_0)$; further improvement would again result by adding a correction for the difference of Φ and Φ_0.

ACKNOWLEDGMENTS

The authors would like to thank W. G. Wagner, Morrel Cohen, J. McKenna, and M. Lax for many helpful and stimulating discussions.

APPENDIX A

In this appendix we discuss a fundamental path integral (or trace) such as (7) for a single one-dimensional oscillator, in terms of which all other path integrals can be immediately evaluated.

The oscillator Hamiltonian $H=p^2/2+\omega^2q^2/2$. Let

$$U=\exp\left\{-i\int_{-\infty}^{+\infty}[H_t+\gamma(t)q_t]dt\right\},$$

$$U'=\exp\left\{i\int_{-\infty}^{+\infty}[H_t'+\gamma'(t)q'_t]dt\right\},$$

and $\rho_a=e^{-\beta H}/Q$. The $\mathrm{Tr}(U\rho_aU'^{-1})\equiv g$ may be done in several ways, for example by writing it in terms of path integrals and performing the resulting Gaussian integrals.[12] As an alternative method we choose to represent the trace in free oscillator eigenfunctions. In this representation $(\rho_a)_{n,n}=e^{-n\beta\omega}(1-e^{-\beta\omega})$ and

$$g\equiv\sum_{m,n}G_{m,n}G_{m,n}{}^{*\prime}e^{-\beta\omega n}(1-e^{-\beta\omega}), \tag{A1}$$

where $G_{m,n}$ is given in reference (9), Eq. (38) and $G_{m,n}{}'$ is the same expression with γ' replaced by γ.

By expanding

$$\exp[(x+i\xi)(y+i\xi^*)]=\sum_t(x+i\xi)^t(y+i\xi^*)^t/t!$$

in powers of x and y, one can show that $G_{m,n}$ may also be written as

$$G_{m,n}=G_{0,0}(m!n!)^{-1/2}e^{\xi\xi^*}\sum_t\frac{t!(i\xi)^{t-m}(i\xi^*)^{t-n}}{(t-m)!(t-n)!}, \tag{A2}$$

where

$$\xi=i(2\omega)^{-1/2}\int_{-\infty}^{+\infty}e^{i\omega t}\gamma(t)dt. \tag{A3}$$

The summation over m and n in our expression for g can be done by the binomial theorem if one uses this expression for $G_{m,n}$; but the one in reference (4), Eq. (38), for $G_{m,n}{}'$, calling $t=r+v$, where r is the free index in $G_{m,n}{}'$, permits summing first over r then over v. The final result is

$$g=G_{00}G_{00}{}'e^{\xi\xi^*}\times\exp[(1-e^{-\beta\omega})^{-1}(i\xi-i\xi')(i\xi^*-i\xi'^*e^{-\beta\omega})]. \tag{A4}$$

Substituting in for G_{00} and $G_{00}{}'$, we find that

$$g=\exp\left(\tfrac{1}{2}i\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[\gamma(t)-\gamma'(t)]\{[\gamma(s)+\gamma'(s)]\times y(\omega,t-s)+i[\gamma(s)-\gamma'(s)]a(\omega,t-s)\}dtds\right), \tag{A5}$$

where

$$\begin{aligned} y(\omega,t-s)&=(1/\omega)\sin\omega(t-s), && t>s\\ &=0, && t<s\end{aligned} \tag{A6}$$

and

$$a(\omega,t-s)=(1/2\omega)\cos\omega(t-s)[1+2P(\beta\omega)]. \tag{A7}$$

If $\gamma(\nu)$, $\gamma'(\nu)$, $y(\nu)$, $a(\nu)$ are the Fourier transforms of $\gamma(t)$, $\gamma'(t)$, etc., so that, for example,

$$\gamma(t)=\int\gamma(\nu)e^{i\nu t}d\nu/2\pi, \tag{A8}$$

then the expression for g can be written in Fourier space as

$$g=\exp\left(\frac{i}{4\pi}\int_{-\infty}^{+\infty}[\gamma(-\nu)-\gamma'(-\nu)]\{[\gamma(\nu)+\gamma'(\nu)]\times y(\omega,\nu)+i[\gamma(\nu)-\gamma'(\nu)]a(\omega,\nu)\}d\nu\right). \tag{A9}$$

[20] One probably cannot get $\chi(\nu)$ from (63) and (64) by iteration for one finds $\chi(-i\mu)$ only approximately this way and one *cannot* pass to an accurate value on the real line from an imperfect knowledge on the negative imaginary axis alone.

Here,[11]

$$y(\omega,\nu)=1/(\nu-i\varepsilon)^2-\omega^2 \tag{A10}$$

and

$$a(\omega,\nu)=(\pi/2\omega)[1+2P(\beta\omega)][\delta(\nu+\omega)+\delta(\nu-\omega)]. \tag{A11}$$

For the case of the Hamiltonian (11), we can represent the particle motion by a path integral on $\mathbf{X}(t)$. Then we have in fact a large number of independent oscillators, each coupled to the particle. The $\mathbf{K}$ mode of frequency $\omega_{\mathbf{K}}$ is coupled by a $\gamma(t)=C_{\mathbf{K}}\exp[i\mathbf{K}\cdot\mathbf{X}(t)]$. Each of these modes contributes a factor like (A9) to g so that the final exponent is a sum of contributions from each oscillator mode.

We shall need the functions (A10), (A11) and superpositions of them (sums for various frequencies) which we call $Y(\nu)$, $A(\nu)$. We shall also need another function $D(u)$ defined as

$$D(u)=\pi^{-1}\int_{-\infty}^{+\infty}\{\sin(\nu u)Y(\nu)+[1-\cos(\nu u)]A(\nu)\}d\nu, \tag{A12}$$

and $\Delta(u)$ defined as $D(u+i\beta/2)$. All these functions are related to $Y(\nu)$, in fact to its imaginary part $\mathrm{Im}Y(\nu)$. We need it only for $\nu>0$, since $\mathrm{Im}Y(-\nu)=-\mathrm{Im}Y(\nu)$. For a single oscillator, from (A5) as we have $\mathrm{Im}y(\nu)=-(\pi/2\omega)[\delta(\nu-\omega)-\delta(\nu+\omega)]$. But $a(\nu)$ can also be written as

$$\frac{\pi}{2\omega}\left(\frac{e^{\beta\nu}+1}{e^{\beta\nu}-1}\delta(\nu-\omega)+\frac{e^{-\beta\nu}+1}{e^{-\beta\nu}-1}\delta(\nu+\omega)\right);$$

hence, $a(\nu)=-\coth(\beta\nu/2)\,\mathrm{Im}y(\nu)$. Further, since the poles of a general $Y(\nu)$ lie above the real axis, the real part of $Y(\nu)$ can be obtained from the imaginary part. Proceeding in this way, we find the following expressions for all the functions in terms of $\mathrm{Im}Y(\nu)$:

$$Y(\nu)=\frac{2}{\pi}\int_0^\infty\frac{\mathrm{Im}Y(\mu)\mu d\mu}{(\nu-i\varepsilon)^2-\mu^2} \tag{A13}$$

$$\mathrm{A}(\nu)=-\left(\frac{e^{\beta\nu}-1}{e^{\beta\nu}+1}\right)\mathrm{Im}Y(\nu), \tag{A41}$$

$$D(u)=-\frac{2}{\pi}\int_0^\infty\left(1-e^{i\nu u}+\frac{2(1-\cos\nu u)}{e^{\beta\nu}-1}\right)\mathrm{Im}Y(\nu)d\nu, \tag{A15}$$

and

$$\Delta(u)=-\frac{2}{\pi}\int\frac{[\cosh(\beta\nu/2)-\cos(\nu u)]}{\sinh(\beta\nu/2)}\mathrm{Im}Y(\nu)d\nu. \tag{A16}$$

Although derived for a single oscillator, these relations are linear and hold for any superposition of oscillators.

APPENDIX B

We give here the details of the calculation of the first order correction to G_1. As explained in the text, we can calculate an expression like (28) by substituting (30) and (31) into (20). If we write

$$g^0=\exp\left[\frac{i}{\pi}\int\eta e^{i\nu\tau}(\epsilon e^{-i\nu\sigma}Y_0+i\eta A_0e^{-i\nu\tau})d\nu\right]$$

for the expression (20) calculated for $K=0$, the result (20) with K included as in (30) and (31) is

$$\begin{aligned}R(K,t,s)&=g^0\exp\left\{\frac{iK^2}{4\pi}\int|e^{i\nu t}-e^{i\nu s}|^2(Y_0+iA_0)d\nu\right.\\&+\frac{iK_x}{2\pi}\int(e^{i\nu t}-e^{i\nu s})(\epsilon e^{-i\nu\sigma}Y_0+i\eta A_0e^{-i\nu\tau})d\nu\\&\left.+\frac{iK_x}{2\pi}\int(e^{-i\nu t}-e^{-i\nu s})\eta e^{i\nu\tau}(Y_0+iA_0)d\nu\right\}.\end{aligned} \tag{B1}$$

We shall ultimately only need the result to the first order in $\epsilon\eta$ [see (9)], so differentiating R with respect to η and ϵ, putting $\epsilon=\eta=0$ and calling the result $2r(K,t,s)$, we get (there is a 1/3 for averaging over directions of $\mathbf{K}$)[14]

$$\begin{aligned}r(K,t,s)=\left\{\right.&\frac{i}{2\pi}\int Y_0e^{i\nu(\tau-\sigma)}d\nu\\&-\frac{K^2}{3(8\pi^2)}\int(e^{i\nu t}-e^{i\nu s})e^{-i\nu\sigma}Y_0d\nu\\&\times\int(e^{i\mu t}-e^{i\mu s})iA_0(\mu)e^{-i\mu\tau}\\&\left.+(e^{i\mu t}-e^{-i\mu s})e^{i\mu\tau}[Y_0(\mu)+iA_0(\mu)]d\mu\right\}\\&\times\exp\left[\frac{iK^2}{4\pi}\int|e^{i\nu t}-e^{i\nu s}|^2(Y_0+iA_0)d\nu\right].\end{aligned} \tag{B2}$$

The first term is $G_0(\tau-\sigma)$ times the path integral with $\epsilon=\eta=0$. Such a term arises no matter what we integrate, so in total it gives $G_0(\tau-\sigma)\int e^{i\Phi_0}(\Phi-\Phi_0)\mathfrak{D}\mathbf{X}\mathfrak{D}\mathbf{X}'$. This term just cancels when we remember that we must divide $(\partial^2g/\partial\eta\partial\epsilon)$ by g evaluated at $\epsilon=\eta=0$ for normalization. Therefore, this does not contribute to G_1, the first correction of G from G_0, and we omit it. The other terms are later to be multiplied by a function of $t-s=u$ only, and integrated on t and s. Hence, we let $t=u+s$ and integrate on all s's to get [note $Y_0(-\nu)=Y_0^*(\nu)$,

$A(-\nu)=A_0(\nu)]$

$$r(K,t-s)=\left\{-\frac{K^2}{12\pi^2}\int 2(1-\cos\nu u)Y_0(\nu)\right.$$
$$\left.[Y_0(\nu)+2iA_0(\nu)]e^{i\nu(\tau-\sigma)}d\nu\right\}$$
$$\times\exp\left[\frac{iK^2}{2\pi}\int(1-\cos\nu u)(Y_0+iA_0)d\nu\right] \quad \text{(B3)}$$

This will make a contribution to $G_1(\tau-\sigma)$. It is already in the form of a Fourier transform so for the contribution to $G_1(\nu)$ we omit the integral on ν and the factor $e^{i\nu(\tau-\sigma)}$. According to (29) we must next multiply $r(K,t-s)$ by $y(\omega_K,u)+ia(\omega_K,u)$ and integrate on u. This is best done by dividing the range of u from 0 to ∞ and from $-\infty$ to 0, and in the latter putting $u\to -u$ so that all integrals are over positive u only.

The integral in the exponent is $-K^2/2$ times

$$-\frac{i}{\pi}\int(1-\cos\nu u)[Y_0(\nu)+iA_0(\nu)]d\nu$$
$$=+i[\mathbf{Y}_0(u)+\mathbf{Y}_0(-u)]+2[\mathbf{A}_0(0)-\mathbf{A}_0(u)] \quad \text{(B4)}$$

[where $\mathbf{Y}_0(t)$, $\mathbf{A}_0(t)$ are the inverse transforms of $Y_0(\nu)$, $A_0(\nu)$]. Expression (46) is equal to $D(u)$, defined in (35c) for $u>0$, and $D(-u)$ for $u<0$, since $\mathbf{Y}_0(u)=0$ for $u<0$. Thus, this term in $r(K,t-s)$ contributes a piece

$$-\frac{K^2}{3\pi}Y_0(\nu)[Y_0(\nu)+2iA_0(\nu)]$$
$$\times\int_0^\infty(1-\cos\nu u)[y(\omega_K,u)+y(\omega_K,-u)$$
$$+2ia(\omega_K,u)]e^{-\frac{1}{2}K^2D(u)}du. \quad \text{(B5)}$$

Adding the three other corresponding pieces from $\exp\{i\mathbf{K}\cdot[\mathbf{X}'(t)-\mathbf{X}'(s)]\}$, etc., multiplying by $|C_K|^2$, and integrating over K [see Eq. (30)] gives the first term in (35). The second term is gotten in an analagous way from Φ_0. We need to expand our expression for $r(K,t-s)$ just to first order in K^2. The terms like $e^{-\frac{1}{2}K^2D(u)}$ are replaced by one. The resulting expression is an integral on u, $\int_0^\infty(1-e^{i\nu u})S_0(u)du$, where

$$S_0(u)=\mathrm{Im}C\left(e^{iwu}+\frac{2\cos wu}{e^{\beta w}-1}\right)=C\sin wu.$$

The integral on u gives $C\nu^2/w(\nu^2-w^2)$, as in the last term of (34).

ANNALS OF PHYSICS: 24, 118–173 (1963)

The Theory of a General Quantum System Interacting with a Linear Dissipative System

R. P. FEYNMAN

California Institute of Technology, Pasadena, California

AND

F. L. VERNON, JR.*

Aerospace Corporation, El Segundo, California

A formalism has been developed, using Feynman's space-time formulation of nonrelativistic quantum mechanics whereby the behavior of a system of interest, which is coupled to other external quantum systems, may be calculated in terms of its own variables only. It is shown that the effect of the external systems in such a formalism can always be included in a general class of functionals (influence functionals) of the coordinates of the system only. The properties of influence functionals for general systems are examined. Then, specific forms of influence functionals representing the effect of definite and random classical forces, linear dissipative systems at finite temperatures, and combinations of these are analyzed in detail. The linear system analysis is first done for perfectly linear systems composed of combinations of harmonic oscillators, loss being introduced by continuous distributions of oscillators. Then approximately linear systems and restrictions necessary for the linear behavior are considered. Influence functionals for all linear systems are shown to have the same form in terms of their classical response functions. In addition, a fluctuation-dissipation theorem is derived relating temperature and dissipation of the linear system to a fluctuating classical potential acting on the system of interest which reduces to the Nyquist-Johnson relation for noise in the case of electric circuits. Sample calculations of transition probabilities for the spontaneous emission of an atom in free space and in a cavity are made. Finally, a theorem is proved showing that within the requirements of linearity all sources of noise or quantum fluctuation introduced by maser-type amplification devices are accounted for by a classical calculation of the characteristics of the maser.

I. INTRODUCTION

Many situations occur in quantum mechanics in which several systems are coupled together but one or more of them are not of primary interest. Problems

* This report is based on a portion of a thesis submitted by F. L. Vernon, Jr. in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the California Institute of Technology, 1959.

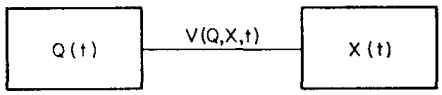


FIG. 1. General quantum systems Q and X coupled by a potential $V(Q, X, t)$

in the theory of measurement and in statistical mechanics present good examples of such situations. Suppose, for instance, that the quantum behavior of a system is to be investigated when it is coupled to one or more measuring instruments. The instruments in themselves are not of primary interest. However, their effects are those of perturbing the characteristics of the system being observed. A more concrete example is the case of an atom in an excited state which interacts with the electromagnetic field in a lossy cavity resonator. Because of the coupling there will be energy exchange between the field and the atom until equilibrium is reached. If, however, the atom were not coupled to any external disturbances, it would simply remain unperturbed in its original excited state. The cavity field, although not of central interest to us, influences the behavior of the atom.

To make the discussion more definite, let us suppose there are two nonrelativistic quantum systems whose coordinates are represented in a general way by Q and X, as in Fig. 1, coupled together through some interaction potential which is a function of the parameters of the two systems. It is desired to compute the expectation value of an observable which is a function of the Q variables only. As is well known, the complete problem can be analyzed by taking the Hamiltonian of the complete system, forming the wave equation as follows:

$$\{H(Q) + H(X) + V(Q, X)\}\psi(Q, X) = -\left(\frac{\hbar}{i}\right)\frac{\partial}{\partial t}\psi(Q, X),$$

and then finding its solution. In general, this is an extremely difficult problem. In addition, when this approach is used, it is not easy to see how to eliminate the coordinates of X and include its effect in an equivalent way when making computations on Q. A satisfactory method of formulating such problems as this in a general way was made available by the introduction of the Lagrangian formulation of quantum mechanics by Feynman. He applied the techniques afforded by this method extensively to studies in quantum electrodynamics. Thus, in a problem where several charged particles interact through the electromagnetic field, he found that it was possible to eliminate the coordinates of the field and recast the problem in terms of the coordinates of the particles alone. The effect of the field was included as a delayed interaction between the particles (*1*, *2*).

The central problem of this study is to develop a general formalism for finding all of the quantum effects of an environmental system (the interaction system) upon a system of interest (the test system), to investigate the properties of this formalism, and to draw conclusions about the quantum effects of specific

interaction systems on the test system. Cases where the interaction system is composed of various combinations of linear systems and classical forces will be considered in detail. For the case in which the interaction system is linear, it will be found that parameters such as impedance, which characterize its classical behavior, are also important in determining its quantum effect on the observed system. Since this linear system may include dissipation, the results have application in a study of irreversible statistical mechanics.

In Section II, after a brief discussion of the Lagrangian formulation of quantum mechanics, a general formulation of the problem is made and certain functionals, called influence functionals, will be defined, which contain the effect of the interaction system (such as system X in Fig. 1) on the test system in terms of the coordinates of the test system only. General properties of these functionals will be derived and their relationship to statistical mechanics will be discussed. To obtain more specific information about the properties of the formalism, we then specialize the discussion to cases where well-defined systems are involved. In Section III, the special cases are considered in which the interaction system is a definite classical force and a random classical force. In Section IV, the influence functionals for exactly linear systems at zero temperatures are derived and then extended to the case that the linear systems are driven by classical forces. In addition, the effect of finite temperatures of linear systems is considered. Then, in Section V, the unobserved systems are again assumed to be general but weakly coupled to the observed system. Within the approximation of weak coupling these general systems also behave as if they were linear. Then finally in Section VI, the results of the analysis are used to prove a general result concerning maser noise.

It is to be emphasized that although we shall talk of general test and interaction systems, the Lagrangian formulation is restricted to cases involving momentum or coordinate operators. Therefore, strictly speaking, systems in which the spin is of importance are not covered by this analysis. However, this has no bearing on the results since their nature is such that their extension to the case where spins are important can be inferred.

An equivalent approach can be made to the problem using the Hamiltonian formulation of quantum mechanics by making use of the ordered operator calculus developed by Feynman (*3*). This approach has been used to some extent by Fano (*4*) and has been developed further by Hellwarth (*5*).[1] Some advantages of this method are that many results may be obtained more simply than by the Lagrangian method and nonclassical concepts such as spin enter the formalism naturally. However, the physical significance of the functions being dealt with are often clearer in the Lagrangian method.

[1] Many of the results obtained in this work have also been obtained by him using ordered operator techniques.

II. GENERAL FORMULATION—INFLUENCE FUNCTIONAL

A. Lagrangian Formulation of Quantum Mechanics

We shall begin the discussion with a brief introduction to the Lagrangian or space-time approach to quantum mechanics and the formal way in which one may set up problems of many variables.[2] Let us suppose that we are considering a single system which has coordinates that are denoted by Q, and that for the time being it is not acted on by any other quantum system. It can be acted on by outside forces, however. The system may be very complicated, in which case Q represents all the coordinates in a very general way. If at a time t the variable Q is denoted by Q_t, then the amplitude for the system to go from position Q_τ at $t = \tau$ to Q_T at $t = T$ is given by

$$K(Q_T, T; Q_\tau, \tau) = \int \exp[(i/\hbar)S(Q)]\mathfrak{D}Q(t) \tag{2.1}$$

in integral which represents the sum over all possible paths $Q(t)$ in coordinate space from Q_τ to Q_T of the functional $\exp[(i/\hbar)S(Q)]$.[3] $S(Q) = \int_\tau^T L(\dot{Q}, Q, t)\, dt$ is the action calculated classically from the Lagrangian for the trajectory $Q(t)$. For the case that Q is a single linear coordinate of position, this is represented in the diagram in Fig. 2. The magnitude of the amplitude for all paths is equal but the phase for each path is given by the classical action along that path in units of $\hbar$. Thus, amplitudes for neighboring paths which have large phases tend to cancel. The paths which contribute the greatest amount are those whose amplitudes have stationary phases for small deviations around a certain path. This is the path for which the classical action is at an extremum and is, therefore, the classical path. Remarkably enough, for free particles and harmonic oscillators, the result of the path integration is

$$K(Q_T, Q_\tau) = (\text{Smooth Function}) \exp[(i/\hbar)S_{\text{cl}}]$$

where S_{cl} is the action evaluated along the classical path between the two end points Q_τ, Q_T. However, for more complicated systems this simple relation does not hold. A discussion of the methods of doing integrals of this type is not included here since methods appropriate for the purposes here are already contained in the literature (*1*, *2*).

Since $K(Q_T, Q_\tau)$ is the amplitude to go from coordinate Q_τ to Q_T, it follows that at $t = T$ the amplitude that the system is in a state designated by $\phi_m(Q_T)$ when initially in a state $\phi_n(Q_\tau)$ is given by

$$\begin{aligned} A_{mn} &= \int \phi_m^*(Q_T) K(Q_T, Q_\tau) \phi_n(Q_\tau)\, dQ_T\, dQ_\tau \\ &= \int \phi_m^*(Q_T) \exp[(i/\hbar)S(Q)] \phi_n(Q_\tau) \mathfrak{D}Q(t)\, dQ_T\, dQ_\tau \end{aligned} \tag{2.2}$$

[2] For a more complete treatment, see ref. *1*.

[3] In subsequent equations $K(Q_T, T; Q_t, t)$ will be written $K(Q_T, Q_t)$.

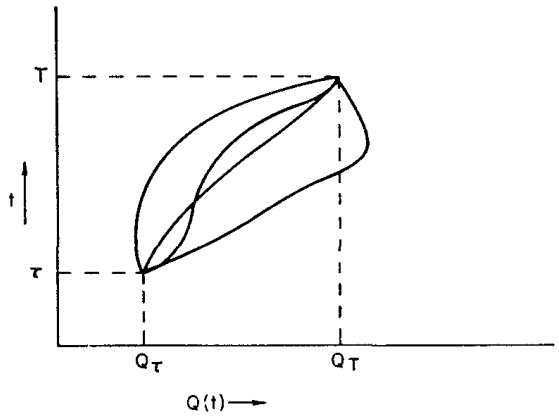


FIG. 2. Space-time diagram showing possible paths for particle to proceed from Q_τ to Q_T

The probability of the transition from $n \rightarrow m$ is given by $|A_{mn}|^2$ and from Eq. (2.2) this can be written in the form of multiple integrals as follows:

$$P_{mn} = \int \phi_m^*(Q_T)\phi_m(Q'_T) \exp\{(i/\hbar)[S(Q) - S(Q')]\} \times \phi_n(Q_\tau)\phi_n^*(Q'_\tau)\mathfrak{D}Q(t)\mathfrak{D}Q'(t)\, dQ_\tau\, dQ'_\tau\, dQ_T\, dQ'_T \tag{2.3}$$

As a example of a more complicated case let us consider two systems whose coordinates are Q and X.[4] The systems are coupled by a potential which can be designated as $V(Q, X)$ and incorporated in the total Lagrangian. We assume that when $V = 0$ the states of Q and X can be described by sets of wave functions $\phi_k(Q)$ and $\chi_p(X)$ respectively. If, initially, Q is in a state $\phi_n(Q_\tau)$ and X is in a state $\chi_i(X_\tau)$, then the amplitude that Q goes from state n to m while X goes from state i to f can be formed in a similar way to that of Eq. (2.2),

$$A_{mf,ni} = \int \phi_m^*(Q_T)\chi_f^*(X_T) \exp[(i/\hbar)S(Q, X)]\phi_n(Q_\tau)\chi_i(X_\tau) \times \mathfrak{D}Q(t)\mathfrak{D}X(t)\, dQ_\tau\, dX_\tau\, dQ_T\, dX_T \tag{2.4}$$

where $S(Q, X)$ represents the classical action of the entire system including both Q and X. The important property of separability afforded by writing the amplitude in this way is now apparent.[5] For instance, if one wishes to know the effect that X has on Q when X undergoes a transition from state i to f, then all

[4] Each system will be denoted by the coordinates that characterize it. Where Q or X means specifically a coordinate, it will be so designated by a statement if it is not obvious.

[5] If system Q represents a harmonic oscillator and the interaction of Q with X were linear and of the form $-\gamma(t, X)Q(t)$, then that part of Eq. (2.4) which involves the Q variables corresponds to the function G_{mn} defined and used by Feynman to eliminate the electromagnetic field oscillators. See ref. 2.

of the integrals on the X variables may be done first. What is left is an expression for A_{mn} for Q and in terms of Q variables only but with the effect of X included. The extension of writing transition amplitudes for large numbers of systems is obvious. In principle the order in which the variables are eliminated is always arbitrary.

B. Definition of Influence Functional

A functional can now be defined which can be used to describe mathematically the effect of external quantum systems upon the behavior of a quantum system of interest.[6]

The fundamental theorem for this work may be stated as follows: For any system, Q, acted on by external classical forces and quantum mechanical systems as discussed above, the probability that it makes a transition from state $\psi_n(Q_\tau)$ at $t = \tau$ to $\psi_m(Q_T)$ at $t = T$ can be written

$$P_{mn} = \int \psi_m^*(Q_T)\psi_m(Q'_T) \exp\{(i/\hbar)[S_0(Q) - S_0(Q')]\}\mathfrak{F}(Q, Q') \times \psi_n^*(Q'_\tau)\psi_n(Q_\tau)\mathfrak{D}Q(t)\mathfrak{D}Q'(t)\, dQ_\tau\, dQ'_\tau\, dQ_T\, dQ'_T \tag{2.5}$$

where $\mathfrak{F}(Q, Q')$ contains all the effects of the external influences on Q, and $S_0(Q) = \int_\tau^T L(\dot{Q}, Q, t)\, dt$, the action of Q without external disturbance. The proof of this is straightforward. Let us examine two coupled systems characterized by coordinates Q and X as represented diagrammatically in Fig. 1. Q will represent the test system and X the quite general interaction system, (excepting only the effects of spin) coupled by a general potential $V(Q, X, t)$ to Q. Assume Q to be initially ($t = \tau$) in state $\psi_n(Q_\tau)$ and X to be in state $\chi_i(X_\tau)$, a product state. The probability that Q is found in state $\psi_m(Q_T)$ while X is in state $\chi_f(X_T)$ at $t = T$ can be written in the manner discussed above and is

$$\begin{aligned} P_{mf,ni} &= |A_{mf,ni}|^2 \\ &= \int \psi_m^*(Q_T)\psi_m(Q'_T)\chi_f^*(X_T)\chi_f(X'_T) \\ &\quad \times \exp\{(i/\hbar)[S_0(Q) - S_0(Q') + S(X) - S(X') + S_I(Q, X) \\ &\qquad - S_I(Q',X')]\} \\ &\quad \times \psi_n^*(Q'_\tau)\psi_n(Q_\tau)\chi_i^*(X'_\tau)\chi_i(X_\tau)\, dX_\tau \\ &\qquad \cdots dQ'_T\mathfrak{D}X'(t) \cdots \mathfrak{D}Q(t) \end{aligned} \tag{2.6}$$

The primed variables were introduced when the integrals for each $A_{mf,ni}$ were combined. Now if all of Eq. (2.6) which involves coordinates other than Q or

[6] Hereafter, the system of interest will be referred to as the test system. Conversely, the system not of primary interest will be called the interaction or environmental system.

Q' is separated out and designated as $\mathfrak{F}(Q, Q')$, then the following expression is obtained

$$\mathfrak{F}(Q, Q') = \int \chi_f^*(X_T)\chi_f(X'_T) \cdot \exp\{(i/\hbar)[S(X) - S(X') + S_I(Q, X) - S_I(Q', X')]\} \times \chi_i^*(X'_\tau)\chi_i(X_\tau)\, dX_\tau\, dX'_\tau \mathfrak{D}X(t)\mathfrak{D}X'(t) \tag{2.7}$$

Incorporation of this expression into Eq. (2.6) yields the desired form of Eq. (2.5). If the path integrals are written in terms of kernels, Eq. (2.7) becomes

$$\mathfrak{F}(Q, Q') = \int \chi_f^*(X_T)\chi_f(X'_T)K_Q(X_T, X_\tau)K_{Q'}^*(X'_T, X'_\tau) \times \chi_i^*(X'_\tau)\chi_i(X_\tau)\, dX_\tau \cdots dX'_T \tag{2.8}$$

where the subscript Q means that the kernel includes the effect of a potential $V(Q, X)$ acting on X during the interval $T > t > \tau$. As can be seen, $\mathfrak{F}$ is a functional whose form depends upon the physical system X, the initial and final states of X, and the coupling between Q and X.

It is to be emphasized that the formulation of $\mathfrak{F}$ is such that it includes all the effects of the interaction system in influencing the behavior of the test system. Thus, if there are two systems A and B which can act on Q, and if

$$\mathfrak{F}_{A \text{ on } Q} = \mathfrak{F}_{B \text{ on } Q},$$

then the effects of A on Q are the same as those of B on Q. It follows that if simplifying assumptions are necessary in finding $\mathfrak{F}_{A \text{ on } Q}$ and $\mathfrak{F}_{B \text{ on } Q}$ (due to the complicated nature of A and B) and if the resulting functions are equal, then within the approximations the effects of A and B on the test system are the same. In the situation where the interaction system is composed of a linear system or combinations of linear systems we shall see that the same form of $\mathfrak{F}$ is always appropriate. To adapt this general form of $\mathfrak{F}$ to a particular linear system it is only necessary to know such quantities as impedance and temperature which determine its classical behavior. In still other situations, very weak coupling between systems is involved. The approximate $\mathfrak{F}$ which can be used in this case to represent the effect of the interaction systems has a form which is independent of the nature of the interaction system. This form is the same as for linear systems. These cases will be considered in more detail in later sections.

C. General Properties of Influence Functionals

There are several general properties of influence functionals which are of interest and which will be useful in subsequent arguments. The first three of these (1, 2, 3) follow directly from the definition of $\mathfrak{F}(Q, Q')$. The last two (4, 5) will require more discussion.

1. If the physical situation is unsure (as for instance if the type of interaction system X, or the initial or final states are not known precisely) but if the probability of the pth situation is w_p and the corresponding influence functional is $\mathfrak{F}_p$, then the effective $\mathfrak{F}$ is given by

$$\mathfrak{F}_{\text{eff}} = \sum_p w_p \mathfrak{F}_p \equiv \langle \mathfrak{F} \rangle \tag{2.9}$$

Thus, in Eq. (2.6) if the initial state of X were not certain but the probability of each initial state were w_i, then P_{mn} for system Q would be given by $\sum_i w_i P_{mf,ni}$. Since the summation involves only the part of Eq. (2.6) involving the X variables, it is a sum over the influence functions for each possible initial state and results in an average influence functional of the type given above.

2. If a number of statistically and dynamically independent partial systems act on Q at the same time and if $\mathfrak{F}^{(k)}$ is the influence of the kth system alone, the total influence of all is given by the product of the individual $\mathfrak{F}^{(k)}$:

$$\mathfrak{F} = \prod_{k=1}^{N} \mathfrak{F}^{(k)} \tag{2.10}$$

Again referring to Eq. (2.6), if there were N subsystems interacting with Q, then the probability that Q makes a transition from state n to m while each of the subsystems makes a transition from its initial to its final state is given by an expression of the same form as Eq. (2.6). The difference in this case being that the term involving the X variables would be replaced by a product of N similar terms—one for each subsystem. Thus, when the term involving all the $X^{(k)}$ variables is separated out the complete influence functional is recognized as a product of the functionals $\mathfrak{F}^{(k)}(Q, Q')$ for each subsystem.

DEFINITION: In many cases it will be convenient to write $\mathfrak{F}$ in the form $\exp[i\Phi(Q, Q')]$. Φ is then called the influence phase. For independent disturbances as considered in 2, the influence phases add. In the event that $i\Phi(Q, Q')$ is a real number we will continue to use the notation Φ; the phase simply becomes imaginary. It will frequently be more convenient to work with Φ rather than $\mathfrak{F}$.

3. The influence functional has the property that

$$\mathfrak{F}^*(Q, Q') = \mathfrak{F}(Q', Q) \tag{2.11}$$

Referring to Eq. (2.7), the definition of the influence functional, this fact follows immediately upon interchanging Q and Q'.

4. In the class of problems in which the final state of the interaction system is arbitrary, which means the final states are to be summed over, then $\mathfrak{F}(Q, Q')$ is independent of $Q(t)$ if $Q(t) = Q'(t)$ for all t. All of the problems we will be concerned with here are of this type.

The validity of this statement can be ascertained by observing Eq. (2.7), the general definition of the influence functional. In particular, for the case where the initial and final states of the interaction system X are i and f respectively, as in

Eq. (2.7), we denote the influence functional by ${}_f\mathfrak{F}_i(Q, Q')$. Let us assume we have no interest in the final state of X which means that ${}_f\mathfrak{F}_i(Q, Q')$ must be summed over all such states. The initial state i can be quite general. Thus, the influence functional for the case of an arbitrary final state is

$$\mathfrak{F}_i(Q, Q') = \sum_f {}_f\mathfrak{F}_i(Q, Q')$$

For clarity in finding the result of letting $Q(t) = Q'(t)$ for all t in $\mathfrak{F}_i(Q, Q')$, we will write out the expression explicitly from Eq. (2.7). It is

$$\mathfrak{F}_i(Q, Q) = \int \sum_f \chi_f^*(X_T)\chi_f(X'_T) \cdot \exp\{(i/\hbar)[S(X) - S(X') + S_I(Q, X) - S_I(Q, X')]\} \times \chi_i^*(X'_\tau)\chi_i(X_\tau)\, dX_\tau \cdots \mathfrak{D}X'(t)$$

Since Q appears in the interaction potentials acting on the X and X' variables respectively, it loses its identity as the coordinate of a quantum system and becomes just a number (which may be, of course, a function of time). Thus $S_I(Q, X)$ may be interpreted as the action of an external potential which drives the X system. The above expression then represents the probability that X, which is in state i initially, is finally in any one of its possible states after being acted on by an external potential (as, for instance, in Eq. (2.3) summed over the final states, m). This result is unity. We have then that $\mathfrak{F}(Q, Q) = 1$ and is independent of $Q(t)$.

5. A more restrictive statement of the property in the above paragraph (4) can be made. In this same class of problems in which the final states are summed over, if $Q(t) = Q'(t)$ for all $t > r$ then $\mathfrak{F}(Q, Q')$ is independent of $Q(t)$ for $t > r$. To see this we write down the influence functional from Eq. (2.8) breaking up the time interval into two parts, before and after r. Setting $Q = Q'$ for $t > r$ and utilizing the closure relation for the sum over final states we have,

$$\mathfrak{F}_r(Q, Q') = \int \delta(X_T - X'_T)K_Q(X_T, X_r)K_Q^*(X'_T, X'_r) \times K_Q(X_r, X_\tau)K_{Q'}^*(X'_r, X'_\tau)\chi_i^*(X'_\tau)\chi_i(X_\tau)\, dX_\tau \cdots dX'_T$$

Examining the parts of the above integral which contain the effects of $t > r$:

$$\int \delta(X_T - X'_T)K_Q(X_T, X_r)K_Q^*(X'_T, X'_r)\, dX_T\, dX'_T = \int K_Q(X'_r, X_T)K_Q(X_T, X_r)\, dX_T = K_Q(X'_r, X_r) = \delta(X_r - X'_r)$$

The expression for $\mathfrak{F}_r(Q, Q')$ becomes then

$$\mathfrak{F}_r(Q, Q') = \int \delta(X_r - X'_r)K_Q(X_r, X_\tau)K_{Q'}^*(X'_r, X'_\tau) \times \chi_r^*(X'_\tau)\chi_i(X_\tau)\, dX_\tau \cdots dX'_r,$$

which is independent of $Q(t)$ for $t > r$. As will be seen later in the specific case of linear systems, this leads to a statement of causality.

D. Statistical Mechanics

Finally it is appropriate to point out explicitly the significance of the influence functional in a study of quantum statistical mechanics. In the class of problems considered here we are only interested in making measurements on the test system and not on the interacting system. Thus, when the expectation value of an operator which acts only on the test system variables is taken, the final states of the interaction system must be summed over. It is equivalent to taking the expectation value of the desired operator in the test system and simultaneously the unit operator in the interaction system. Therefore, only the influence functional where the final states of the interaction system are summed over will be of interest to us.

Starting with the coordinate representation of the density matrix (6) for the test and interaction systems, $\rho(Q, X; Q', X')$, we will show the part played by the influence functional in obtaining an expression for $\rho(Q_T, Q'_T)$, that is, with the X coordinates eliminated, in terms of its value at an earlier time τ, $\rho(Q_\tau, Q'_\tau)$. First, we recall that the definition of ρ is as follows:

$$\rho(Q, X; Q', X') = \langle \psi(Q, X)\psi^*(Q', X') \rangle_{\text{av}} \tag{2.12}$$

where $\psi(Q, X)$ represents the wave function for one of the systems in an ensemble of systems each representing one of the possible states of the Q, X system (7). The average, represented by $\langle\ \rangle_{\text{av}}$, is taken over the ensemble. The trace of the density matrix is

$$\text{Tr}\, \rho(Q, X; Q', X') = \int\int \rho(Q, X; Q, X)\, dQ\, dX \tag{2.13}$$

and the expectation value of an operator A which operates on the Q variables only is

$$\langle A \rangle = \int\int\int \rho(Q, X; Q', X) A(Q, Q')\, dQ\, dQ'\, dX \tag{2.14}$$

In the above

$$\begin{aligned} A(Q', Q) &= \textstyle\sum_{i,j} A_{ij}\phi_i^*(Q)\phi_j(Q'), \\ A_{ij} &= \int \phi_i^*(Q) A \phi_j(Q)\, dQ, \end{aligned} \tag{2.15}$$

and $\phi_i(Q)$ is one of a set of complete orthonormal eigenfunctions. From Eq. (2.14) then we see that the formal expression which we wish to derive is

$$\int \rho(Q_T, X_T; Q'_T, X_T)\, dX_T = \rho(Q_T, Q'_T) \tag{2.16}$$

in terms of $\rho(Q_\tau, Q'_\tau)$. From the rules given in Section II.A for propagation of a wave function with time we can easily find ρ_T in terms of ρ_τ. Thus

$$\rho(Q_T, X_T; Q'_T, X'_T) = \int \exp\{(i/\hbar)[S_0(Q) - S_0(Q') + S(X) - S(X') + S_I(Q, X) - S_I(Q', X')]\} \times \rho(Q_\tau, X_\tau; Q'_\tau, X'_\tau)\mathfrak{D}Q(t) \cdots dX'_\tau \quad (2.17)$$

Now, for simplicity let us assume that initially the two systems are independent so that

$$\rho(Q_\tau, X_\tau; Q'_\tau, X'_\tau) = \rho(Q_\tau, Q'_\tau)\rho(X_\tau, X'_\tau)$$

Then eliminating the X_T coordinate as indicated in Eq. (2.16) we have

$$\rho(Q_T, Q'_T) = \int \{\int \delta(X_T - X'_T) \exp[(i/\hbar)[S(X) - S(X') + S_I(Q, X) - S_I(Q', X')]]\rho(X_\tau, X'_\tau) \times \mathfrak{D}X(t) \cdots dX'_\tau\} \exp[(i/\hbar)[S_0(Q) - S_0(Q')]]\rho(Q_\tau, Q'_\tau)\mathfrak{D}Q(t) \cdots dQ'_\tau$$

The expression inside the braces is identified as $\mathfrak{F}(Q, Q')$ for the case in which the final state of X is summed over. Therefore, the following result is obtained:

$$\rho(Q_T, Q'_T) = \int \mathfrak{F}(Q, Q') \exp[(i/\hbar)[S_0(Q) - S_0(Q')]]\rho(Q_\tau, Q'_\tau)\mathfrak{D}Q(t) \cdots dQ'_\tau \quad (2.18)$$

Thus, if the density matrix of the test system Q is represented by $\rho(Q_\tau, Q'_\tau)$ at some initial instant τ, the density matrix $\rho(Q_T, Q'_T)$ at some later time T is given by Eq. (2.18). The entire influence of the interaction system is contained in $\mathfrak{F}(Q, Q')$.

E. Use of Influence Functionals

At this point we need to consider how influence functionals can be used in the analysis of a problem. For clarity the discussion will be specialized to a particular problem but the principle is valid more generally. Suppose we wish to know the probability that a test system Q makes a transition from an initial state $\phi_n(Q_\tau) \exp[(-i/\hbar)E_n\tau]$ to a final state $\phi_m(Q_T) \exp[(-i/\hbar)E_mT]$ when coupled to an interaction system. The formal expression for this probability is, from Eq. (2.5),

$$P_{nm} = \int \phi_m^*(Q_T)\phi_m(Q'_T) \exp\{(i/\hbar)[S_0(Q) - S_0(Q')]\}\mathfrak{F}(Q, Q') \times \phi_n^*(Q'_\tau)\phi_n(Q_\tau)\, dQ_\tau \cdots dQ'(t) \quad (2.19)$$

This is formally exact but except in special cases it cannot be evaluated exactly. Furthermore, to obtain any specific answers to the problem the characteristics

of Q must be known as well as knowing the influence functional. However, by using perturbation theory we may find general expressions for transition probabilities to as many orders as desired. For example, if the interaction system is a linear system at zero temperature, we will find that $\mathfrak{F}_0(Q, Q')$ is of the form $\exp[i\Phi_0(Q, Q')]$. The perturbation expansion is obtained by writing $\exp[i\Phi_0(Q, Q')]$ in terms of a power series and evaluating the path integral corresponding to each term in the expansion. In many cases the coupling between Q and the interaction system is small enough that only a few terms in the expansion are necessary. In Appendix I the basic procedure for finding the perturbation expansion is demonstrated by finding the specific expression up to second order in the potentials involved for transition probability of a test system when acted on by a linear interaction system at zero temperature. Calculation of transition probabilities represent only one piece of information that one might desire to know about a test system. For instance, it is more usually desired to find the expectation value of an operator in the test system. To calculate this one needs to know the density matrix describing the test system when it is coupled to an interaction system. The exact expression for the required density matrix is given in Section II,D. Again in the general case, one runs into the difficulty of making an exact calculation and is forced to make calculations using perturbation theory. The same procedure of expanding the influence functional into a power series and performing the required path integrations yields useful perturbation expressions.

III. INFLUENCE FUNCTIONALS FOR CLASSICAL POTENTIALS

In this section we will derive specific forms and properties of influence functionals for the effects of classical potentials on the test system. These represent the simplest form of influence functionals and their properties follow directly from the general properties obtained in the previous section. These forms will then be extended to the case where the classical potential represents Brownian noise.

A. Properties of Influence Functionals for Classical Potentials

The first step is to find the influence functional for a definite classical potential acting on the test system, Q. If the potential energy term in the Lagrangian is of the form $V(Q, t)$, then it can be ascertained readily by referring to the fundamental definition of $\mathfrak{F}(Q, Q')$ that

$$\mathfrak{F}(Q, Q') = \exp\{-(i/\hbar)\textstyle\int_\tau^T [V(Q, t) - V(Q', t)]\, dt\} \tag{3.1}$$

or equivalently the influence phase is

$$\Phi(Q, Q') = -(1/\hbar)\textstyle\int_\tau^T [V(Q, t) - V(Q', t)]\, dt \tag{3.2}$$

The next degree of complication is to have several potentials, $\sum_k V_k(Q, t)$ acting on Q simultaneously. However, since the sum of all these potentials represents an equivalent potential, say $V(Q, t) = \sum_k V_k(Q, t)$, then it is obvious that the total influence functional $\mathfrak{F}(Q, Q')$ is the product of the individual $\mathfrak{F}_k(Q, Q')$. More specifically,

$$\begin{aligned}\mathfrak{F}(Q, Q') &= \exp\{-(i/\hbar)\textstyle\int_\tau^T [V(Q, t) - V(Q', t)]\, dt\} \\ &= \exp \textstyle\sum_k \{-(i/\hbar)\int_\tau^T [V_k(Q, t) - V_k(Q', t)]\, dt\} \\ &= \textstyle\prod_k \mathfrak{F}_k(Q, Q'),\end{aligned} \tag{3.3}$$

or

$$\begin{aligned}\Phi(Q, Q') &= -(1/\hbar)\textstyle\sum_k \int_\tau^T [V_k(Q, t) - V_k(Q', t)]\, dt \\ &= \textstyle\sum_k \Phi_k(Q, Q')\end{aligned} \tag{3.4}$$

The same result follows directly from Section II.C.3 which gives the total influence functional for several statistically and dynamically independent systems acting on Q. The total influence functional for all the systems (in this case potentials) is the product of the functionals for the individual systems.

Another property of the classical influence functionals is obtained by inspection of Eq. (3.1). We notice that for any classical $\mathfrak{F}(Q, Q')$ if conditions are such that $Q(t) = Q'(t)$, then $\mathfrak{F}(Q, Q') = 1$ and is independent of t for all times that the two variables are equal. It follows that the influence phase is zero for this condition.

Finally, from Section II.C.1 we find that if the potential is uncertain but the probability of each $V_r(Q, t)$ is w_r then the average functional is given by

$$\begin{aligned}\langle \mathfrak{F}(Q, Q') \rangle &= \textstyle\sum_r w_r \exp\{-(i/\hbar)\int_\tau^T [V_r(Q, t) - V_r(Q', t)]\, dt\} \\ &= \textstyle\sum_r w_r \mathfrak{F}_r(Q, Q')\end{aligned} \tag{3.5}$$

In the following Sections we will assume a probability distribution w_r appropriate to Brownian noise and will be able to derive a specific form for the average influence functional.

B. Specific Functionals for Random Potentials

Let us now suppose that the potential has known form, $V(Q)$, but unknown strength $C(t)$ as a function of time so that the total potential is $V(Q, t) = C(t)V(Q)$. The average influence functional for two cases involving this type of potential will be particularly useful in the discussion contained in Sections IV and V. These cases are: (1) when $C(t)$ is characterized by any coupling strength (average magnitude of C) with a purely Gaussian distribution, and (2) when $C(t)$ is composed of large number of very weak potentials (acting on the test

system simultaneously) whose distributions are stationary but not necessarily Gaussian.

1. *Gaussian Noise*

First, we consider the situation when $C(t)$ is Gaussian noise with a power spectrum $\Phi(\nu)$ and a correlation function $R(\tau) = (2/\pi)\int_0^\infty \Phi(\nu) \cos \nu\tau \, d\nu$ then $\langle \mathfrak{F} \rangle$ is given by

$$\begin{aligned} \langle \mathfrak{F} \rangle &= \langle \exp\{(i/\hbar)\textstyle\int_\tau^T C(t)[V(Q)-V(Q')]\,dt\}\rangle \\ &= \exp\{-\hbar^{-2}\textstyle\int_\tau^T\int_\tau^t R(t-s)[V(Q_t)-V(Q'_t)][V(Q_s)-V(Q'_s)]\,ds\,dt\} \end{aligned} \tag{3.6}$$

Expressed in Fourier transform notation this becomes

$$\langle \mathfrak{F} \rangle = \exp\{-(\pi\hbar^2)^{-1}\textstyle\int_0^\infty \phi(\nu) \mid V_\nu(Q)-V_\nu(Q') \mid^2 d\nu\} \tag{3.7}$$

where

$$V_\nu(Q)-V_\nu(Q') = \textstyle\int_\tau^T [V(Q)-V(Q')]e^{-i\nu t}dt \tag{3.8}$$

Expressions of the type given in Eq. (3.6) are common for operations in which it is required to find the characteristic function, $F(i\zeta) = \langle e^{i\zeta f(T)} \rangle$ for $f(T)$ represented by integrals of the form $f(T) = \int_{a(T)}^{b(T)} A(T, t)x(t)\,dt$ where $x(t)$ is a Gaussian process. The result will not be worked out here as it may be found in standard references (*8*).[7] The equivalent expression for $\mathfrak{F}$ in terms of frequency components, Eq. (3.7), is obtained from Eq. (3.6) in a direct manner using the definitions for $R(t)$ and Eq. (3.8).

2. *Brownian Noise*

The Gaussian behavior of Brownian noise, characterized by the typical Gaussian probability distribution, may be the result of the cumulative effects of many small statistically independent sources, none of which is truly Gaussian. How that comes about can be seen as follows. The effect of these small sources on a test system may be represented by an influence functional of the same form as that of Eq. (3.6) where now $C(t) = \sum_{i=1}^N C_i(t)$, N is a very large number, and the $C_i(t)$ are independent random variables. Application of the central-limit theorem to this situation shows that the probability distribution appropriate to $C(t)$ is asymptotically normal subject to the following conditions (*9*):

(a) The average values,

$$\langle C_i(t) \rangle < \infty$$

and

$$\mu_{i,2} \equiv \langle \mid C_i(t) - \langle C_i(t) \rangle \mid^2 \rangle < \infty,$$

[7] See, for example, pp. 372–373 where it is shown that the characteristic function $F(i\xi)$ appropriate to the integral given above for a Gaussian process $x(t)$ is $\exp[-\frac{1}{2}\xi^2\int\int_{a(T)}^{b(T)} \cdot A(T, t)A(T, s)K_x(t, s)\,ds\,dt]$ where the covariance $K_x(t, s) = \langle x_t x_s \rangle$ is the correlation function corresponding to $R(t-s)$ in Eq. (3.6).

(b) The absolute moments

$$\mu_{i,2+\delta} \equiv \langle | C_i(t) - \langle C_i(t) \rangle |^{2+\delta} \rangle$$

exist for some $\delta > 0$, and

(c) Making use of the definition

$$\mu_l = \sum_{i=1}^{N} \mu_{i,l} ,$$

then

$$\lim_{N \to \infty} \frac{\mu_{2+\delta}}{(\mu_2)^{1+\delta/2}} \to 0$$

The condition of independence on the large number of variables and the finite average values required by (a) and (b) above assures that no one component dominates the total distribution. Condition (c) is sufficient to ensure that all higher order correction terms tending to deviate from a normal distribution vanish in the limit of large N. It should be recognized that if the C_i possesses finite third moments $\mu_{i,3}$ the correction terms arising from these moments decrease as $N^{-1/2}$. However, for the cases in which we are interested, the number of the component forces C_i is essentially infinite and higher order terms are negligible.

IV. INFLUENCE FUNCTIONALS FOR LINEAR SYSTEMS

Linear systems are of considerable interest both because of the large number of situations in which they are involved and because they are amenable to exact calculation. In this section the influence functional for arbitrary combinations of oscillators will be found by direct extension of the analysis of a single oscillator. All linear systems which are lossless and those which contain certain kinds of loss can be represented by distributions of oscillators. Situations in which dissipation arises from sources other than distributions of perfect oscillators will be covered in Section V. The same conclusions apply for all linear systems, however, as will be discussed subsequently. For clarity, we will restrict our attention initially to linear interaction systems at zero temperature and not acted on by classical forces. The effects of finite temperature and forces can then be included so that their significance is more apparent.

A. Zero Temperature Linear System

The result to be proven involves the assumption that the interaction system (X) is linearly coupled to the test system (Q). The total Lagrangian for the system is

$$L_{\text{total}} = L_0(\dot{Q}, Q, t) + L(\dot{X}, X, t) + L_I(Q, X) \tag{4.1}$$

where $L_I = \gamma QX$, and $L(\dot{X}, X, t)$ is the part of the Lagrangian involving the X system above. The situation is the same as that shown in Fig. 1 except the interaction potential is given by $V(Q, X, t) = -\gamma QX$ and the X system is linear. Our fundamental theorem for linear systems is as follows:

The influence phase for the effect of X on Q can be written as follows:

$$\Phi(Q\,Q') = \frac{1}{2\pi\hbar}\int_0^\infty \left[\frac{Q'_\nu(Q_{-\nu} - Q'_{-\nu})}{(i\nu Z_\nu)} + \frac{Q_{-\nu}(Q_\nu - Q'_\nu)}{(-i\nu Z_{-\nu})}\right] d\nu \tag{4.2}$$

$\Phi_0(Q, Q')$ is found by studying the properties of X alone.[8] *Q_ν is the Fourier transform of $\gamma(t)Q(t)$ and Z_ν is a classical impedance function which relates the reaction of X to an applied force. Z_ν is found by taking the classical system corresponding to X (that is, whose Lagrangian is $L(\dot{X}, X, t)$) and finding the response of the coordinate X to a driving force $f(t)$ which is derived from the potential $-f(t)X(t)\cdot f(t)$ is considered to be applied at $T = 0$ subject to the initial conditions that $X(0) = \dot{X}(0) = 0$. Z_ν is defined by the expression*

$$Z_\nu = f_\nu/(i\nu X_\nu) \tag{4.3}$$

where

$$f_\nu = \int_0^\infty f(t)e^{-i\nu t}\,dt \quad \text{and} \quad X_\nu = \int_0^\infty X(t)e^{-i\nu t}\,dt$$

In the time domain, Eq. (4.2) can be expressed as

$$i\Phi(Q, Q') = -(1/2\hbar)\int_{-\infty}^{\infty}\int_{-\infty}^{t}\gamma_t\gamma_s(Q_t - Q'_t) \cdot[Q_s F^*(t-s) - Q'_s F(t-s)]\,ds\,dt \tag{4.4}$$

where Im $F(t)$, which we will call $B(t)$ is, for $t > 0$, the classical response of X to a force $f(t) = \delta(t)$. Re $F(t)$, which for this zero temperature case we call $A_0(t)$, is the correlation function for the zero point fluctuation of the variable X, a point discussed at more length below. The relations connecting these quantities are then,

$$\begin{aligned} F(t) &= A_0(t) + iB(t) \\ \frac{1}{i\nu Z\nu} &= \int_0^\infty B(t)e^{-i\nu t}\,dt \end{aligned} \tag{4.5a}$$

and the inverse relations

$$\begin{aligned} A_0(t) &= -\frac{2}{\pi}\int_0^\infty \mathrm{Im}\left(\frac{1}{i\nu Z_\nu}\right)\cos\nu t\,d\nu \\ B(t) &= -\frac{2}{\pi}\int_0^\infty \mathrm{Im}\left(\frac{1}{i\nu Z_\nu}\right)\sin\nu t\,d\nu \end{aligned} \tag{4.5b}$$

[8] More generally, the part of the interaction represented by Q could be represented by function of Q such as $V(Q)$. In this case Q in the influence phase would be replaced by $V_\nu(Q)$, the Fourier transform of $V(Q(t))$.

$A_0(t)$ and $B(t)$ are related as follows:

$$A_0(t) = \frac{2}{\pi}\int_0^\infty \frac{sB(s)}{s^2 - t^2}\,ds \tag{4.5c}$$

These relationships may be written in many forms. Two additional forms are

$$F(t) = \frac{2i}{\pi}\int_0^\infty \left(\frac{1}{i\nu Z_\nu}\right)\cos \nu t\, d\nu$$

and

$$\begin{aligned} \textstyle\int_{-\infty}^{\infty} F(|\,t\,|)e^{-i\nu t}dt &= 2i(i\nu Z_\nu)^{-1} && \text{for } \nu > 0 \\ &= 2i(-i\nu Z_{-\nu})^{-1} && \text{for } \nu < 0 \end{aligned} \tag{4.5d}$$

All the poles of $1/i\nu Z_\nu$ have positive imaginary parts and this impedance function has the additional property that

$$(1/i\nu Z_\nu) = (1/-i\nu Z_{-\nu})^*$$

In the case of finite temperatures, the influence phase can be written in the same form as Eq. (4.4) except that $\mathrm{Re}F(t) = A(t)$, that is, without the subscript 0, and a more general relation exists connecting $A(t)$ and Im $(1/i\nu Z_\nu)$ (see Section IV,C).

$\mathfrak{F}(Q, Q')$ for Single Lossless Harmonic Oscillator

To prove the above theorem, we consider first a test system, Q, which is coupled to a simple harmonic oscillator whose mass is m, characteristic frequency ω, and displacement coordinate X. The complete Lagrangian for X and Q can be written

$$L_{\text{total}} = L_0(\dot{Q}, Q, t) + \tfrac{1}{2}m\dot{X}^2 - \tfrac{1}{2}m\omega^2X^2 + QX \tag{4.6}$$

and the total action is written similarly.[9]

$$S_{\text{total}} = S_0(Q) + \textstyle\int_\tau^T (\frac{1}{2}m\dot{X}^2 - \frac{1}{2}m\omega^2X^2 + QX)\,dt$$

If X is assumed to be initially in the ground state (corresponding to zero temperature) then to within a normalizing constant $\chi_i(X) = e^{-m\omega X^2/2\hbar}$. The final state of X is assumed to be arbitrary which means the final states are to be summed over. Therefore, in Eq. (2.7), the definite state $\chi_f^*(X_T)\chi_f(X'_T)$ will be replaced by the sum $\sum_n \Phi_n^*(X_T)\Phi_n(X'_T) = \delta(X_T - X'_T)$. The $\Phi_n(X)$ represent the energy eigenfunctions of the harmonic oscillator. With this in-

[9] The interaction Lagrangian QX could be written more transparently as γQX where Q and X are the coordinates of the system involved and γ is a coupling factor which may or may not be a function of time. For simplicity in writing the lengthy expressions to follow, γ has been incorporated into an effective coordinate Q since no loss in generality results.

formation available the influence functional is completely defined and from Section II.B can be written

$$\mathfrak{F}_0(Q, Q') = \int \delta(X_T - X'_T) K_Q(X_T, X_\tau) K^*_{Q'}(X'_T, X'_\tau) \times \exp[-(m\omega/2\hbar)(X_\tau^2 + X_\tau'^2)]\, dX_\tau \cdots dX'_T \tag{4.7a}$$

where the subscript Q, Q' refer to the interaction potentials $-QX$ and $-Q'X'$ acting on the X and X' systems respectively. The subscript 0 on $\mathfrak{F}_0(Q, Q')$ indicates zero temperature. For the harmonic oscillator

$$\begin{aligned} K_Q(X_T, X_\tau) &= N \exp\{(i/\hbar)[S(X) - S_I(Q, X)]_{\text{classical}} \\ &= N \exp\{[i\omega/2\hbar \sin\omega(T-\tau)][(X_T^2 + X_\tau^2)\cos\omega(T-\tau) \\ &\quad - 2X_T X_\tau + (2X_T/\omega)\textstyle\int_\tau^T Q_t \sin\omega(t-\tau)dt + (2X_\tau/\omega) \\ &\quad \cdot \textstyle\int_\tau^T Q_t \sin\omega(T-t)\,dt - (2/\omega^2)\int_\tau^T \int_\tau^t Q_t Q_s \sin\omega(T-t)\sin\omega(s-\tau)\,ds\,dt\} \end{aligned} \tag{4.7b}$$

where N is a normalizing factor depending only on ω and the time interval $T - \tau$.[10] Thus, Eq. (4.7a) represents a Gaussian integral over the four X variables since S is itself quadratic in the X variables. When the integrals are carried out the following result is obtained for the influence phase:

$$i\Phi_0(Q, Q') = -(2\hbar m\omega)^{-1}\int_\tau^T \int_\tau^t (Q_t - Q'_t) \times (Q_s e^{-i\omega(t-s)} - Q'_s e^{i\omega(t-s)})\, ds\, dt \tag{4.8}$$

Thus, $F(t-s)$ in Eq. (4.4) corresponds in this case to $e^{+i\omega(t-s)}/m\omega$ and from the definition given above $B(t-s) = (1/m\omega)\sin\omega(t-s)$.[11] Rewriting Eq. (4.8) in transform notation we have

$$\Phi_0(Q, Q') = (2\pi\hbar)^{-1}\int_0^\infty \left\{ \frac{Q'_\nu(Q_{-\nu} - Q'_{-\nu})}{-m[(\nu - i\epsilon)^2 - \omega^2]} + \frac{Q_{-\nu}(Q_\nu - Q'_\nu)}{-m[(\nu + i\epsilon)^2 - \omega^2]} \right\} d\nu \tag{4.9}$$

where

$$Q_\nu = \int_{-\infty}^{\infty} Q_t e^{-i\nu t}\, dt$$

The function $\{-m[(\nu - i\epsilon)^2 - \omega^2]\}^{-1} = (m\omega)^{-1}\int_0^\infty \sin\omega t e^{-i\nu t}\, dt$ corresponds to $1/i\nu Z_\nu$ of Eq. (4.2).[12]

[10] See ref. 2, Section 3.

[11] The finite time interval indicated by the limits T and τ can be interpreted as turning the coupling (between Q and X) on at $t = \tau$ and off at $t = T$. However, since the interaction system is to be considered in most cases as part of the steady-state environment of Q, it is really more meaningful to extend these limits over an infinite range of time ($\tau \to -\infty$, $T \to +\infty$). The possibility of allowing X to interact with Q over a finite range of time can be taken care of by giving the coupling factor (already included in the variable Q) the proper time dependence.

[12] ϵ which occurs in $i\nu Z_\nu$ is a convergence factor which was inserted in taking the Fourier transform $(1(t)/m\omega)\sin\omega t$ where $1(t)$ is the unit step function and is kept to show the location of the poles with respect to the ν axis when doing integrations of the type $\int_0^\infty H(\nu)[i\nu Z_\nu]^{-1}\, d\nu$.

Having obtained the expression for the influence phase we now turn to the classical problem of finding the response of $X(t)$ to a driving force, $f(t)$, applied at $t = 0$ with the initial conditions $X(0) = \dot{X}(0) = 0$. Starting with the Lagrangian of the unperturbed oscillator from Eq. (4.6), we add to it a potential term $-f(t)X(t)$. This potential has the same form as the coupling potential $-QX$ used in the quantum calculation. However, it is to be emphasized that the response of X to a force has nothing to do with the system Q outside of the type of coupling involved; therefore, $f(t)$ will symbolize the force in the classical problem. The complete Lagrangian is

$$L(\dot{X}, X, t) = \tfrac{1}{2}m\dot{X}^2 - \tfrac{1}{2}m\omega^2X^2 + fX \tag{4.10}$$

and the equation of motion derived from it is,

$$m\ddot{X} + m\omega^2X = f \tag{4.11}$$

Its solution under the initial conditions stated above is

$$X(t) = (m\omega)^{-1}\int_0^t f(s) \sin \omega(t - s)\, ds \tag{4.12}$$

or alternatively, in terms of Fourier transforms, is

$$X_\nu = f_\nu\{-m[(\nu - i\epsilon)^2 - \omega^2]\}^{-1} \tag{4.13}$$

Therefore, $B(t - s)$ in this case is a Green's function which yields the response of $X(t)$ to an impulse force $f(s) = \delta(s)$ and its transform yields $1/i\nu Z_\nu$. Thus a classical calculation of the ratio X_ν/f_ν under quiescent initial conditions yields the proper function for $1/i\nu Z_\nu$.

Distribution of Oscillators—Representation of Loss

The results of the preceding section are easily extended to the situation where the interaction system is a distribution of oscillators. First, we consider the case of independent oscillators coupled to the test system. It is assumed that there is a distribution of oscillators such that $G(\Omega)\, d\Omega$ is the weight of oscillators whose natural frequency is in the range between Ω and $\Omega + d\Omega$. More specifically, $G(\Omega)\, d\Omega$ is the product of the number of oscillators and the square of their coupling constants divided by the mass in $d\Omega$. Thus, we have a situation represented by the diagram of Fig. 3. Each oscillator is assumed to be initially in the ground state and finally in an arbitrary state; the coupling is again assumed to be linear. The total action is then given by

$$S[Q, X(\Omega)] = S_0(Q) + \int_{t'}^{T}\int_0^\infty G(\Omega)[\tfrac{1}{2}\dot{X}^2 - \tfrac{1}{2}X^2\Omega^2 + QX]\, d\Omega\, dt \tag{4.14}$$

For the general properties of influence functionals already described we know that when independent disturbances act on Q the influence functional is a product of the ones for each individual disturbance. Since $\mathfrak{F}_0(Q, Q') = \exp[i\Phi_0(Q, Q')]$

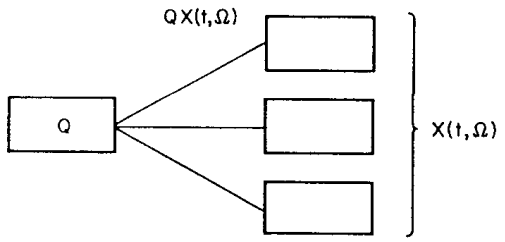


FIG. 3. Test system Q coupled to a distribution of oscillators

for the case of a single oscillator, the total influence phase for the distribution is the sum of the individual phases,

$$\Phi_0(Q, Q') = \int_0^\infty G(\Omega)\Phi_{0,\Omega}(Q, Q')\, d\Omega \tag{4.15}$$

More explicitly,

$$\Phi_0 = (2\pi\hbar)^{-1} \int_0^\infty G(\Omega)\, d\Omega \cdot \int_0^\infty \left\{ \frac{Q'_\nu(Q_{-\nu} - Q'_{-\nu})}{-[(\nu - i\epsilon)^2 - \Omega^2]} + \frac{Q'_{-\nu}(Q_\nu - Q'_\nu)}{-[(\nu + i\epsilon)^2 - \Omega^2]} \right\} d\nu \tag{4.16}$$

For this case then, the form of Eq. (4.2) is obtained if we put

$$(i\nu Z_\nu)^{-1} = \lim_{\epsilon \to 0} \int_0^\infty G(\Omega)[(\nu - i\epsilon)^2 - \Omega^2]^{-1}\, d\Omega \tag{4.17}$$

or[13]

$$(Z_\nu)^{-1} = (\pi/2)G(\nu) - i\nu \int_0^\infty G(\Omega)(\nu^2 - \Omega^2)^{-1}\, d\Omega \tag{4.18}$$

Thus the effects of all the oscillators are included in the influence phase through the expression for Z_ν, Eq. (4.17). Now, however, because of the continuous distribution of oscillators, Z_ν has a finite real part. We will now show that this real part represents dissipation by arriving at the same impedance function classically.

As before, we take the part of the Lagrangian from Eq. (4.14) having to do with the oscillators, except that the coupling potential $-Q(t)\int_0^\infty G(\Omega)X(\Omega, t)\, d\Omega$ is replaced by $-f(t)\int_0^\infty G(\Omega)X(\Omega, t)\, d\Omega$, a classical potential. $X(\Omega, t)$ is the coordinate of the oscillator in the distribution whose frequency is Ω while the total coordinate of the complete linear system with which $f(t)$ is interacting is

[13] Eq. (4.18) is obtained from (4.17) using the identity

$$\lim_{\epsilon \to 0} [(\nu - i\epsilon)^2 - \Omega^2]^{-1} = (\nu^2 - \Omega^2)^{-1} + (i\pi/2\Omega)[\delta(\nu - \Omega) - \delta(\nu + \Omega)]$$

$\int_0^\infty G(\Omega)X(\Omega, t)\,d\Omega = X(t)$. It is the relationship between $f(t)$ and $X(t)$ in which we are interested in this classical case:

$$L[\dot{X}(\Omega), X(\Omega), t] = \int_0^\infty G(\Omega)\,d\Omega[\tfrac{1}{2}\dot{X}(\Omega)^2 - \tfrac{1}{2}\Omega^2 X(\Omega)^2] + f(t)\int_0^\infty G(\Omega)X(\Omega)\,d\Omega \tag{4.19}$$

The equations of motion are the infinite set represented by

$$\ddot{X}(\Omega) + \Omega^2 X(\Omega) = f(t) \tag{4.20}$$

They result from varying L with respect to the independent variables $X(\Omega)$. For quiescent initial conditions and for $f(t)$ applied at $t = 0$, this solution is expressed

$$X_\nu(\Omega)/f_\nu = -[(\nu - i\epsilon)^2 - \Omega^2]^{-1} = [i\nu Z_\nu(\Omega)]^{-1}$$

The relation of the total coordinate X_ν to f_ν is obtained simply,

$$X_\nu/f_\nu = [\int_0^\infty X_\nu(\Omega)G(\Omega)\,d\Omega](f_\nu)^{-1} = -\int_0^\infty G(\Omega)[(\nu - i\epsilon)^2 - \Omega^2]^{-1}\,d\Omega = (i\nu Z_\nu)^{-1} \tag{4.21}$$

Referring to Eq. (4.17) it is seen that the same expression for Z_ν is obtained in the quantum and classical cases. In addition, since Z_ν is now identified with a classical impedance, the real part represents resistance while the imaginary part corresponds to reactance. Therefore, at least for the case that loss is represented by distributions of oscillators, its effect can be included in the influence functional by using the appropriate impedance expression. The spontaneous emission of a particle in free space represents a good example of such a loss mechanism. A demonstration of this point is included in Appendix II where the oscillator distribution is related to the probability of spontaneous emission starting from the influence functional representing the effect of free space.

The relationship

$$(i\nu Z_\nu)^{-1} = \int_0^\infty B(t)e^{-i\nu t}dt$$

has already been established during the course of the derivation of the influence phase for the single oscillator. Now the inverse relation between $F(t)$ and $1/i\nu Z_\nu$ can be written for the zero temperature case. In the time domain the influence phase for the distribution of oscillators is

$$i\Phi(Q, Q') = -(2\hbar)^{-1}\int_0^\infty G(\Omega)\Omega^{-1}\,d\Omega\int_{-\infty}^{\infty}\int_{-\infty}^{t}(Q_t - Q'_t)(Q_s e^{-i\Omega(t-s)} - Q'_s e^{i\Omega(t-s)})\,ds\,dt$$

Comparing this with Eq. (4.4) it is evident that

$$F(t) = \int_0^\infty G(\Omega)\Omega^{-1}e^{i\Omega t}\,d\Omega$$

But from Eq. (4.18),

$$\mathrm{Im}(i\nu Z_\nu)^{-1} = -\pi G(\nu)/2\nu$$

Therefore, it can be immediately written that

$$F(t) = -(2/\pi)\int_0^\infty \mathrm{Im}(i\nu Z_\nu)^{-1} e^{i\nu t}\, d\nu \tag{4.22}$$

as was given in Eq. (4.5b).

The results above can now be extended by a simple argument to include all linear systems composed entirely of distributions of oscillators. To do this it need only be shown that the general system can be reduced to a distribution of oscillators independently coupled to the test system, which was the situation just considered. Explicitly, suppose there exists a test system Q, coupled to an assemblage of oscillators which are also interconnected with each other. For instance, the situation might be as in Fig. 4, where each of the X_n components of the total interaction system could also represent a system of oscillators. However, it is well known[14] that such a linear system may be represented by an equivalent set of oscillators (the normal modes of the total system) independently coupled to Q.[15] Or, stated another way, the classical representation of the Lagrangian in normal modes finds new linear combinations of the X_n which makes the total Lagrangian, except for the coupling, a sum of individual quadratic forms with no cross terms. But, this same transformation of variables can be made on the expression for $\mathfrak{F}(Q, Q')$ (see Eq. (2.7)). The effect of this transformation is to change the $\mathfrak{D}X(t)$ volume by a numerical factor, since the transformation is linear.[16] Thus, in effect, we get the sum of independent systems in the quantum mechanical case also. From this argument it is concluded that the results above regarding a distribution of independent oscillators coupled to a test system, apply to any linear interaction system. Therefore, it has been found that the influence functional for all linear systems has exactly the same form $\exp[i\Phi_0(Q, Q')]$ where $\Phi_0(Q, Q')$ is a quadratic functional of the Q and Q'. $\Phi_0(Q, Q')$ is adapted to a particular linear system only through the classical response of that linear system to a force. Thus, the procedure for finding the influence functional for a linear system has been reduced to a classical problem. The fact that eliminating the coordinate of an oscillator always yields an in-

[14] This point is considered more fully in Section IV,B on classical forces.

[15] The fact that one or more of the X_n might represent continuous distributions of oscillators need not be bothersome since in principle they represent the behavior of the total system in terms of its infinite set of normal modes.

[16] The only result of such a numerical factor would be to change the normalization of $\mathfrak{F}(Q, Q')$. However, we already know that for the case that the final states of the interaction system are summed over $\mathfrak{F}(Q, Q') = 1$. Therefore the normalization of $\mathfrak{F}(Q, Q')$ is not changed by the transformation and thus is not dependent upon the coordinates chosen to represent the interaction system.

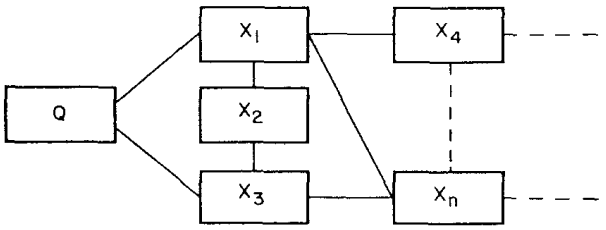


FIG. 4. Test system coupled to an arbitrary assemblage of oscillators

fluence functional which is quadratic in the potential applied to that oscillator, is a basic property of linear systems. For example, where the coupling Lagrangian is linear between an oscillator of coordinate X and another system of coordinate Q, the elimination of the X coordinate yields an influence phase which is quadratic in Q as has already been shown. If Q were the coordinate of another oscillator coupled to P, then elimination of the Q coordinates would yield an influence phase quadratic in P, etc. This can be understood mathematically by observing that the Lagrangian for all the oscillators with linear coupling is always quadratic. Doing the path integral to eliminate a coordinate is basically a process of completing the square and performing Gaussian integrals. This process of completing the square also yields quadratic terms. It is therefore not surprising that the influence phase for any linear system should be always of the same quadratic form.

It is to be emphasized that the analysis so far presented has been concerned entirely with systems whose complete behavior can be described by combinations of lossless oscillators at zero temperature. The only example of such a system is the electromagnetic field in free space. In all other physical situations linear behavior is an approximation to the actual behavior. However, this approximation may be very good over a wide range of operating conditions. In Section V the problem of approximately linear systems will be considered in detail. The results will be found to be the same as for perfect oscillators to the extent that linear behavior is realized.

Form of Influence Functionals for Linear Systems and Classical Forces as Deduced from Properties of Influence Functionals

So far, we have found the influence phase for classical potentials, uncertain classical potentials, and linear systems at zero temperature. By studying Eqs. (3.1), (3.6), and (4.4), we see that the general form for the influence functional in which all three of these were acting on Q is

$$\begin{aligned}\mathfrak{F}(Q, Q') = \exp\{&\int_\tau^T iC_1(t)(Q_t - Q'_t)\,dt - \int_\tau^T \int_\tau^t A_1(t-s)(Q_t - Q'_t) \\ &\cdot(Q_s - Q'_s)\,ds\,dt - \int_\tau^T \int_\tau^t iB_1(t-s)(Q_t - Q'_t)(Q_s + Q'_s)\,ds\,dt\}\end{aligned} \tag{4.23}$$

The exponent is written solely in terms of Q for simplicity although when the potentials are not linear in Q (as $XV(Q)$), the same general form exists, except that it is written in terms of $V(Q)$. We now observe that there are other possible combinations of the Q, Q' variables not represented here such as terms in $(Q_t + Q'_t)$, $(Q_t + Q'_t)(Q_s + Q'_s)$. To see if such terms are possible, let us form a hypothetical functional containing all possible forms up to second order in Q.

$$\mathfrak{F}(Q, Q') = \exp\{\int_\tau^T [iC_1(t)(Q_t - Q'_t) + D_1(t)(Q_t + Q'_t)]\,dt$$
$$- \int_\tau^T \int_\tau^t [A_1(t-s)(Q_t - Q'_t)(Q_s - Q'_s) + iB_1(t-s)(Q_t - Q'_t)(Q_s + Q'_s)$$
$$+ iD_2(t-s)(Q_t + Q'_t)(Q_s - Q'_s) + D_3(t-s)(Q_t + Q'_t)(Q_s + Q'_s)]\,ds\,dt\} \tag{4.24}$$

That the coefficients of the Q's inside the double integrals should be functions of $(t - s)$ is evident since the functional should not depend on the absolute time. We now will try to eliminate terms in the exponent by using the general properties of $\mathfrak{F}(Q, Q')$ given in Section II. First, we know $\mathfrak{F}(Q, Q') = \mathfrak{F}^*(Q', Q)$. This implies that all the functions A_1, B_1, C_1, D_1, D_2, and D_3 are real. Next, we know that $\mathfrak{F}(Q, Q') = 1$ if $Q'(t) = Q(t)$. Hence, D_1, D_3 are zero. This leaves only one term which we did not have before, that of $D_2(t - s)$. Now we apply the property of these functionals requiring that if $Q_t = Q'_t$ for $t > t_0$ then $\mathfrak{F}(Q, Q')$ is independent of Q for $t > t_0$. This statement is obviously true for the $C_1(t)$ and $A_1(t - s)$ terms. Consider now the $B_1(t - s)$ term. For $t > t_0$, $Q_t - Q'_t = 0$ and therefore this term is also legitimate. As for the $D_2(t - s)$ term, let us consider $t > t_0$, but $s < t_0$. Then $Q_s - Q'_s \neq 0$. Furthermore, $Q_t + Q'_t = 2Q_t \neq 0$. Therefore, $D_2(t - s)$ must also be zero. The fact that $D_2(t - s) = 0$ is actually a statement of causality, i.e., that the effect due to an applied force cannot precede the time the force was applied. To see this, let us change the limits of integration on this term

$$\int_\tau^T \int_\tau^t D_2(t-s)(Q_t + Q'_t)(Q_s - Q'_s)\,ds\,dt$$
$$= \int_\tau^T \int_t^T D_2(s-t)(Q_t - Q'_t)(Q_s + Q'_s)\,ds\,dt$$

Now the integrand is of the same form as that of the $B(t - s)$ term. However, for a fixed t the integration over s is over the range of $s > t$. This amounts to a sum over the future rather than a sum over past histories of the variable Q.

The conclusion to be drawn is that there are three possible types of terms up to second order in Q and Q' when definite classical forces, indefinite classical forces, and linear systems act on Q. Terms of this type have already been derived during the course of our analysis. Therefore, there are no major types of phenomena which have not been noticed. In the light of the above discussion we would expect the effects of additional phenomena, if they are described by terms of

second order or less in Q and Q', to be contained in one or more of the three forms of exponents shown in Eq. (4.23). For instance, in the case of a linear system at finite temperature, it will be found that the effect of temperature is to change the effective value of $A_1(t-s)$ in the exponent of Eq. (4.23) from its minimum value which occurs at zero temperature. It should be pointed out that although $A_0(t-s)$ (see Eq. (4.5b)) occurs in a term which has the form of an uncertain classical potential acting on the test system, at zero temperature one must be careful about this interpretation, for the existence of a random classical potential implies a random fluctuation of the variables of the interaction system which could induce transitions in the test system either upwards or downwards in energy. However, if the interaction system is already in its lowest state it can only induce downward transitions in the test system so that the term in $A_0(t-s)$ by itself is not sufficient. Thus, as has already been found, the exponent of the zero temperature influence functional contains two terms, one in $A_0(t-s)$ and the other in $B(t-s)$, which are related through Eq. (4.5c). Together they give the whole picture, i.e., that there is a zero point, random fluctuation of the variables of the interaction system but that this fluctuation can induce only those transitions in the test system which, through spontaneous emission, give up energy to the interaction system.

B. Influence Functionals for Driven Linear Systems

It is to be expected that if a classical force is applied to a linear interaction system which in turn is coupled to a test system, the effect of the interaction system is to modify the character of the force applied to the test system. In this section we will find the exact form for the influence functional of this effective force. *If a linear system is coupled to $Q(t)$ through one of its coordinates $X(t)$ and if a classical force $C(t)$ is coupled to another coordinate $Y(t)$, then $\Phi(Q, Q')$ representing the effect of both the linear system and the force is*

$$\Phi(Q,Q') = \Phi_0(Q,Q') + (2\pi\hbar)^{-1}\int_0^\infty \left[\frac{C_\nu(Q_{-\nu}-Q'_{-\nu})}{i\nu z_\nu} + \frac{C_{-\nu}(Q_\nu - Q'_\nu)}{-i\nu z_{-\nu}}\right] d\nu \tag{4.25}$$

where Φ_0 is the influence phase of the linear system is the absence of a classical force $C(t)$, and z_ν is a transfer impedance function which modifies the effect of C_ν on Q. The impedance z_ν is found by computing the classical response of the coordinate X to the force C with all other potentials acting on the linear system (including those due to coordinates of external systems such as Q) set equal to zero. The result of the calculation yields $i\nu z_\nu = C_\nu/X_\nu$. Alternatively, in the time domain,

$$\Phi(Q,Q') = \Phi_0(Q,Q') + \hbar^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{t}(Q_t - Q'_t)b(t-s)C_s\,ds\,dt \tag{4.26}$$

where[17]

$$(i\nu z_\nu)^{-1} = \int_0^\infty b(t)e^{-i\nu t}\,dt$$

The theorem can be stated in the form of a diagram as shown in Fig. 5. In this figure $f(t) = \int_{-\infty}^{t} F(s)b(t-s)\,ds$. It will be convenient to work in the frequency domain.

First we recall the influence phase for a classical force acting directly on Q. From this expression we will be able to identify the character of the force acting on Q in more complicated expressions. If the potential is of the form $-C(t)Q(t)$, we have, from Section III,

$$\Phi(Q, Q') = (2\pi\hbar)^{-1}\int_0^\infty [C_\nu(Q_{-\nu} - Q'_{-\nu}) + C_{-\nu}(Q_\nu - Q'_\nu)]\,d\nu \qquad (4.27)$$

Classical Potential and Q Coupled to the Same Coordinate

Before developing the general situation we consider the simpler situation where Q is coupled to a linear system through the potential $-QX$, and a force $F(t)$ is applied to the same system through the potential $-FX$. The Lagrangian for the complete system is

$$L\ (\text{system}) = L_0(\dot{Q}, Q, t) + (F + Q)X + L(\dot{X}, \dot{Y}, \cdots X, Y, t) \qquad (4.28)$$

where $X, Y, \cdots$ represent all the coordinates of the linear system. If $F = 0$,

$$\Phi(Q, Q') \equiv \Phi_0(Q, Q') = (2\pi\hbar)^{-1}\int_0^\infty \left[\frac{Q'_\nu(Q_{-\nu} - Q'_{-\nu})}{(i\nu Z_\nu)} + \frac{Q_{-\nu}(Q_\nu - Q'_\nu)}{(-i\nu Z_{-\nu})}\right] d\nu \qquad (4.29)$$

If $F \neq 0$ it is evident from Eq. (4.28) that the required influence phase can be found by replacing Q by $Q + F$ and Q' by $Q' + F$. Notice that F does not carry the prime notation since it is not a coordinate. If this substitution is made in Eq. (4.29) we have

$$\Phi(Q, Q') = (2\pi\hbar)^{-1}\int_0^\infty \left[\frac{(Q'_\nu + F_\nu)(Q_{-\nu} - Q'_{-\nu})}{i\nu Z_\nu} + \frac{(Q_{-\nu} + F_{-\nu})(Q_\nu - Q'_\nu)}{(-i\nu Z_{-\nu})}\right] d\nu = \Phi_0(Q, Q') + (2\pi\hbar)^{-1} \cdot \int_0^\infty \left[\left(\frac{F_\nu}{i\nu Z_\nu}\right)(Q_{-\nu} - Q'_{-\nu}) + \left(\frac{F_{-\nu}}{-i\nu Z_{-\nu}}\right)(Q_\nu - Q'_\nu)\right] d\nu \qquad (4.30)$$

[17] The notation z_ν, $b(t-s)$ was chosen to avoid confusion with Z_ν and $B(t-s)$ which are the impedance and response function respectively of the linear system as seen by the test system.

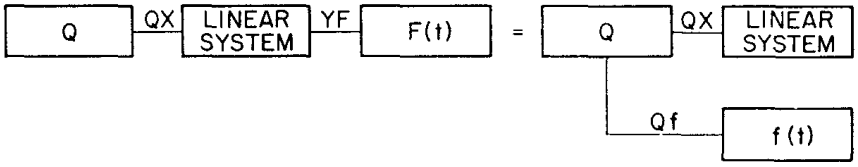


FIG. 5. Equivalent influences of a linear system and a force, $F(t)$, acting on a test system, Q.

As might be expected the total effect of the linear system and the driving force consists of two separate terms, one describing the effect of the linear system alone, and the other describing the effect of the driving force. Comparison of Eqs. (4.26) and (4.30) shows that the effective force applied to Q is, in transform language, $F_\nu/i\nu Z_\nu$ and, further, shows that F_ν is modified by $1/i\nu Z_\nu$, the classical impedance function of the interaction system. In this special example where F and Q are both coupled to the same coordinate, Z_ν is both the correct impedance to be used in $\Phi_0(Q, Q')$, i.e., that impedance seen by the test system, and is the transfer impedance z_ν which modifies F_ν. This is not true generally as we shall see in the next section. In addition, it is interesting to observe that no unexpected quantum effects appear because of the addition of a force to the interaction system. The only effect of the interaction system is to modify the characteristics of $F(t)$ in an entirely classical way.

Classical Forces Acting Through a General Linear System

Having obtained an idea of the type of results to expect in the above simplified analysis we now proceed to the more general case. Let the N coordinates of the interaction system be represented by X_i, $i = 1 \cdots N$. Its coupling to the test system, $Q(t)$, and to the driving force $C(t)$ is given by the potentials $-X_nQ$ and $-X_kC$, respectively. Thus we are assuming, for simplification in writing, that Q is coupled only to the variable X_n and the force $C(t)$ is applied just to the variable X_k. Again the interaction system is assumed to be composed entirely of harmonic oscillators. The Lagrangian is

$$L \text{ (system)} = L(\dot{Q}, Q, t) + \sum_{i,j} [\tfrac{1}{2}(T_{ij}\dot{X}_i\dot{X}_j - V_{ij}X_iX_j)] + X_nQ + X_kC \tag{4.31}$$

It is well known in the theory of linear systems that new coordinates may be defined by means of a linear transformation of the X_i. These new coordinates will be chosen as the eigenvectors, Y_l, of the interaction system (*10*). Thus,

$$X_i = \sum_{l=1}^{n} a_{il}Y_l \qquad i = 1, 2, \cdots n$$

Assuming the a_{il} to be properly normalized, the Lagrangian may be rewritten as follows

$$L(\dot{Q}, Q, \dot{Y}_l, Y_l, t) = L_0(\dot{Q}, Q, t) + \sum_l [\tfrac{1}{2}(\dot{Y}_l^2 - \omega_l^2 Y_l^2) + Y_l(a_{nl}Q + a_{kl}C)] \tag{4.32}$$

Since these are now independent oscillators coupled to Q the influence phase can be written down immediately,

$$\Phi(Q, Q') = \sum_l \Phi_l(Q, Q') = \sum_l \frac{1}{2\pi\hbar} \int_0^\infty \left[Q'_\nu (Q_{-\nu} - Q'_{-\nu}) \left(\frac{a_{nl}^2}{i\nu Z_l(\nu)} \right) + Q_{-\nu}(Q_\nu - Q'_\nu) \left(\frac{a_{nl}^2}{-i\nu Z_l(-\nu)} \right) \right] d\nu + \sum_l \frac{1}{2\pi\hbar} \cdot \int_0^\infty \left[\frac{a_{nl}\, a_{kl}\, C_\nu}{i\nu Z_l(\nu)} (Q_{-\nu} - Q'_{-\nu}) + \frac{a_{nl}\, a_{kl} C_{-\nu}}{-i\nu Z_l(-\nu)} (Q_\nu - Q'_\nu) \right] d\nu \tag{4.33}$$

where

$$\frac{1}{i\nu Z_l(\nu)} = \left. \frac{Y_l(\nu)}{a_{nl}\, Q_\nu} \right|_{C_\nu = 0} \equiv \left. \frac{Y_l(\nu)}{a_{kl}\, C_\nu} \right|_{Q_\nu = 0} \tag{4.34}$$

calculated classically.

This can be written in the form of Eq. (4.25) if we make the correspondence

$$\frac{1}{i\nu Z_\nu} = \sum_l \frac{a_{nl}^2}{i\nu Z_l(\nu)}$$

and

$$\frac{1}{i\nu z_\nu} = \sum_l \frac{a_{nl}\, a_{kl}}{i\nu Z_l(\nu)} \tag{4.35}$$

Using Eqs. (4.34) and (4.35) we can now show that $1/i\nu Z_\nu$ and $1/i\nu z_\nu$ are equivalent to $X_n(\nu)/Q_\nu$ and $X_n(\nu)/C_\nu$ respectively. Thus

$$\frac{1}{i\nu Z_\nu} = \sum_l \frac{a_{nl}^2}{i\nu Z_l(\nu)} = \sum_l a_{nl}^2 \left(\frac{Y_l(\nu)}{a_{nl}\, Q_\nu} \right) \Bigg|_{C_\nu = 0} = \sum_l \left. \frac{a_{nl}\, Y_l(\nu)}{Q_\nu} \right|_{C_\nu = 0} = \left. \frac{X_n(\nu)}{Q_\nu} \right|_{C_\nu = 0} \tag{4.36}$$

and

$$\frac{1}{i\nu z_\nu} = \sum_l \frac{a_{nl}\, a_{kl}}{i\nu Z_l(\nu)} = \sum_l a_{nl}\, a_{kl} \left. \frac{Y_l(\nu)}{a_{kl}\, C_\nu} \right|_{Q_\nu = 0} = \left. \frac{\sum_l a_{nl}\, Y_l(\nu)}{C_\nu} \right|_{Q_\nu = 0} = \left. \frac{X_n(\nu)}{C_\nu} \right|_{Q_\nu = 0} \tag{4.37}$$

Equation (4.36) is a mathematical expression of the argument used earlier to find the appropriate impedance function to be used in the influence functional for a linear system acting on Q. The additional information obtained here in this regard is that when other forces are present they are to be set equal to zero when this computation is made. Equation (4.37) states the new result that the transfer impedance which modifies C_ν in its effect on the test system is to be found by computing the ratio of C_ν to the coordinate $X_n(\nu)$ to which the test system is coupled. The total force acting on the test system when several forces are acting on the interaction system is simply the sum of these forces each modified by the appropriate transfer impedance determined in the above described manner.

C. Linear Systems at Finite Temperatures

The forms of influence functionals which are possible for linear systems have been established by an argument which utilized the general properties discussed in Section II. Each of these forms has already occurred in the analyses of classical potentials, random potentials, and zero temperature linear systems. Therefore, the results to be expected here are one or more of the forms already obtained.

The discussion is begun again with a single oscillator as our linear system, for simplicity. From this, the extension to distributions of oscillators is immediate, as it was in Section IV.A. The complete problem is set up in the same way as for zero temperature except that the initial state of the oscillator is not simply the ground state or any definite eigenstate. The effect of temperature is to make the initial state uncertain and it is properly represented by a sum over all states weighted by the Boltzmann factor $e^{-\beta E_n}$ where $\beta = 1/kT$, T being the temperature in this case. The final state is again arbitrary: therefore, a formal expression for the influence functional is

$$\begin{aligned}\mathfrak{F}(Q, Q') = \int \delta\,(X_T - X'_T) K_Q(X_T, X_\tau) K^*_{Q'}(X'_T, X'_\tau) \\ \cdot \sum_n N^{-1} e^{-\beta E_n} \phi_n(X_\tau) \phi_n^*(X'_\tau)\, dX_\tau \cdots dX'_T\end{aligned} \tag{4.38}$$

where N, the normalization constant, is $\sum_n e^{-\beta E_n}$. The ϕ_n represent the energy eigenfunctions of the oscillator unperturbed by external forces. The first problem is to find a closed form for the expression $\sum_n \phi_n(X)\phi_n(X')e^{-\beta E_n}$. This can be done by noticing that its form is identical with the kernel which takes a wave function from one time to another if we make the correspondence that β represents an imaginary time interval. If the times involved are t_2 and t_1, this kernel is

$$\begin{aligned}K_0(X_2, X_1) &= \sum_n \phi_n(X_2)\phi_n^*(X_1) \exp\left[-(i/\hbar)E_n(t_2 - t_1)\right] \\ &= \exp\left[(i/\hbar)S_{\mathrm{cl}}\right]\end{aligned} \tag{4.39}$$

for the harmonic oscillator, and where the subscript 0 indicates the absence of external forces. For the harmonic oscillator the expression for S is easily obtained in terms of the initial and final positions X_1 and X_2 (2). Thus,

$$S_{c1} = m\omega[2 \sin \omega(t_2 - t_1)]^{-1}[(X_1^2 + X_2^2) \cos \omega(t_2 - t_1) - 2X_1X_2] \quad (4.40)$$

Utilizing Eqs. (4.42) and (4.43) and making the correspondence $\beta = i(t_2 - t_1)/\hbar$, $X_1 = X_\tau$, and $X_2 = X_\tau'$, we find that

$$\begin{aligned} &\sum \phi_n(X_\tau)\phi_n^*(X'_\tau) \exp(-\beta E_n) \\ &\quad = \exp\{-m\omega[2\hbar \sinh(\beta\hbar\omega)]^{-1}[(X_\tau^2 + X_\tau'^2) \cosh(\beta\hbar\omega) - 2X_\tau X'_\tau]\} \end{aligned} \quad (4.41)$$

Using this closed expression for the average initial state of the oscillator and the kernel for the driven harmonic oscillator (Eq. (4.7b)) the influence functional can be evaluated by evaluation of a series of Gaussian integrations just as in the zero temperature case. The result expressed in the frequency domain is,

$$\begin{aligned} i\Phi(Q, Q') &= i\Phi_0(Q, Q') \\ &\quad - (\pi\hbar^2)^{-1} \int_0^\infty \pi\hbar[2m\omega(e^{\beta\hbar\omega} - 1)]^{-1}\delta(\nu - \omega)|Q_\nu - Q'_\nu|^2 \, d\nu \end{aligned} \quad (4.42)$$

Thus, the influence phase is made up of two terms, the first of which is the effect of the oscillator at zero temperature. The second is recognized as having the same form which was found for an uncertain classical potential with a Guassian distribution. Therefore, the effect of finite temperature is to introduce a noisy potential acting on Q at the frequency of the original oscillator. The power spectrum of the noise produced by the finite temperature is

$$\phi(\nu) = \pi\hbar[2m\nu(e^{\beta\hbar\nu} - 1)]^{-1}\delta(\nu - \omega) \quad (4.43)$$

To indicate more clearly the relationship of $\phi(\nu)$ to the characteristics of the linear system it is instructive to extend this expression to the case of a distribution of oscillators $G(\Omega)$ all at the same finite temperature. The resulting influence phase is

$$\begin{aligned} i\Phi(Q, Q') &= i \int_0^\infty \Phi_0(Q, Q')G(\Omega) \, d\Omega \\ &\quad + (\pi\hbar^2)^{-1} \int_0^\infty \pi\hbar G(\nu)[2\nu(e^{\beta\hbar\nu} - 1)]^{-1} |Q_\nu - Q'_\nu|^2 \, d\nu \end{aligned} \quad (4.44)$$

where in the distribution m has been set equal to unity. The first term is again the influence phase for zero temperature, while the second term again has the form of a noisy potential whose power spectrum

$$\phi(\nu) = (\hbar\pi/2)G(\nu)[\nu(e^{\beta\hbar\nu} - 1)]^{-1}$$

Recalling the analysis of the distribution of oscillators, it is found from Eq. (4.18) that $\pi G(\nu)/2 = \mathrm{Re}(1/Z_\nu)$. Therefore, the power spectrum can be written

$$\phi(\nu) = \hbar \, \mathrm{Re}\,(1/Z_\nu)[\nu(e^{\beta\hbar\nu} - 1)]^{-1} \quad (4.45)$$

In the time domain the influence phase is

$$i\Phi = i\Phi_0 - (\pi\hbar^2)^{-1}\int_0^\infty d\nu G(\nu)\hbar[\nu(e^{\beta\hbar\nu}-1)] \times \int_{-\infty}^{\infty}\int_{-\infty}^{t}(Q_t - Q'_t)(Q_s - Q'_s)\cos\nu(t-s)\,ds\,dt \tag{4.46}$$

Comparing this with Eq. (3.6) for random classical forces we see that the correlation function of the noise due to the finite temperature is

$$\begin{aligned} R(t-s) &= \int_0^\infty \hbar G(\nu)[\nu(e^{\beta\hbar\nu}-1)]^{-1}\cos\nu(t-s)\,d\nu \\ &= -(2/\pi)\int_0^\infty \hbar\,\mathrm{Im}\,(i\nu Z_\nu)^{-1}(e^{\beta\hbar\nu}-1)^{-1}\cos\nu(t-s)\,d\nu \end{aligned} \tag{4.47}$$

Finally, if we write Eq. (4.46) in terms of $F(t-s)$ as, for instance, in Eq. (4.4), we find that

$$F(t) = \int_0^\infty [G(\nu)/\nu]\coth(\beta\hbar\nu/2)\cos\nu t\,d\nu + i\int_0^\infty [G(\nu)/\nu]\sin\nu t\,d\nu$$

Thus,

$$\begin{aligned} A(t) &= -(2/\pi)\int_0^\infty \mathrm{Im}\,(i\nu Z_\nu)^{-1}\coth(\beta\hbar\nu/2)\cos\nu t\,d\nu \\ &= -(2/\pi)\int_0^\infty \mathrm{Im}\,(i\nu Z_\nu)^{-1}[1+2(e^{\beta\hbar\nu}-1)^{-1}]\cos\nu t\,d\nu \end{aligned} \tag{4.48}$$

and

$$B(t) = -(2/\pi)\int_0^\infty \mathrm{Im}\,(i\nu Z_\nu)^{-1}\sin\nu t\,d\nu$$

which are the more general counterparts of Eq. (4.5b). Notice, however, that only the relation for $A(t)$ changes with temperature.

Thus, if a linear interaction system is initially at a finite temperature, the resulting effect is the same, as far as the test system is concerned, as if the linear system were at zero temperature and, in addition, a random classical potential were connected independently to the test system. The power spectrum of the random potential is given by Eq. (4.45) and is related both to the temperature and to the dissipative part of the impedance of the linear system. The theorem is stated in terms of a diagram in Fig. 6 where the power spectrum $\phi(\nu)$ of the random force C_ν is defined by

$$\phi(\nu) = 4\pi\langle C_\nu C_{-\nu'}\rangle\delta(\nu+\nu') \tag{4.49}$$

This fluctuation dissipation theorem has a content which is different from those stated by Callen and Welton (*11*), Kubo (*12*), and others. It represents still another generalization of the Nyquist theorem which relates noise and resistance in electric circuits (*13*). These previously stated fluctuation dissipation theorems related the fluctuations of some variable in an isolated system, which is initially at thermal equilibrium, to the dissipative part of the impedance of the isolated system. This would be equivalent in our case to relating $\langle Q^2\rangle$ when the test system

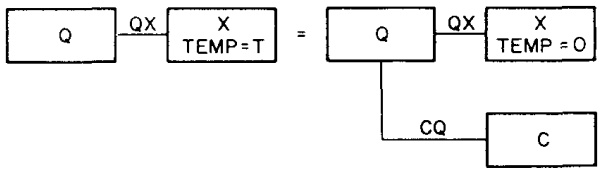


FIG. 6. Equivalent influences of a linear system at a finite temperature acting on a test system.

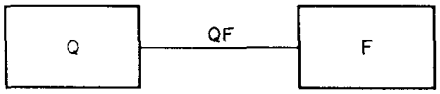


FIG. 7. Linear classical force F acting on a test system

is at equilibrium ($F = 0$) to the dissipative part of F_ν/Q_ν where F is a classical force acting on Q through a linear potential as shown in Fig. 7. However, we have shown that the effect of an external quantum system at thermal equilibrium on a test system can be separated into two effects, a zero point quantum term, which cannot be classified as pure noise, and a random potential term. Using this influence functional approach we can find $\langle Q^2 \rangle$ due not only to the internal fluctuations of Q but also due to the effect of X.

The fact that all the derivations so far have been exact, which is a consequence of dealing only with systems made up of distributions of oscillators, brings up two interesting aspects of the theory. The first one is that Z_ν does not depend on the temperature, only on the distribution of oscillators. Yet in any real finite system, it is to be expected that the temperature of a system does affect its impedance. The second aspect is that when a force $f(t)$, of any magnitude whatever, is applied to the linear system, its temperature does not change although it is obvious that if z_ν has a finite real part the linear system must absorb energy. For example, we have shown that the effect on a test system of a linear system at a finite temperature acted on by a force $f(t)$ is the same as the effect of a linear system at zero temperature, a force dependent only on the temperature of the linear system, and the force $f(t)$ modified by the transfer characteristics of the linear system. Figure 8 shows the situation where $f_\nu = F_\nu/i\nu z_\nu$ and $C(t)$ is the random force with a power spectrum given in Eq. (4.49). The influence phase acting on Q is

$$i\Phi = i\Phi_0 + i(2\pi\hbar)^{-1} \int_{-\infty}^{\infty} (f_\nu/i\nu z_\nu)(Q_{-\nu} - Q'_{-\nu})\, d\nu$$
$$- (\pi\hbar^2)^{-1} \int_0^{\infty} \phi(\nu) |Q_\nu - Q'_\nu|^2\, d\nu$$

As can be seen, the addition of the driving force $f(t)$ (which is denoted in the above expression by its Fourier transform f_ν) does not change either the imped-

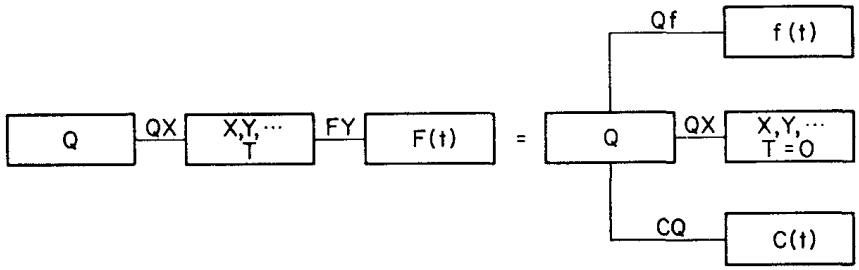


FIG. 8. Equivalent influences of a linear system at a finite temperature and a force acting on Q.

ance characteristics or the temperature of the interaction system (which would be reflected as changes in $\Phi_0(Q, Q')$, z_ν, and $\phi(\nu)$).

The fact that the impedance characteristics of the interaction system are not temperature dependent is a direct result of linearity. The incremental response of a perfectly linear system due to a driving force is independent of its initial state of motion or of that induced by other linearly coupled forces. Therefore, the random motion of the coordinate of the interaction system implied by a finite temperature does not affect its response to a driving force.

To see the reason that temperature of the linear interaction system has no dependence on the applied force $F(t)$, let us consider a linear system represented by a very large box whose dimensions we will allow to become infinite. In addition, we assume that a test system is located in the box and is receiving portions of a classical (very large amplitude) electromagnetic wave which is being transmitted by an antenna also within the box. As long as the box is of finite size and has lossless walls, any signal transmitted by the antenna will find its way either to the test system or back to the antenna. Transmission of the signal will be effected through the modes of the cavity which have a discrete frequency distribution. This system exhibits no loss since each mode represents a lossless harmonic oscillator. For the rest of the discussion it is convenient to separate the energy transmitted by the antenna into parts which reach the test system directly or through reflection from the walls. As the dimensions of the box are allowed to become infinite (i.e., the distribution of oscillators describing the electromagnetic behavior of the box becomes continuous), the time required for the energy to reach the test system by reflection from the walls also becomes infinite. Therefore, since part of the energy is lost from the antenna and test systems, the box has, in effect, dissipation. However, because of the volume of the box is infinite, its temperature is not changed by this lost energy. In other words, the box has an infinite specific heat. In addition, the energy transmitted by the antenna will generally have different characteristics from that of the Gaussian noise associated with temperature and even if the average background

energy content of the box were changed it could not be properly described by a temperature parameter.

All the results so far suggest that any linear system can be handled by the same rules that have been developed for systems of oscillators. This will be developed fully in the next section. However, we will assume this to be true now and conclude this section by applying the theorem just derived to obtain Nyquist's result for noise from a resistor. Take as an example an arbitrary circuit as the test system connected to a resistor $R_T(\nu)$ at a temperature T as shown in Fig. 9. The resistor comprises the interaction system. The interaction between the test system and the resistor is characterized by a charge $Q(t)$ flowing through the test system and resistor and a voltage $V(t)$ across the terminals. Let us associate $Q(t)$ with the coordinate of the test system and $V(t)$ with the coordinate of $R(\nu)$. The interaction part of the Lagrangian is symbolically represented by $-Q(t)V(t)$ since the current voltage relationship in $R(\nu)$ is opposite to that of a generator. The quantity $i\nu Z_\nu$ appearing in the influence functional is given by

$$-[Q_\nu/V_\nu] = -(i\nu R_\nu)^{-1} = i\nu Z_\nu$$

Thus, it follows that

$$\text{Re}\,(Z_\nu)^{-1} = \nu^2 R_\nu$$

Then the results of this section tell us that this situation, as shown in Fig. 9, may be replaced by a resistor at zero temperature (i.e., a resistor with thermal fluctuations appropriate to zero temperature but with the same magnitude of resistance it has at temperature T) and a random classical voltage whose power spectrum is

$$\phi(\nu) = \hbar\nu R_\nu(e^{\beta\hbar\nu} - 1)^{-1} \tag{4.50}$$

as shown in Fig. 10.

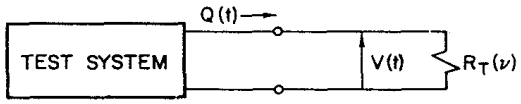


FIG. 9. Test system acted on by a linear system represented by an electrical resistance R at a finite temperature.

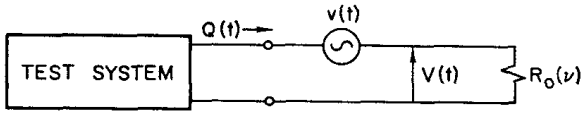


FIG. 10. Equivalent interaction system of a resistor at finite temperature acting on the test system Q.

The mean value of this voltage is

$$\begin{aligned}\langle v^2(t)\rangle &= (2/\pi)\int_0^\infty \phi(\nu)\,d\nu \\ &= \int_0^\infty 2\hbar\nu R_\nu[\pi(e^{\beta\hbar\nu}-1)]^{-1}\,d\nu\end{aligned} \tag{4.51}$$

For high temperature $\beta \ll 1$, and we find that over the frequency range where $\beta\hbar\nu \ll 1$, $\phi(\nu) = kTR_\nu$. If R_ν is constant over the frequency range of interest whose limits are ν_2 and ν_1, then as a result of the noise power in this range

$$\langle v^2(t)\rangle = 4kTR(f_2 - f_1) \tag{4.52}$$

where f is the circular frequency. This is the famous Nyquist result for noise from a resistor. Notice that the noise voltage generator must be placed in series with the resistor as a consequence of interacting directly with the coordinate of the test system, which in this case is $Q(t)$, the charge flowing in the circuit. If the coordinates were chosen such that $Q(t)$ were the coordinate of the resistor and $V(t)$ that of the test system, then the noise generator would become a current source interacting with the voltage $V(t)$ so that the situation would be as shown in Fig. 11, where $\langle i^2\rangle = (4kT/R)(f_2 - f_1)$ in the high temperature limit.

It is worth mentioning, but obvious from the derivation, that if there were many sources of dissipation coupled to the test system, each at different temperatures, then there would be a fluctuating potential associated with each source of dissipation with a power spectrum characteristic to the temperature involved. The case of different temperatures represents a nonequilibrium condition in that the hot resistors are always giving up energy to the cold ones. However, when the resistors are represented by continuous distributions of oscillators as in the case of an infinite box (free space), the temperatures do not change because of the infinite specific heat of the ensemble of oscillators.

V. WEAKLY COUPLED SYSTEMS

We are now faced with the problem of finding influence functionals whose behavior is in some sense linear but whose total behavior is not representable by systems of perfect oscillators. There are many examples of this. The concepts of resistance, electric and magnetic polarizations, etc. are basic quantities which characterize the classical electrical behavior of matter. However, for an accurate description of this behavior, these quantities can be constant, independent of the applied electromagnetic field only in the range of approximation that the

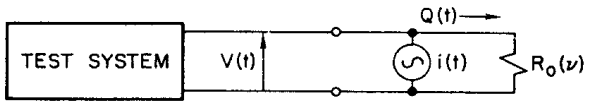


FIG. 11. Equivalent interaction system of a resistor at finite temperature when the coordinate associated with the resistor is the charge flowing through it.

magnitude of this applied field does not become too large. We will consider here the analysis of such systems and the approximations which make linear analysis valid. Insofar as linear behavior is obtained, the results of this section are basically the same as those obtained in previous sections with regard to finding influence functionals. However, it is interesting to see the same principles come out of the analysis another way. In addition, it will be found that expressions corresponding to Z_ν which appear in the influence functional are actually closed forms which can be used to compute such quantities as the conductivity from a knowledge of the unperturbed quantum characteristics of a system. These expressions have been derived before by several authors but it is interesting to find that they also appear in the influence functional quite naturally. The results will then be applied to the case of a beam of non-interacting particles passing near a test system such as a cavity. This analysis naturally lends itself to a discussion of noise in beam-type maser amplifiers. In Appendix III, the results of this and Section IV are used to compute the spontaneous emission of a particle in a cavity.

A. Interaction Systems with Coupling Potentials of the Form $-V(Q)U(P)$

The specific result to be shown here can be stated as follows: If a general interaction system, P, is coupled to a test system Q so that the interaction potential is small and of the form $-V(Q)U(P)$, then the effect of the test system is that of a sum of oscillators whose frequencies correspond to the possible transitions of the interaction system. Therefore, to the extent that second order perturbation theory yields sufficient accuracy, the effect of an interaction system is that of a linear system.

To show this, we shall first assume the interaction system to be in an eigenstate $\phi_a(P_\tau) \exp\{-(i/\hbar)E_a\tau\}$ at the beginning of the interaction and in an arbitrary state at the end of the interaction consistent with the usual procedure we have followed. Also, for convenience in writing, the interaction potential will be assumed $-U(P)Q$. The influence functional is then

$$\begin{aligned}\mathfrak{F}(Q, Q') = \int \delta(P_T - P'_T) \exp\{(i/\hbar)[S(P) - S(P') \\ + \int_\tau^T (QU - Q'U')\, dt]\}\phi_a^*(P'_\tau)\phi_a(P_\tau)\, dP_\tau \cdots \mathfrak{D}P'(t)\end{aligned} \tag{5.1}$$

where in the above we have written U for $U(P)$ and U' for $U(P')$. Since the magnitude of the interaction is assumed to be small, the perturbation approach can be used to good advantage. Thus, expanding the interaction part of the exponent and keeping terms to second order in Q only,

$$\begin{aligned}\mathfrak{F}(Q, Q') = \int \delta(P_T - P'_T) \exp\{(i/\hbar)[S(P) - S(P')]\} \\ \times \{1 + (i/\hbar)\int_\tau^T (QU - Q'U')\, dt + (i/\hbar)^2 \int_\tau^T \int_\tau^t (Q_tU_t - Q'_tU'_t) \\ \times (Q_sU_s - Q'_sU'_s)\, ds\, dt\}\phi_a^*(P'_\tau)\phi_a(P_\tau)\, dP_\tau \cdots \mathfrak{D}P'(t)\end{aligned} \tag{5.2}$$

The evaluation of this is done in an entirely straightforward manner and the following is obtained,[18]

$$\mathfrak{F}(Q, Q') = 1 + (iU_{aa}/\hbar) \int_\tau^T (Q_t - Q'_t)\, dt - \sum_b |U_{ab}/\hbar|^2 \times \int_\tau^T \int_\tau^t (Q_t - Q'_t)\{Q_s \exp[-i\omega_{ba}(t - s)] - Q'_s \exp[i\omega_{ba}(t - s)]\}\, ds\, dt \tag{5.3}$$

where

$$U_{ab} = \int \phi_a^*(P) U \phi_b(P)\, dP \quad \text{and} \quad \omega_{ba} = (E_b - E_a)/\hbar$$

All terms which involve U_{aa} will be disregarded. This is because U_{aa} represents the average value of the operator $U(P)$ in an eigenstate of the interaction system alone. Even if it is not zero, it will be noted from Eq. (5.2) that terms involving U_{aa} can be written

$$1 + (iU_{aa}/\hbar) \int_\tau^T (Q_t - Q'_t)\, dt + (1/2!)[(iU_{aa}/\hbar) \int_\tau^T (Q_t - Q'_t)\, dt]^2 + \cdots = \exp\{(i/\hbar) \int_\tau^T U_{aa}(Q_t - Q'_t)\, dt\}$$

When $\mathfrak{F}(Q, Q')$ is used to make a calculation on the test system, this term has the effect of adding a constant potential $V(Q(t)) = -U_{aa}Q(t)$ to the unperturbed test system. Disregarding U_{aa}, the influence functional then becomes

$$\mathfrak{F}(Q, Q') = 1 + \alpha_a \tag{5.4}$$

where

$$\alpha_a = \left(\frac{1}{\hbar}\right) \int_{-\infty}^{\infty} \int_{-\infty}^{t} \gamma_t \gamma_s (Q_t - Q'_t)[Q_s F_a^*(t - s) - Q'_s F_a(t - s)]\, ds\, dt$$

and

$$F_a(t - s) = \sum_b (|U_{ba}|^2/\hbar) \exp[i\omega_{ba}(t - s)]$$

where the limits have been extended and the factors $\gamma_t \gamma_s$ inserted to allow finite

[18] A typical term in Eq. (5.2) is as follows:

$$\int \delta(P_T - P'_T) \exp\{(i/\hbar)[S(P) - S(P')]\} (i/\hbar)^2 \int_\tau^T \int_\tau^t Q_t U_t Q_s U_s\, ds\, dt \times \phi_a^*(P'_\tau)\phi_a(P_\tau)\, dP_\tau \cdots \mathfrak{D}P(t)$$

Taking the time integrations outside the path integral and replacing the path integrations by propagating kernels, this becomes

$$(i/\hbar)^2 \int_\tau^T \int_\tau^t Q_t Q_s\, ds\, dt \int \delta(P_T - P'_T) K^*(P'_T, P'_\tau) K(P_T, P_t) \cdot U_t K(P_t, P_s) U_s K(P_s, P_\tau) \phi_a^*(P'_\tau) \phi_a(P_\tau) dP_\tau \cdots dP'_T$$

Remembering that $K(P_T, P_t) = \sum \phi_n(P_T) \phi_n^*(P_t) \exp\{-iE_n(T - t)/\hbar\}$, this expression becomes simply,

$$-\sum_b |U_{ab}/\hbar|^2 \int_\tau^T \int_\tau^t Q_t Q_s \exp[-i\omega_{ba}(t - s)]\, ds\, dt$$

coupling time if necessary. If the strength of coupling is sufficiently weak then Eq. (5.4) can be rewritten

$$\mathfrak{F} \approx \exp(\alpha_{\iota}) \tag{5.5}$$

In this form we recognize $\mathfrak{F}(Q, Q')$ as that describing the effect of a sum of harmonic oscillators independently connected to the test system each of whose "weights" is $U_{ba}U_{ab}/\hbar^2$. The complete response function for the system of oscillators is

$$B_a(t-s) = \operatorname{Im} F_a(t-s) = \sum_b (2|U_{ba}|^2/\hbar) \sin \omega_{ba}(t-s)$$

where the subscript a on B_a refers to the initial eigenstate. According to previous definition, the mass of each individual oscillator is identified by $m = \hbar(2|U_{ba}|^2\omega_{ba})^{-1}$ and its characteristic frequency by ω_{ba}. Therefore, to the extent that second order perturbation theory yields satisfactory accuracy, any system may be considered as a sum of harmonic oscillators. This is equivalent in classical mechanics to the theory of the motion of a particle having small displacements around an equilibrium position. Its motion, to a first approximation, is also that of a harmonic oscillator, if the first effective term in a power series expansion of the potential around that equilibrium position is quadratic in the displacement.

In this part of the discussion we should again point out the motivation for writing the approximate influence function, Eq. (5.4), in terms of the approximate exponential of Eq. (5.5), apart from the obvious advantage of making the form agree with that of exactly linear systems. Frequently, we deal with a test system which is influenced by another system which is actually made up of a large number of very small systems. Examples of such an interaction system would be a beam of atoms or the electrons in a metal. Although the expression for the influence functional for any one of the subsystems is only good to second order, their individual effects are so small that this accuracy is very good and Eqs. (5.4) or (5.5) is equally valid. However, when the sum of the effects of the subsystems is not small, then the two forms above do not describe the situation equally insofar as the composite effect of the interaction system is concerned. We know that when these subsystems are dynamically and statistically independent, the total influence functional is simply a product of the individual ones. In such a case the influence functional obtained by using Eq. (5.5) as follows

$$\mathfrak{F}_k(Q, Q') \approx \exp\{i\textstyle\sum_k \Phi_k(Q, Q')\}$$

yields much greater accuracy than that obtained from Eq. (5.4) where we would find

$$\mathfrak{F}(Q, Q') \approx 1 + i\textstyle\sum_k \Phi_k(Q, Q')$$

$\Phi_k(Q, Q')$ being the influence phase for the kth subsystem (see Appendix IV).

The expression given by Eq. (5.5) has additional implications which are not immediately apparent from the analogy with the harmonic oscillator. No assumption was made as to whether the initial state was necessarily the lowest of the system. Therefore, ω_{ba} could be either positive or negative. Suppose for a moment that the interaction system has only two states ($\phi_a(P)$ and $\phi_b(P)$) and that the initial state, a, is the lower one ($\omega_{ba} > 0$). It is obvious that the only effect it can have on the test system Q is to absorb energy from Q. However, if the initial state, a, is the upper state (corresponding to $\omega_{ba} < 0$) then the interaction system can only give up energy to the test system. It can do this in two ways, through spontaneous emission or through coherent emission due to some coherent driving force exerted on it by the test system. So for the case $\omega_{ba} > 0$ we expect the influence functional to show that the interaction system has the effect of a cold system characterized by a dissipative impedance (or positive resistance). Conversely, for $\omega_{ba} < 0$ it is expected that $\mathfrak{F}(Q, Q')$ will be characterized by a negative resistance and a random potential due to the spontaneous emission transitions. This situation is made more obvious if we translate Eq. (5.5) into transform notation. Thus,

$$\Phi(Q, Q') = (2\pi\hbar)^{-1} \int_0^\infty \left[\frac{Q'_\nu (Q_{-\nu} - Q'_{-\nu})}{(i\nu Z_{ba,\nu})} + \frac{Q_{-\nu}(Q_\nu - Q'_\nu)}{(-i\nu Z_{ba,-\nu})} \right] d\nu - (\pi\hbar^2)^{-1} \int_0^\infty |Q_\nu - Q'_\nu|^2 \pi |U_{ba}|^2 \delta(\nu + \omega_{ba})\, d\nu \tag{5.6}$$

where

$$i\nu Z_{bc,\nu} = -\hbar(2\omega_{ba}|U_{ba}|^2)^{-1}[(\nu - i\epsilon)^2 - \omega_{ba}^2] \tag{5.7}$$

First of all we notice that the sign of $Z_{ba,\nu}$ changes with that of ω_{ba} and therefore its dissipative part can be positive or negative as was argued above. Secondly, for $\omega_{ba} < 0$ there is a random potential acting on Q whose power spectrum is given by $\phi(\nu) = \pi|U_{bc}|^2\delta(\nu + \omega_{ba})$. Of course when $\omega_{ba} > 0$, the integral involving this term disappears indicating the noise potential does not exist and the effect is the same as that of a harmonic oscillator initially in the ground state with $1/2m\omega$ identified with $|U_{ba}|^2/\hbar$.

In a real physical situation it is not likely that the interaction system will be in a definite state initially. So to extend the above results to a more general case we assume that the initial state is described as a sum over states weighted by a density matrix $\rho(\tau)$ which is diagonal in the energy eigenstates of the system. For example, if the system is initially in temperature equilibrium, $\rho = e^{-\beta H}/T_r(e^{-\beta H})$ where H is the Hamiltonian operator such that $\rho_{mn} = \delta_{mn} \exp(-\beta E_n)/\sum_n \exp(-\beta E_n)$. The influence functional becomes

$$\mathfrak{F}(Q, Q') = \int \delta(P_T - P'_T) \exp\{(i/\hbar)[S(P) - S(P') + \int_t^T (UQ - U'Q')\, dt]\} \sum_a \rho_{aa}\phi_a^*(P'_\tau)\phi_a(P_\tau)\, dP_\tau \cdots \mathfrak{D}P'(t) \tag{5.8}$$

Within the limit of small coupling, then we can simply extend the influence phase of Eq. (5.5) by summing over all initial states weighted by the initial density matrix ρ_{aa}. If this is done, we obtain the usual form of the influence phase, Eq. (5.4), with a response function $B(t-s)$ given by

$$B(t-s) = \sum_{a,b} (2\rho_{a'}|U_{ab}|^2/\hbar) \sin \omega_{ba}(t-s) \tag{5.9}$$

Again $\Phi(Q, Q')$ is the phase for a sum of oscillators, each of whose weights is $\rho_{aa}|U_{ib}|^2/\hbar^2$.[19] In Section IV it was shown that $B(t-s)$ was the classical response of the linear system to an impulse of force applied to the coordinate $U(P)$. However, the expression above is in a form which is familiar to us only when we think of the interaction system, P, as consisting of a sum of oscillators. In this connection $B(t-s)$ is the total classical response of the oscillators describing the system P to an impulse of force applied to the "coordinate" $U(P)$.[20] To obtain a more direct interpretation of Eq. (5.9) we now calculate the linear classical response $B(t-s)$ of the interaction system to an applied impulse of force in terms of its unperturbed quantum characteristics. In so doing we will show that Eq. (5.9) is indeed this expression. Therefore, we will again have the result that the influence functional for a general, linear interaction system is formed simply from a knowledge of its classical characteristics just as in the case of systems of perfect oscillators. By the "classical characteristics" of the quantum mechanical system we mean the expected value of $U(P)$ as a function of time after a potential $-f(t)U(P)$ is applied starting at $t = 0$. Thus

$$\langle U(P_t)\rangle = \int \psi^*(P_t)U(P_t)\psi(P_t)\, dP_t \tag{5.10}$$

where $\psi(P_t)$ represents the state of system P at t. Using the path integral representation for the development of a wave function with time, as outlined in Section II, this can be written

$$\langle U(P_t)\rangle = \int U_t\delta(P_t - P'_t) \exp\{(i/\hbar)[S_0(P) - S_0(P') + \int_0^t f(s)(U_s - U'_s)\, ds]\}\psi^*(P'_0)\psi(P_0)\, dP_0 \cdots \mathfrak{D}P'(t)$$

Alternatively, this expression can evidently be written

$$\langle U(P_t)\rangle = -i\hbar \frac{\partial}{\partial f} \mathfrak{F}(f_t\,; f'_t)\,|_{f=f'}$$

[19] It is interesting to notice that the relative populations of any two levels may be described by an effective temperature $T_e = 1/k\beta_e$. For instance if the probabilities of occupation of states a and b are ρ_{aa} and ρ_{bb} respectively, we use the definition $\rho_{aa}/\rho_{bb} = \exp\beta_e(E_b - E_a)$. If $\rho_{aa} = 0$ this is described by setting $T_e = 0+$, meaning to approach zero from the positive side. Similarly, if the two states were inverted $\rho_{bb} = 0$ and $T_e = 0-$. This device has been used widely in the description of such situations.

[20] $U(P)$ may be regarded as a coordinate which is a function of other coordinates P in terms of which we choose to describe the interaction system.

The initial definite state $\psi_0^*(P')\psi_0(P)$ will be replaced by an average state described by the following density matrix diagonal in the energy representation, $\sum \rho_{aa}\phi_a^*(P'_0)\phi_a(P_0)$ so that

$$\psi^*(P'_t)\psi(P_0) = \sum_a \rho_{aa}\phi_a^*(P'_0)\phi_a(P_0)$$

Assuming $f(s)$ to be small in magnitude, Eq. (5.10) can be written to first order

$$\begin{aligned}\langle U_t\rangle = \sum_a \rho_{aa} \int U_t \delta(P_t - P'_t) \exp\{(i/\hbar)[S_0(P) - S_0(P')]\} \\ \cdot\{1 + (i/\hbar)\int_0^t f_s U_s\, ds - (i/\hbar)\int_0^t f_s U'_s\, ds\} \\ \cdot\phi_a^*(P'_0)\phi_a(P_0)\, dP_0 \cdots \mathfrak{D}P'(t) \\ = \sum_a \rho_{aa}\{U_{aa} + \sum_b (i/\hbar)|U_{ab}|^2 \int_0^t f_s \exp[-i\omega_{ba}(t-s)]\, ds \\ - \sum_b (i/\hbar)|U_{ab}|^2 \int_0^t f_s \exp[i\omega_{ba}(t-s)]\, ds\}\end{aligned} \tag{5.11}$$

Again assuming $U_{aa} = 0$, Eq. (5.11) becomes

$$\langle U_t\rangle = \sum_{a,b} (2\rho_{aa}|U_{ab}|^2/\hbar) \int_0^t f_s \sin\omega_{ba}(t-s)\, ds \tag{5.12}$$

For $f(s) = \delta(s)$, then, the classical response function is[21]

$$B(t) = \sum_{a,b} (2\rho_{aa}|U_{ab}|^2/\hbar) \sin\omega_{ba}t \tag{5.13}$$

which is identical with the response function found in the influence functional.

Since we have found expressions identifiable as classical response functions and impedances, it remains to show that associated with the dissipative part of the impedance is a noise potential. The impedance is simply obtained from[22]

$$\begin{aligned}(i\nu Z_\nu)^{-1} &= \int_0^\infty B(t)e^{-i\nu t}\, dt \\ &= \sum_{a,b} (-2\rho_{aa}\omega_{ba}|U_{ba}|^2/\hbar)[(\nu - i\epsilon)^2 - \omega_{ba}^2]^{-1}\end{aligned} \tag{5.14}$$

To obtain the power spectrum it is only necessary to sum the influence phase of Eq. (5.6) over all initial states weighted by ρ_a . Thus we find

$$\phi(\nu) = \sum_{a,b} \pi\rho_{aa}|U_{ab}|^2 \delta(\nu + \omega_{ba}) \tag{5.15}$$

and we now wish to relate this to the real part of $1/Z_\nu$. From Eq. (5.14) it is found that

$$\text{Re}\,(Z_\nu)^{-1} = \sum_{a,b} (\pi\nu|U_{ba}|^2/\hbar)(\rho_{bb} - \rho_{aa})\delta(\nu + \omega_{ba}) \tag{5.16}$$

[21] It should be noted that implicit in the use of first order perturbation theory to obtain Eq. (5.13) is the fact that for this relation for the response function to hold as a steady-state description of the linear system, the initial distribution must not be significantly disturbed by the application of the driving force.

[22] Expressions of this kind have been used by several authors to compute quantities such as the conductivity of materials. See, for instance, ref. *14*.

Rewriting $\phi(\nu)$

$$\phi(\nu) = \sum_{a,b} \pi \mid U_{ab} \mid^2 (\rho_{bb} - \rho_{aa})[(\rho_{bb}/\rho_{aa}) - 1]^{-1} \delta(\nu + \omega_{ba}) \qquad (5.17)$$

If the average initial state of the interaction system is one of temperature equilibrium, then $\rho_{aa} = e^{-\beta E_a} / \sum_n e^{-\beta E_n}$ and $\rho_{bb}/\rho_{aa} = e^{\beta(E_a - E_b)} \equiv e^{\beta\hbar\omega_{ab}}$. Taking advantage of the characteristics of $\delta(\nu + \omega_{ba})$ so that ν can replace ω_{ab}, from Eqs. (5.16) and (5.17)[23]

$$\begin{aligned} \phi(\nu) &= \sum_{a,b} (\pi/\hbar) \mid U_{ba} \mid^2 \nu(\rho_{bb} - \rho_{aa}) \delta(\nu + \omega_{ba}) \hbar[\nu(e^{\beta\hbar\nu} - 1)]^{-1} \\ &= \hbar \operatorname{Re} (Z_\nu)^{-1} [\nu(e^{\beta\hbar\nu} - 1)]^{-1} \end{aligned} \qquad (5.18)$$

Therefore, again we find that the power spectrum is related to the dissipative parts of the impedance when the initial state is one of temperature equilibrium.[24] The power spectrum given in the form of Eq. (5.17) again illustrates the origin of thermal noise and identifies it as being just another aspect of spontaneous emission. Pound has also discussed this (*15*). The only contribution to the noise power spectrum is through the possible downward transitions of each possible state, a, weighted by the statistical factor, ρ_{aa}.

B. Beam of Particles Interacting with a Cavity

In the cases just considered the interaction system provided a steady-state environment for the test system. In contrast, let us now examine the situation where the interaction system is made up of a large number of independent particles coupled to the test system at different times. As an example consider a beam of noninteracting, identical particles which interact weakly with a resonant cavity as might occur, for instance, in a gas maser. We assume that the beam is not necessarily in temperature equilibrium but that the initial state of the particles entering the cavity would be properly represented by a density matrix diagonal in the energy representation. Such a situation would occur if the beam were prepared by passing it through a beam separator whose function would be to eliminate certain particles from the beam depending on their energy levels. For the purposes of simplifying the analysis we assume the molecules to be two-level quantum systems and that before they enter the cavity all of them are in

[23] Notice that if P were a two-level system initially in the lower state, then $\rho_{bb} = 0$ and $T_e = 0+$. In this case $\phi(\nu) = 0$ which agrees with the required result for $\nu > 0$ (Eq. (5.6) for $\omega_{ba} > 0$). If initially P were in the upper state then $\rho_{aa} = 0$ and $T_e = 0-$ yielding $\phi(\nu) = -(\hbar/\nu)\mathrm{Re}(1/Z_\nu)$ agreeing with Eq. (5.6) for $\omega_{ba} < 0$. This is the power spectrum of the so-called spontaneous emission noise from an inverted two-level system.

[24] It may be disturbing that $\mathrm{Re}(1/Z_\nu)$ contains singular forms such as $\delta(\nu + \omega_{ba})$. However, the infinite sums over the distribution of states of which it is a coefficient can be replaced by integrals over densities of states in most practical situations and as part of an integrand $\delta(\nu + \omega_{ba})$ is not unrealistic.

the lower state or all in the upper state. It is easy to extend these results to the case where the beam is mixed with a certain fraction in the upper state and a certain fraction in the lower state initially. Since the beam is assumed to be composed of noninteracting particles, we can consider the total beam as composed of two independent beams appropriate to the two possible initial states of the constituent particles. The influence phase for the complete beam is simply the sum of the influence phases for the two beams. In addition, we assume that the beam is characterized by a spatial density such that the number of particles passing a given point along the beam in a time dt is Ndt and if t_0 is the time a molecule passes a reference point in the cavity, $\gamma(t - t_0)$ describes the coupling between the molecule and cavity. Thus, the beam is a univelocity beam. Again, in a real case where the beam is characterized by a distribution of velocities, the total beam may be split up into many univelocity beams. The total influence phase is simply a sum of those for each component beam.

Let us call the coordinates of the cavity Q, representing the test system and the coordinates of a particle in the beam P. The interaction between beam and particles is given by $L_I(Q, P) = \gamma(t - t_0)QP$. Under these circumstances the influence functional for the effect of the beam on the cavity can be written down immediately:

$$i\Phi_B(Q, Q') = -N \mid P_{ab}/\hbar \mid^2 \int_{-\infty}^{\infty} \int_{-\infty}^{t} ds\, dt \left[\int_{-\infty}^{\infty} \gamma(t - t_0)\gamma(s - t_0)\, dt_0 \right] \times (Q_t - Q'_t)(Q_s e^{-i\omega_{ba}(t-s)} - Q'_s e^{i\omega_{ba}(t-s)}) \tag{5.19}$$

The integral involving the coupling parameters can be used to define a new function Γ,

$$\int_{-\infty}^{\infty} \gamma(t - t_0)\gamma(s - t_0)\, dt_0 = \int_{-\infty}^{\infty} \gamma(\xi)\gamma(\xi - t + s)\, d\xi \equiv \Gamma(t - s) \tag{5.20}$$

Therefore,

$$i\Phi(Q, Q') = -N \mid P_{ab}/\hbar \mid^2 \int_{-\infty}^{\infty} \int_{-\infty}^{t} \Gamma(t - s) \times (Q_t - Q'_t)(Q_s e^{-i\omega_{ba}(t-s)} - Q'_s e^{i\omega_{ba}(t-s)})\, ds\, dt \tag{5.21}$$

From this we can identify the response function of the beam as

$$B(t - s) = (2/\hbar)N \mid P_{ab} \mid^2 \Gamma(t - s) \sin \omega_{ba}(t - s) \tag{5.22}$$

In transform notation the influence functional for the effect of the beam has the same form as has been previously derived with

$$(i\nu Z_\nu)^{-1} = (i/\hbar)N \mid P_{ab} \mid^2 (\Gamma_{\nu - \omega_{ba}} - \Gamma_{\nu + \omega_{ba}}) \tag{5.23}$$

where

$$\Gamma_{\nu + \omega_{ba}} = \int_{-\infty}^{\infty} 1(r)\Gamma(r) e^{-i(\nu + \omega_{ba})r}\, dr \tag{5.24}$$

and with a power spectrum

$$\phi(\nu) = \tfrac{1}{2}N \mid P_{ab} \mid^2[\Gamma_{\nu+\omega_{ba}} + \Gamma_{-\nu-\omega_{ba}}] \tag{5.25}$$

Previously, we have found that for a linear system initially in the ground state the power spectrum is zero over the range of positive ν thus indicating a zero noise potential due to the linear system. However, ϕ_ν is not necessarily zero for $\nu > 0$ in Eq. (5.25) in the case that the beam is initially in its lowest state ($\omega_{ba} > 0$). This is because no restrictions were placed on the time variation of $\gamma(t - t_0)$. However, in practical situations the coupling between a cavity and a particle in a beam passing through the cavity varies adiabatically so that for all practical purposes $\phi(\nu)$ is really zero for $\nu > 0$. To point this out more clearly, let us assume that $\gamma(t - t_0) = 1(t - t_0)1(t_0 + \tau_0 - t)$, that is, the coupling is turned on at t_0 and off at $t_0 + \tau_0$, the time of transit in the cavity being τ_0. We find by evaluating Eq. (5.20) that

$$\Gamma(t - s) = (s - t + \tau_0)1(s - t + \tau_0)$$

and from this we can find

$$\begin{aligned}\Gamma_{\nu+\omega_{ba}} &= \int_{-\infty}^{\infty} (\tau_0 - r)1(\tau_0 - r)1(r) \exp[-i(\nu + \omega_{ba})r]\, dr \\ &= \{1 - i\tau_0(\nu + \omega_{ba}) - \exp[-i(\nu + \omega_{ba})\tau_0]\}(\nu + \omega_{ba})^{-2}\end{aligned}$$

Therefore, from Eq. (5.25),

$$\phi(\nu) = (\tfrac{1}{2})N \mid P_{ab} \mid^2 \tau_0^2(\sin^2\theta/\theta^2) \tag{5.26}$$

where

$$\theta = \tfrac{1}{2}(\nu + \omega_{ba})\tau_0$$

To find the effect of suddenly turning the coupling on and off we find the ratio of the total noise power to the noise power for $\nu > 0$. This is given by

$$\frac{\int_{-\infty}^{\infty} \phi(\nu)\, d\nu}{\int_0^{\infty} \phi(\nu)\, d\nu} > \frac{\int_{-\infty}^{\infty} (\sin^2\theta/\theta^2)\, d\theta}{\int_{\omega_{ba}\tau_0}^{\infty} d\theta/\theta^2} = \pi\omega_{ba}\tau_0$$

For ammonia molecules at a temperature of $T = 290°\text{K}$, the average velocity $v \approx 6 \times 10^4$ cm/sec. For a microwave cavity of 10 cm length $\tau_0 = 2 \times 10^4$ sec. For the 3–3 line of ammonia $\omega_{ba} \approx 1.5 \times 10^{11}$ rad/sec. For our case then $(\pi\omega_{ba}\tau_0) \approx 10^8$. From this, we can conclude that even in this unfavorable case of coupling time variation, the $\phi(\nu)$ is negligible for $\nu > 0$.

Examination of $\text{Re}(1/Z_\nu)$, derived from Eq. (5.23), reveals terms of the same form as those just discussed,

$$\begin{aligned}\text{Re}(1/Z_\nu) &= \tfrac{1}{2}[(1/Z_\nu) + (1/Z_\nu^*)] \\ &= (\nu N/2\hbar) \mid P_{ab} \mid^2[\Gamma_{\nu+\omega_{ba}} - \Gamma_{\nu-\omega_{ba}} - \Gamma_{-\nu+\omega_{ba}} + \Gamma_{-\nu-\omega_{ba}}]\end{aligned} \tag{5.27}$$

By the same type of argument as above, the terms in $\Gamma_{\nu-\omega_{ba}}$ can be neglected when $\nu > 0$, $\omega_{ba} < 0$ and conversely, when $\nu > 0$, $\omega_{ba} > 0$ the terms in $\Gamma_{\nu+\omega_{ba}}$ can be neglected. Since this is the case, we can write

$$\phi(\nu) \approx \hbar \operatorname{Re}(1/Z_\nu)[\nu(e^{-\beta_e \hbar\omega_{ba}} - 1)]^{-1} \tag{5.28}$$

where $\beta_e = 1/kT_e$, describes the relative initial populations of the two states. Therefore, we conclude that in most practical cases the power spectrum can be written in the form of Eq. (5.28). In cases where the transients cannot be neglected for very low ω_{ba} and very short transit times, however, $\phi(\nu)$ is not so simply related to $\operatorname{Re}(1/Z_\nu)$ and must be written in the form given by Eq. (5.25).

It is to be noted that $\phi(\nu)$ of Eq. (5.28) is not precisely of the form for the Nyquist relation because of the appearance of $e^{\beta\hbar\omega_{ba}}$ in the denominator rather than $e^{\beta\hbar\nu}$. This is a consequence of the finite coupling time between each part of the beam and the cavity which results in a nonequilibrium condition. If the coupling times were infinite, the expression for $\operatorname{Re}(1/Z_\nu)$ would contain forms such as $\delta(\nu + \omega_{ba})$, a situation discussed earlier, so that the Nyquist form then results. However, when the coupling time is long as in masers,

$$\Gamma_{\nu+\omega_{ba}} + \Gamma_{-\nu-\omega_{ba}} \approx 4\pi\delta(\nu + \omega_{ba})$$

so that the true Nyquist relation may be used with negligible error.

VI. SOURCES OF NOISE IN MASERS

Having developed the theory of linear systems in detail we are now in a position to discuss the sources of noise in linear quantum-mechanical devices such as maser amplifiers. The subject of maser noise has been explored by many authors (*15–19*). The details of the treatment given the subject differ, but the principles are essentially the same. The amplifiers are considered as operating at signal levels high enough (classical) that a signal entering a maser may be considered as a group of photons whose number is large enough that the amplification process increases the signal in a continuous fashion. The sources of noise were found to be those derived from the thermal noise arising from the sources of dissipation, and those derived from spontaneous emission from the "active" quantum material. They go further and define an effective temperature of the active quantum system so that the noise it produces is related to the negative resistance of the active materials. Our analysis of linear systems has also shown that these same sources of noise exist. However, if the signal level entering the maser is very small such that its strength can be characterized by a few quanta per second, a serious question arises as to the nature of the signal out of the maser (here assumed to be at a classical level). An additional fluctuation of the output signal or "quantum" noise might be expected due solely to what might be termed

a "shot noise" effect created by each individual photon entering the cavity. It has been shown that no such signal exists by other authors (*20*, *21*). We wish to show how the same result follows from the theory developed here in an extremely transparent way. We will show that the only fluctuations in output signal which are to be expected are those noise sources computed classically. No additional "shot noise" does, in fact, appear.

Let us suppose that we have a beam-type maser amplifier in which all participating systems used meet the requirement of linearity. There may be one or more beams interacting with various electromagnetic resonators which can be coupled together in any way desired. The output of the maser is connected to a detector of some sort which perhaps consists of a resistor in which the current is to be measured. To the input of the maser system we now apply an incoming classical signal of large magnitude and of frequency ω through an attenuator whose value of attenuation may be varied at will as shown in Fig. 12. In practice such a situation could arise if the classical wave originates from a distant antenna with a very large magnitude of output, so large that all quantum effects in the wave are effectively obscured. The long distance would then play the role of the attenuator.

Now, if the classical wave were attenuated by a large amount so that only a few photons/sec were entering the maser, the only uncertainty in the signal in the output of the maser caused by the maser itself arises from those sources of noise which can be arrived at by a classical calculation of the characteristics of the maser. There is no extra quantum fluctuation introduced by the maser into the output signal due to the small number of quanta entering the maser.

It is true that the amplitude of the signal output from the maser might itself be so small that it is still on a quantum level. In this case the detector output would be uncertain due to the inherently small magnitude of the signal from the maser. However, if this is the case, we may put as many amplifiers in series as necessary to bring the output signal back to a classical level. When this is done the signal applied to the detector consists of the original signal modified by the transfer characteristics of the maser system (and attenuator) and noise signals which arise from all the possible sources computed classically. The proof of this assertion is not difficult. We divide up the total system into a test system, here the detector, and an interaction system which consists of the maser, attenuator, and classical signal, $C(t)$. Then, to find the effect of the interaction system on the

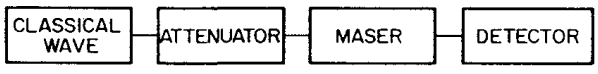


FIG. 12. System in which a classical wave is attenuated to a very low level (a few photons/sec), then amplified by a maser and detected.

detector we need only to look at the influence functional. However, we already know that it can be written as follows:

$$\mathfrak{F}(Q, Q') = \exp\left\{\left(\frac{i}{2\pi\hbar}\right)\int_0^\infty \left[\frac{Q'_\nu(Q_{-\nu} - Q'_{-\nu})}{(i\nu Z_\nu)} + \frac{Q_{-\nu}(Q_\nu - Q'_\nu)}{(-i\nu Z_{-\nu})}\right] d\nu \right.$$
$$\left. + \left(\frac{i}{2\pi\hbar}\right)\int_{-\infty}^\infty \left(\frac{C_\nu}{i\nu z_\nu}\right)(Q_{-\nu} - Q'_{-\nu})\, d\nu - \left(\frac{1}{\pi\hbar^2}\right)\int_0^\infty |Q_\nu - Q'_\nu|^2 \sum_i \phi_i(\nu)\, d\nu\right\}$$

where $\phi_i(\nu)$ represents the power spectrum of the noise from the ith source, z_ν is the classical transfer characteristic of the complete interaction system, and Z_ν is the impedance of the maser system as seen by the detector. All the terms in the influence functional are familiar in view of the derivations which have been presented previously. The first term in the exponent of the influence functional is recognized as describing a linear system at zero temperature (Section IV), the linear system in this case being the maser. This term describes the spontaneous emission of the detector back into the maser (see Appendix I). Furthermore, it can be deduced that this spontaneous emission can be thought of as resulting from a noise generator created by the detector (test system) acting on the interaction system in the usual classical way, i.e., whose power spectrum was related to the dissipative part of the detector impedance and to the temperature by the Nyquist relation

$$\phi(\nu) = \hbar\nu R_\nu[\exp(\hbar\nu/kT) - 1]^{-1}$$

where R_ν is the detector resistance and T its temperature. The second term in the above composite influence functional is easily interpreted and is simply the effect of a classical voltage, related to the input voltage by the classical transfer characteristic of the maser, acting on the detector (see Section IV.B). The last term represents the effect on the detector of random noise voltages associated with the various classical noise sources in the maser (Sections IV.C and V). Both positive and negative resistances are such noise sources. In either case the power spectrum of the noise from a particular resistance is computed from the same relation as given above.

If R_ν is negative the effective temperature of R_ν will also be negative always giving a positive power spectrum. Therefore, if we were to compute the current in the detector due to the interaction system (maser) using the influence functional we would find components of current due to: (1) the noise voltage generated by the detector itself, the power spectrum of which is related to the resistance of the detector by the generalized Nyquist relation given above; (2) a classical voltage related to the input voltage $C(t)$ by the classical transfer characteristic of the maser; and (3) random noise voltages associated with the various classical noise sources (resistances) in the maser. Therefore, the maser

simply acts as a classical amplifier with sources of noise which can be predicted from considerations of its classical characteristics.

APPENDIX I

In Section II.E the problem of making perturbation calculations using influence functionals was outlined. Here we will calculate in detail the probability, to second order in the coupling potentials, that a test system which is in a definite state at $t = \tau$ and finds itself in another state at $t = T$ when interacting with a zero temperature interaction system. Let us call the initial and final states $\phi_n(Q_\tau)$ and $\phi_m(Q_T)$ respectively and, for simplicity, assume that these are eigenstates. The formal expression for the transition probability is given by Eq. (4.24) and for this case,

$$\mathfrak{F}(Q, Q') = \exp\{-(2\hbar)^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{t}\gamma_t\gamma_s(Q_t - Q'_t) \times [Q_sF^*(t-s) - Q'_sF(t-s)]\,ds\,dt\} \tag{I.1}$$

Then to second order

$$P_{nm} \approx \int\phi_m^*(Q_T)\phi_m(Q'_T)\exp\{(i/\hbar)[S_0(Q) - S_0(Q')]\} \times \{1 - (2\hbar)^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{t}\gamma_t\gamma_s(Q_t - Q'_t)[Q_sF^*(t-s) - Q'_sF(t-s)]\,ds\,dt\} \times \phi_n^*(Q'_\tau)\phi_n(Q_\tau)\,dQ_\tau\cdots\mathfrak{D}Q'(t) \tag{I.2}$$

Making use of the fact that

$$\int\exp[(i/\hbar)S_0(Q)]\mathfrak{D}Q(t) = K(Q_T, Q_\tau) = \int K(Q_T, Q_t)K(Q_t, Q_\tau)\,dQ_t \quad (\text{where } T < t < \tau)$$

and writing (I.2) as a sum of integrals

$$\begin{aligned} P_{nm} \approx{}& \int\phi_m^*(Q_T)\phi_m(Q'_T)\exp\{(i/\hbar)[S_0(Q) - S_0(Q')\} \cdot\phi_n^*(Q'_\tau)\phi_n(Q_\tau)\,dQ_\tau\cdots\mathfrak{D}Q'(t) \\ &- (2\hbar)^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{t}ds\,dt\gamma_s\gamma_t\int\phi_m^*(Q_T)\phi_m(Q'_T) \cdot\exp\{(i/\hbar)[S_0(Q) - S_0(Q')]\} \\ &\times (Q_t - Q'_t)[Q_sF^*(t-s) - Q'_sF(t-s)] \cdot\phi_n^*(Q'_\tau)\phi_n(Q_\tau)\,dQ_\tau\cdots\mathfrak{D}Q'(t) \end{aligned} \tag{I.3}$$

Replacing $K(Q_T, Q_t)$ by $\sum_k \phi_k(Q_T)\phi_k^*(Q_t)\exp[-iE_k(T-t)/\hbar]$ and taking matrix elements we have

$$P_{nm} = \delta_{nm}[1 - \textstyle\sum_k(2\hbar)^{-1}\,|\,Q_{nk}\,|^2 f(\nu_{nk})] + (2\hbar)^{-1}\,|\,Q_{nm}\,|^2 f(\nu_{nm}) \tag{I.4}$$

where

$$f(\nu_{nm}) = \int_{-\infty}^{\infty}\int_{-\infty}^{t}\gamma_t\gamma_s[F^*(t-s)e^{i\nu_{nm}(t-s)} + F(t-s)e^{-i\nu_{nm}(t-s)}]\,ds\,dt \quad \text{(I.5)}$$

$f(\nu)$ can be simplified by restating the integral over t and s in terms of frequency. To do this we replace the upper limit t by $+\infty$ and multiply the integrand by a step function $1(t-s)$. Then utilizing the convolution theorem in the form

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} M(t)N(s)R(t-s)\,ds\,dt = (2\pi)^{-1}\int M_{-\nu}N_\nu R_\nu\,d\nu$$

where

$$M(\nu) = \int_{-\infty}^{\infty} M(t)e^{-i\nu t}\,dt$$

$f(\nu)$ becomes

$$f(\nu) = (2\pi)^{-1}\int_{-\infty}^{\infty} |\gamma_{\nu-\omega}|^2(F_\nu + F_\nu^*)\,d\nu$$

where

$$F_\nu = \int_{-\infty}^{\infty} 1(t)F(t)e^{-i\nu t}\,dt$$

From Eqs. (4.5a) and (4.5b) we have that

$$F(t) = -(2/\pi)\int_0^{\infty} \operatorname{Im}(1/i\eta Z_\eta)e^{i\eta t}\,d\eta$$

Therefore,

$$\begin{aligned} F_\nu + F_\nu^* &= -(2/\pi)\int_0^{\infty} d\eta \operatorname{Im}(1/i\eta Z_\eta)\int_0^{\infty} dt(e^{i(\eta-\nu)t} + e^{-i(\eta-\nu)t}) \\ &= -4\operatorname{Im}(1/i\nu Z_\nu) = (4/\nu)\operatorname{Re}(1/Z_\nu) \end{aligned}$$

and

$$f(\nu) = (2/\pi)\int_{-\infty}^{\infty} |\gamma_{\eta-\nu}|^2\eta^{-1}\operatorname{Re}(1/Z_\eta)\,d\eta \quad \text{(I.6)}$$

Thus, to second order

$$\begin{aligned} P_{nm} = \delta_{nm}\{1 - \textstyle\sum_k (\pi\hbar)^{-1}|Q_{nk}|^2\int_0^{\infty}\nu^{-1}|\gamma_{\nu-\nu_{nk}}|^2\operatorname{Re}(Z_\nu)^{-1}\,d\nu\} \\ + (\pi\hbar)^{-1}|Q_{nm}|^2\int_0^{\infty}\nu^{-1}|\gamma_{\nu-\nu_{nm}}|^2\operatorname{Re}(Z_\nu)^{-1}\,d\nu \end{aligned} \quad \text{(I.7)}$$

For the special case that the coupling γ is

$$\gamma = 0 \quad \text{for} \quad t > T/2 \quad \text{and} \quad t < -T/2$$

$$\gamma = 1 \quad \text{for} \quad -T/2 < t < T/2$$

$$|\gamma_{\nu-\nu_{nk}}|^2 = \left|\int_{-T/2}^{T/2} e^{-i(\nu-\nu_{nk})t}dt\right|^2 = 4(\nu-\nu_{nk})^{-2}\sin^2(\nu-\nu_{nk})T/2$$

$$\to 2\pi T\delta(\nu-\nu_{nk}) \quad \text{for large } T$$

Then,

$$P_{nm} = \delta_{nm}\{1 - \textstyle\sum_{k,\nu_{nk}>0} 2T(\hbar\nu_{nk})^{-1} \mid Q_{nk} \mid^2 \mathrm{Re}(Z_{\nu_{nk}})^{-1}\} + 2T(\hbar\nu_{nm})^{-1} \mid Q_{nm} \mid^2 \mathrm{Re}(Z_{\nu_{nm}})^{-1} \quad \text{for } \nu_{nm} > 0 \qquad (1.8)$$
$$= 0 \text{ for } \nu_{nm} < 0$$

Thus, it is seen that the P_{nm} now is proportional to the matrix element and to the dissipative part of the impedance $Z_{\nu_{nm}}$. Appropriately enough, no transition is possible to energy states such that $\nu_{nm} < 0$ since the interaction system, being at zero temperature initially, can give up no energy to the test system.

APPENDIX II

A. Influence Phase for Effect of Free Space on an Atom

As an illustration, the influence phase for the effect of free space on an atom will be calculated. This problem is more complicated than the idealized systems considered in deriving the formalism since the interaction here is of the form $\mathbf{Q}\cdot\mathbf{X}$, $\mathbf{Q}$ and $\mathbf{X}$ being vectors rather than QX where Q and X are scalars. This difficulty could be overcome by writing the influence phase in tensor notation or by recasting the problem so that the interaction is of the form QX. The latter will be done to adhere more closely to the point of view of the derivations. Since a linear system is being dealt with, it is only necessary to determine a suitable coordinate for the atom and find the impedance function Z_r for the effect of free space. It is assumed that the atom is made up of a system of particles of mass m_n, charge e_n, and position $\mathbf{r}_A + x_n$ where $\mathbf{r}_A$ is the position of the center of charge of the atom. If the transverse part of the radiation field in the box is expanded into a series of plane waves each representing independent harmonic oscillations (*22*), then the nonrelativistic Lagrangian for the complete system consisting of the atom and the field in the box can be written (*2*)

$$L(\dot{X}, X, \dot{q}_{\mathbf{k}}, q_{\mathbf{k}}, t) = L_A + c^{-1}\textstyle\sum_n e_n\dot{\mathbf{X}}_n\cdot\mathbf{A}^{\mathrm{tr}}(\mathbf{r}_A + \mathbf{r}_n) + \tfrac{1}{2}\sum_{\mathbf{k}}\sum_{r=1}^{4}[(\dot{q}_{\mathbf{k}}^{(r)}) - c^2k^2(q_{\mathbf{k}}^{(r)})^2] \qquad (\text{II}.1)$$

where L_A is the Lagrangian of the atom unperturbed by outside forces and

$$\mathbf{A}^{\mathrm{tr}}(\mathbf{X}) = (8\pi c^2)^{1/2}\textstyle\sum_{\mathbf{k}}[\mathbf{e}_\lambda(q_{\mathbf{k}}^{(1)}\cos(\mathbf{k}\cdot\mathbf{X}) + q_{\mathbf{k}}^{(3)}\sin(\mathbf{k}\cdot\mathbf{X})) + \mathbf{e}_l(q_{\mathbf{k}}^{(2)}\cos(\mathbf{k}\cdot\mathbf{X}) + q_{\mathbf{k}}^{(4)}\sin(\mathbf{k}\cdot\mathbf{X}))] \qquad (\text{II}.2)$$

Here $\mathbf{e}_\lambda$ and $\mathbf{e}_l$ are two mutually orthogonal polarization vectors, each orthogonal to the propagation vector $\mathbf{k}$. Now we assume that the radiation field of the box is constant over the particle, i.e., that $\mathbf{A}(\mathbf{r}_A + \mathbf{r}_n) \approx \mathbf{A}(\mathbf{r}_A)$, the dipole approximation.[25] This permits one to replace $\sum_n e_n\dot{\mathbf{X}}_n$ by $\mathbf{j}$, the current operator for

[25] This is equivalent to taking $\mathbf{A}(\mathbf{k}_A + \mathbf{x}_n)$, expanding it in a series of $\mathbf{k}\cdot\mathbf{x}_n$ since this is assumed small, and keeping only those terms which keep the interaction term of the Lagrangian linear. Since the interaction is of the form $e_n\mathbf{x}_n\cdot\mathbf{A}(\mathbf{k}_A + \mathbf{x}_n)$ for the nth particle, then $\mathbf{A}$ can only contain constant terms.

the atom. In addition, even though $\mathbf{e}_\lambda \cdot \mathbf{e}_l = 0$, which fixes their relative orientations, their absolute directions in a plane perpendicular to $\mathbf{k}$ are still arbitrary. Choosing $\mathbf{e}_l$ so that

$$\mathbf{e}_l \cdot \mathbf{j} = 0 \tag{II.3}$$

We assume the box to be very large so that $\sum_{\mathbf{k}} \to (16\pi^3)^{-1} \int d^3\mathbf{k}$, because the mode corresponding to $\mathbf{k}$ and $-\mathbf{k}$ is the same. Combining (II.1), (II.2), and (II.3) the following total Lagrangian is obtained.

$$\begin{aligned} L_T = L_A + \sum_{r=1,3} \int \frac{d^3\mathbf{k}}{16\pi^3} \left[\frac{1}{2} (\dot{q}_{\mathbf{k}}^{(r)})^2 - \frac{1}{2} k^2 c^2 (q_k^{(r)})^2 \right] \\ + (8\pi)^{1/2} \int \frac{d^3\mathbf{k}}{16\pi^3} (\mathbf{j} \cdot \mathbf{e}_\lambda) q_{\mathbf{k}}^{(1)} \cos(\mathbf{k} \cdot \mathbf{r}_A) \\ + (8\pi)^{1/2} \int \frac{d^3\mathbf{k}}{16\pi^3} (\mathbf{j} \cdot \mathbf{e}_\lambda) q_{\mathbf{k}}^{(3)} \sin(\mathbf{k} \cdot \mathbf{r}_A) \end{aligned} \tag{II.4}$$

Thus, the number of each of the two sets of oscillators ($q_{\mathbf{k}}^{(1)}$ and $q_{\mathbf{k}}^{(3)}$) in a volume of k space $d^3\mathbf{k}$ is $d^3\mathbf{k}/16\pi^3$. The coupling strength of the $q_{\mathbf{k}}^{(1)}$ and $q_{\mathbf{k}}^{(3)}$ oscillators with the atom is $(8\pi)^{1/2}(\mathbf{j} \cdot \mathbf{e}_\lambda) q_{\mathbf{k}}^{(1)} \cos(\mathbf{k} \cdot \mathbf{r}_A)$ and $(8\pi)^{1/2}(\mathbf{j} \cdot \mathbf{e}_\lambda) q_{\mathbf{k}}^{(3)} \cos(\mathbf{k} \cdot \mathbf{r}_A)$ respectively. If $\mathbf{j}$ is oriented along the $\theta = 0$ axis in polar coordinate representation, $\mathbf{j} \cdot \mathbf{e}_\lambda = j \sin\theta$. Then choosing j as the atom coordinate, the impedance for the two oscillators of frequency kc can be found from the rule found in Section IV to be

$$\begin{aligned} [i\nu Z_\nu(\mathbf{k})]^{-1} &= (1/j_\nu)[(8\pi)^{1/2} \sin\theta \cos(\mathbf{k} \cdot \mathbf{r}_A) q_{\mathbf{k}}^{(1)} + (8\pi)^{1/2} \sin\theta \sin(\mathbf{k} \cdot \mathbf{r}_A) q_{\mathbf{k}}^{(3)}]_\nu \\ &= -\frac{8\pi \sin^2\theta \cos^2(\mathbf{k} \cdot \mathbf{r}_A)}{(\nu - i\epsilon)^2 - k^2c^2} - \frac{8\pi \sin^2\theta \sin^2(\mathbf{k} \cdot \mathbf{r}_A)}{(\nu - i\epsilon)^2 - k^2c^2} \\ &= -8\pi \sin^2\theta [(\nu - i\epsilon)^2 - k^2c^2]^{-1} \big|_{\epsilon \to 0} \end{aligned} \tag{II.5}$$

The total effect of all the oscillators is

$$\begin{aligned} (Z_\nu)^{-1} &= (16\pi^3)^{-1} \int Z_\nu^{-1}(\mathbf{k})\, d^3\mathbf{k} = (16\pi^3)^{-1} \int k^2 \sin\theta\, d\theta\, d\phi\, dk (Z_\nu^{-1}) \\ &= -i4\nu(3\pi c^3)^{-1} \int_0^\infty \Omega^2\, d\Omega [(\nu - i\epsilon)^2 - \Omega^2]^{-1} \big|_{\epsilon \to 0} \\ &= (2\nu^2/3c^3) - i(4\nu/3\pi c^3) \int_0^\infty \Omega^2\, d\Omega (\nu^2 - \Omega^2)^{-1} \end{aligned} \tag{II.6}$$

where the substitution $\Omega = kc$ has been made. Thus, the effect of free space is characterized by $(i\nu Z_\nu)^{-1}$. The equivalent distribution of oscillators coupled to j_ν is

$$G(\Omega) = (4\Omega^2/3\pi c^3) \tag{II.7}$$

B. Spontaneous Emission Probability of an Atom in Free Space

To compute the transition probability for this atom, we use second order perturbation theory developed in Appendix I for a system initially in state $\phi_n(X_\tau)$ and finally in state $\phi_m(X_T)$ when acted on by an influence functional for a linear system at zero temperature. The expression is

$$\begin{aligned} P_{nm} &= 2T \mid j_{nm} \mid^2 (\hbar\nu_{nm})^{-1} \operatorname{Re}(Z_{\nu_{nm}})^{-1}, \quad \nu_{nm} > 0, \quad n \neq m \\ &= 0, \quad \nu_{nm} < 0 \end{aligned} \tag{II.8}$$

From Eq. (II.7) we find

$$\operatorname{Re}(Z_{\nu_{nm}})^{-1} = (\pi/2)G(\nu_{nm}) = 2\nu_{nm}^2/3c^3 \tag{II.9}$$

Using this in Eq. (II.8)

$$\begin{aligned} P_{nm} &= 4\nu_{nm} \mid j_{nm} \mid^2 T/3\hbar c^3 \\ &= 4\nu_{nm}^3 e^2 \mid X_{nm} \mid^2 T/3\hbar c^3 \end{aligned} \tag{II.10}$$

where to obtain the last, more familiar form, the substitution $j_{nm} = \nu_{nm}X_{mn}$ has been made. This is the first order spontaneous emission probability for an atom in free space.

Now we can form an expression for the intensity of radiation per unit time. The power radiated from the dipole is

$$\begin{aligned} \hbar\nu_{nm}P_{nm}/T &= 2\nu_{nm}^2 e^2 \mid X_{nm} \mid^2 \operatorname{Re}(Z\nu_{nm})^{-1} \\ &= 4e^2 \mid X_{nm} \mid^2 \nu_{nm}^4/3c^2 \end{aligned} \tag{II.11}$$

an expression which is almost the same as that for power radiated from a classical dipole. The expression becomes exactly the same if we apply the correspondence principle by replacing the matrix element of the time average of the coordinate of the oscillator by its corresponding classical quantity. Thus, if X is the coordinate of the corresponding classical oscillator (X_0 is its maximum value) then[26]

$$2 \mid X_{mn} \mid^2 \rightarrow \langle X^2 \rangle = \tfrac{1}{2}X_0^2$$

[26] That this correspondence is true can be seen easily as follows. Consider the dipole, a harmonic oscillator as above, to be in a high quantum state, ϕ_n. Classically the motion of the dipole can be described as $X = X_0 \sin \omega t$. We wish to relate the classical value of $\langle X^2 \rangle$ to its matrix element. In a quantum mechanical sense,

$$\langle X^2 \rangle = \int \phi_n^*(X) X^2 \phi_n(X)\, dX = \iint \phi_n^*(X) X \textstyle\sum_k \phi_k(X)\phi_k^*(X') X' \phi_n(X')\, dX\, dX'$$

Since matrix elements exist, in the case of a harmonic oscillator only for $k = n - 1$, $k = n + 1$, we have

$$\langle X^2 \rangle = \mid X_{n,n-1} \mid^2 + \mid X_{n,n+1} \mid^2$$

For very high quantum numbers these two terms become nearly equal since

$$\mid X_{n,n-1} \mid^2 = n\hbar/2m\omega, \quad \mid X_{n,n+1} \mid^2 = (n+1)\hbar/2m\omega$$

Thus, as $n \rightarrow \infty$ $\langle X^2 \rangle = 2 \mid X_{n,n-1} \mid^2$. But in the classical case, $\bar{X}^2 = \frac{1}{2}X_0^2$. Therefore, $\mid X_{n,n-1} \mid^2 \Rightarrow \frac{1}{4}X_0^2$.

If $|X_{mn}|^2$ is replaced by $X_0^2/4$ then Eq. (II.11) becomes the expression for the power radiated from a classical dipole. Our purpose, however, in doing this example, was to show for a specific problem that the effect of a distribution of oscillators interacting on a system is the same as the effect of loss on the system.[27] This has been done by relating the energy lost from the radiating dipole to the distribution. It is not surprising that a sea of oscillators should give this effect. If the dimensions of the box are allowed to be finite then energy emitted from the system under observation is reflected from the walls and eventually finds its way back to be absorbed again. This is equivalent to saying that the number of oscillators comprising the electromagnetic field in the box is infinite with a finite frequency spacing between the modes. Since the oscillators are independent there is no coupling between them and energy coupled into one of the oscillators from the test system must eventually return to it. If the dimensions of the box are allowed to get infinitely large, energy emitted from the test system never gets reflected and thus never returns. In oscillator language this means that the frequency spacing between oscillators has become infinitesimal, so close that a little of the energy absorbed by each one gradually leaks into nearby modes and eventually is completely gone.

APPENDIX III. SPONTANEOUS EMISSION OF AN ATOM IN A CAVITY

In this calculation as in the free space calculation the dipole approximation will be used in computing the spontaneous emission probability. The linear coordinates inside the cavity will be represented by the vector Q while the time varying coordinates of the single cavity mode being considered will be $X(t)$. The Lagrangian of the system may be written

$$L_T \equiv L(Q_n, Q_n, t) + c^{-1}\sum_n e_n \dot{\mathbf{Q}}_n \cdot \mathbf{A}(\mathbf{Q}_t + \mathbf{Q}_n, t) + L_{\text{cavity}} \quad \text{(III.1)}$$

where $\mathbf{Q}_p$ is the atom coordinate, $\mathbf{Q}_n + \mathbf{Q}_p$ is the particle coordinate in the atom, and $\mathbf{A}$ is the vector potential of the cavity field. The interaction term is the one of interest, since from it we find the terms that we wish to solve for classically. This term will be put into more convenient form. Let us write

$$\mathbf{A}(\mathbf{Q}, t) = \mathbf{a}(\mathbf{Q})X(t) \quad \text{(III.2)}$$

[27] If the power radiated from the oscillator is related to the classical expression $\frac{1}{2}I^2R$ then from Eq. (II.11) it can be seen that R is proportional to $\mathrm{Re}[1/Z(\nu)]$ which in turn is related to the distribution of oscillators. One might expect $\mathrm{Im}[1/Z(\nu)]$ to be replaced to the reactance seen by an oscillating dipole, a quantity which is known to be infinite classically. From Eq. (II.7),

$$\mathrm{Im}(1/Z_\nu) = 4\nu/3\pi c^3 \int_0^\infty \Omega^2(\nu^2 - \Omega^2)^{-1}\, d\Omega = 4\nu\Omega/3\pi c^3 \,|_{\Omega\to\infty}$$

The integral is linearly divergent. This factor is also related to the infinite self energy of a point charge which occurs both classically and in quantum electrodynamics. Here this divergence does not bother us since it never enters into the calculation.

where

$$\int \mathbf{a}(\mathbf{Q})\cdot\mathbf{a}(\mathbf{Q})\, d^3\mathbf{Q} = 4\pi c^2 \tag{III.3}$$

If $\mathbf{A}$ does not vary much over the atom, then $\mathbf{A}(\mathbf{Q}_p) \approx \mathbf{A}(\mathbf{Q}_p + \mathbf{Q}_n)$ and the interaction term is written

$$c^{-1}\sum_n e_n\dot{\mathbf{Q}}_n\cdot\mathbf{a}(\mathbf{Q}_p)X(t) \quad = \quad (j/c)\mid \mathbf{a}(\mathbf{Q}_p)\mid X(t) \tag{III.4}$$

where $j = (\sum e_n\dot{\mathbf{Q}}_n)\cdot\mathbf{a}(\mathbf{Q}_p)/\mid \mathbf{a}(\mathbf{Q}_p)\mid$, the component of the atom current in the direction of the cavity field.

Let us now determine the ratio

$$j_\nu[\mid \mathbf{a}(\mathbf{Q}_p)\mid X_\nu c^{-1}]^{-1} = i\nu Z_\nu \tag{III.5}$$

classically. The wave equation appropriate for this calculation (high Q)[28] is

$$\nabla^2\mathbf{A} - (1/c^2)\ddot{\mathbf{A}} + (\omega/c^2Q)\dot{\mathbf{A}} - (4\pi/c)\dot{\mathbf{P}} = 0 \tag{III.6}$$

The atom is located at $\mathbf{Q}_p$ and, since the dipole moment is induced, its direction on the average is the same as that of the field $\mathbf{A}$ in the cavity. We have then

$$\dot{\mathbf{P}} = j\delta(\mathbf{Q} - \mathbf{Q}_p)\mathbf{a}(\mathbf{Q}_p)/\mid \mathbf{a}(\mathbf{Q}_p)\mid \tag{III.7}$$

Substituting (III.7) and (III.2) into (III.6) we obtain

$$\begin{aligned}[-\omega^2X + (\omega/Q)\dot{X} - \ddot{X}]c^{-2}\mathbf{a}(\mathbf{Q}) \\ - \quad (4\pi/c)j\delta(\mathbf{Q}-\mathbf{Q}_p)\mathbf{a}(\mathbf{Q}_p)/\mid \mathbf{a}(\mathbf{Q}_p)\mid \quad = \quad 0\end{aligned} \tag{III.8}$$

where ω is the resonant frequency of the cavity. Multiplying by $\mathbf{a}(Q)$, integrating over Q, and taking Fourier transforms, (III.8) becomes

$$(\nu^2 - \omega^2 + i\nu\omega/Q)X_\nu - c^{-1}\mid \mathbf{a}(\mathbf{Q}_p)\mid j_\nu = 0. \tag{III.9}$$

We find that the ratio (III.5)

$$i\nu Z_\nu = [\nu^2 - \omega^2 + (i\nu\omega/Q)]c^2/\mid \mathbf{a}(\mathbf{Q}_p)^2$$

The influence phase for this is (although it is unnecessary to write it)

$$\Phi(j, j') = \frac{1}{2\pi\hbar}\int_0^\infty \left[\frac{j'_\nu(j_{-\nu} - j'_{-\nu})}{(i\nu Z_\nu)} + \frac{j_{-\nu}(j_\nu - j'_\nu)}{(-i\nu Z_{-\nu})}\right] d\nu$$

From second order perturbation theory we know

$$P_{nm} = 2T\mid j_{nm}\mid^2 (\hbar\nu_{nm})^{-1}\,\mathrm{Re}(Z_{nm})^{-1} \quad \text{for} \quad \nu_{nm} > 0$$

[28] Q is used here as the dissipation factor of the cavity, $\omega_0 L/R$, while $\mathbf{Q}$ is a vector representing the linear coordinates inside the cavity.

Noting that

$$\mathrm{Re}(Z\nu_{nm})^{-1} = \omega\nu_{nm}^2 \mid \mathbf{a}(\mathbf{Q}_p)\mid^2 Q^{-1}c^{-2}[(\nu_{nm}^2 - \omega^2)^2 + \omega^2\nu_{nm}^2 Q^{-2}]^{-1}$$

and defining a cavity from factor $f^2 = V \mid \mathbf{a}(\mathbf{Q}_p)\mid^2/4\pi c^2$

$$P_{nm} = (8\pi \mid j_{nm}\mid^2 f^2\omega\nu_{nm}T)(\hbar VQ)^{-1}[(\nu_{nm}^2 - \omega^2) + \omega^2\nu_{nm}^2Q^{-2}] \quad \text{(III.10)}$$

At resonance this expression reduces to

$$P_{nm} = 8\pi \mid j_{nm}\mid^2 f^2QT/\hbar V\nu_{nm}^2 \quad \text{(III.11)}$$

The quantity usually computed is the ratio of the transition probability in the cavity at resonance to that in free space. This ratio is

$$\frac{P_{nm}\,(\text{cavity})}{P_{nm}\,(\text{free space})} = \frac{[8\pi \mid j_{nm}\mid^2 Qf^2/\nu_{nm}^2 V\hbar]T}{[4 \mid j_{nm}\mid^2\nu_{nm}/3\hbar c^3]T} = \frac{6\pi c^3Qf^2}{V\nu_{nm}^3}$$

At resonance, the ratio increases with respect to Q as one might expect and decreases with respect to the cavity volume V and ν_{nm}^3. This expression agrees with the one given by E. M. Purcell (*23*) although the form factor in his calculation was left out. This does not matter since for a particle located near the maximum field point in a cavity the magnitude of f is of the order of unity.

APPENDIX IV

It is to be demonstrated that

$$G(X) = \lim_{N\ \text{large}} \prod_{k=1}^{N}(1 + X_k) \to \exp[\textstyle\sum_k X_k] \quad \text{(IV.1)}$$

where the X_k are small but not necessarily equal to each other, and where the total sum $\sum_k X_k$ is finite. Rewriting the expression for $G(X)$ we have (where the summations on all indices go from 1 to N)

$$\begin{aligned} G(X) &= 1 + \textstyle\sum_k X_k + \tfrac{1}{2}\sum_{j\neq k} X_jX_k + (1/6)\sum_{j\neq k\neq \ell} X_jX_kX_\ell + \cdots \\ &= 1 + \textstyle\sum_k X_k + \tfrac{1}{2}!\sum_{j,k} X_jX_k(1 - \delta_{jk}) + (\tfrac{1}{3}!)\sum_{j,k,\ell} X_jX_kX_\ell \\ &\quad \times (1 - \delta_{jk} - \delta_{k\ell} - \delta_{j\ell} + 2\delta_{jk}\delta_{k\ell}) \end{aligned} \quad \text{(IV.2)}$$

As N is allowed to get very large the contribution of the terms involving quantities such as δ_{jk} becomes less significant. For instance, in the third term

$$\textstyle\sum_{j,k}(X_jX_k)(1 - \delta_{jk}) \sim (NX)^2 - (NX)^2/N \quad \text{(IV.3)}$$

and for very large N only the leading term in this sum is important. Thus, we have the result that

$$\begin{aligned} G(X)_{\text{large } N} &\approx 1 + \textstyle\sum_k X_k + \tfrac{1}{2}!(\sum_k X_k)^2 + \tfrac{1}{3}!(\sum_k X_k)^3 \\ &\approx \exp[\textstyle\sum_k X_k] \end{aligned} \quad \text{(IV.4)}$$

RECEIVED: April 5, 1963

REFERENCES

1. R. P. FEYNMAN, *Rev. Mod. Phys.* **20,** 367 (1948).
2. R. P. FEYNMAN, *Phys. Rev.* **80,** 440 (1950).
3. R. P. FEYNMAN, *Phys. Rev.* **84,** 108 (1951).
4. V. FANO, *Rev. Mod. Phys.* **29,** 74 (1957).
5. R. W. HELLWARTH (private discussion).
6. R. MCWEENY, *Rev. Mod. Phys.* **32,** 335 (1960).
7. R. C. TOLMAN, "The Principles of Statistical Mechanics," Chapter 9. Oxford Univ. Press, New York, 1930.
8. D. MIDDLETON, "An Introduction to Statistical Communication Theory." McGraw-Hill, New York, 1960.
9. D. MIDDLETON, "An Introduction to Statistical Communication Theory," pp. 362–364. McGraw-Hill, New York, 1960.
10. H. GOLDSTEIN, "Classical Mechanics," Chap. 10. Addison Wesley, Reading, Mass, 1953.
11. H. B. CALLEN AND J. H. WELTON, *Phys. Rev.* **83,** 34 (1951).
12. R. KUBO, *J. Phys. Soc. Japan* **12,** 570 (1957).
13. H. NYQUIST, *Phys. Rev.* **32,** 110 (1928).
14. H. NAKANO, *Progr. Theoret. Phys. (Kyoto)* **17,** 145 (1957).
15. R. V. POUND, *Ann. Phys. (N.Y.),* **1,** 24 (1957).
16. M. W. MUELLER, *Phys. Rev.* **106,** 8 (1957).
17. K. SHINODA, T. C. WANG, AND C. H. TOWNES, *Phys. Rev.* **102,** 1308 (1956).
18. M. W. P. STRANBERG, *Phys. Rev.* **106,** 617 (1957).
19. K. SHINOLA, H. TAKAHASI, AND C. H. TOWNES, *J. Phys. Soc. Japan* **12,** 686 (1957).
20. W. H. WELLS, *Ann. Phys. (N.Y.),* **12,** 1 (1961).
21. J. P. GORDON, *Proc. IRE* **50,** 1898 (1962).
22. E. FERMI, *Rev. Mod. Phys.* **4,** 87 (1932).
23. E. M. PURCELL, *Phys. Rev.* **69,** 681 (1946).

IV. Liquid Helium

This commentary is supplied by: Alexander L. Fetter, Physics Department, Stanford University.

Between 1953 and 1958, Feynman published a seminal series of papers on the atomic theory of superfluid helium. Superfluidity in liquid helium had been discovered in the 1930's, and the early understanding of this phenomenon relied on Landau's phenomenological theory (1941, 1947) of phonons and rotons as elementary excitations ("quasiparticles"). In this context, the Bose–Einstein statistics of helium atoms (and the existence of a Bose–Einstein condensate) played essentially no role. [For a summary of the early history see, for example, A. Griffin, in *Bose–Einstein Condensation in Atomic Gases*, edited by M. Inguscio, S. Stringari, and C.E. Wieman (Italian Physical Society, 1999).] A significant part of Feynman's central contribution was the demonstration that these phenomenological concepts arose directly from the fundamental quantum mechanics of interacting bosonic atoms with strong repulsive cores.

One of his earliest helium papers [21] showed in detail how the symmetric character of the many-body wave function severely restricts the allowed class of low-lying excited states. Specifically, all such states correspond to density fluctuations (phonons), and he then proposed [24] the explicit model wave function $\psi_{\mathbf{k}} = \rho^{\dagger}_{\mathbf{k}}\phi$, where $\varphi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ is the symmetric exact ground state wave function of the N helium atoms with energy E_0, and $\rho^{\dagger}_{\mathbf{k}} = \Sigma_{j=1}^{N} \exp(i\mathbf{r}_j \cdot \mathbf{k})$ is the symmetric operator that creates a density fluctuation. Here, φ is real, nodeless, and vanishes whenever any two atoms approach each other closer than an effective hard core diameter, ensuring that $\psi_{\mathbf{k}}$ indeed incorporates the proper many-body correlations arising from the hard cores and associated excluded volume. In addition, $\psi_{\mathbf{k}}$ is an eigenstate of the total momentum operator with the eigenvalue $N\hbar\mathbf{k}$, so that states with different eigenvalues are orthogonal. Hence $\psi_{\mathbf{k}}$ serves as a suitable variational trial function that yields an upper bound to the excitation energy $\varepsilon_k = E_k - E_0$ as a function of the continuous variable k.

In this way, Feynman obtained the elegant and simple expression $\varepsilon_k \approx \hbar^2 k^2/2mS(k)$, where $S(k)$ is the static structure function that had been measured independently (for example, with X-ray scattering). At long wavelengths ($k \to 0$), $S(k) \to \hbar k/2ms$, where s is the speed of compressional sound waves, reproducing Landau's linear quasiparticle spectrum $\varepsilon_k \approx \hbar sk$ as $k \to 0$. For large k, in contrast, $S(k) \to 1$, but it has a peak at $k \sim 2$ Å^{-1}, associated with the nearest neighbor separation; this latter feature produces a dip in the quasiparticle spectrum, qualitatively similar to Landau's conjectured roton minimum. In practice, the calculated position $\Delta \approx 19.1$ K of the minimum was roughly twice the value (9.6 K) that fit the heat capacity. Subsequently, Feynman and Cohen [32] improved the variational trial function by including the hydrodynamic "backflow" associated with the rapid motion of a single atom through the fluid, and they found the much lower value $\Delta \approx 11.5$ K. Finally, Cohen and Feynman [38] used their trial function to study the inelastic neutron scattering from superfluid helium, predicting that the scattering would be dominated by the excitation of a single quasiparticle, allowing a direct measurement of energy spectrum ε_k (subsequent measurements fully confirmed this behavior).

Separately, Feynman [27] considered states with macroscopic superfluid flow, arguing that they must have the simple form $\psi_{\text{flow}} = \exp[i\Sigma_{j=1}^{N} s(\mathbf{r}_j)]\phi$, where $s(\mathbf{r})$ is a function that varies only slowly over an interatomic distance. The resulting single-particle current

corresponds to a velocity field $v_s = \hbar \nabla s/m$, which agrees with Landau's assertion that the superfluid velocity is necessarily irrotational. Feynman recognized that the single-valuedness of the many-body wave function implies the quantization of circulation $\oint \mathbf{v}_s \cdot d\mathbf{s}$ in units of $2\pi\hbar/m == h/m$, as suggested by Onsager (1949). He then considered the detailed behavior of a single quantized vortex line, where the centrifugal barrier near the center creates a hole of radius a (a few atomic diameters); as a result, he proposed that bulk superfluid helium rotating at an angular speed ω is filled by an array of such linear vortices with an areal density $2\omega m/h$. Finally, he suggested that the creation of vortices can account for the observed low values of the critical velocity for the onset of dissipation. Landau's original analysis of the critical velocity associated with the creation of a roton suggested a value of order 60 m/s, whereas the measured values were at least 100 times smaller and depended on the geometrical configuration of the fluid. For flow from an aperture of lateral width d, Feynman's vortex model yielded $v_c \sim (\hbar/md)\ln(d/a)$, in reasonable agreement with experiments.

In fact, Feynman's first substantial paper [20] on helium dealt with the λ transition at $T_\lambda \approx 2.17$ K, which signals the formation of the new phase He II (and the onset of superfluidity); this paper is necessarily quite different from those focusing on the low temperature behavior discussed above. Feynman expressed the exact quantum-mechanical partition function as a path integral and then mapped it rigorously onto a classical polymer problem. The resulting picture of superfluidity near the λ transition relies on the appearance of macroscopic ring exchanges that depend critically on the Bose–Einstein statistics. He showed that the strong interactions do not qualitatively change the transition temperature T_λ. This paper was far ahead of its time, as the detailed implementation of its theoretical program required the development of high speed computers. Modern path integral Monte Carlo techniques (in effect, following Feynman's formulation) have yielded a quantitative theory of bulk liquid helium. [For a review, see, for example, D.M. Ceperley, *Rev. Mod. Phys.* **67**, 279 (1995).]

Selected Papers

[20] Atomic theory of the lambda transition in liquid helium. *Phys. Rev.* **91** (1953): 1291–1301.

[21] Atomic theory of liquid helium near absolute zero. *Phys. Rev.* **91** (1953): 1301–1308.

[24] Atomic theory of the two-fluid model of liquid helium. *Phys. Rev.* **94** (1954): 262–277.

[27] Application of quantum mechanics to liquid helium. *Progress in Low Temperature Physics*, Vol. 1. Amsterdam, North-Holland (1954), pp. 17–53.

[32] With M. Cohen. Energy spectrum of the excitations in liquid helium. *Phys. Rev.* **102** (1956): 1189–1204.

[38] With M. Cohen. Theory of inelastic scattering of cold neutrons from liquid helium. *Phys. Rev.* **107** (1957): 13–24.

THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 91, NO. 6 SEPTEMBER 15, 1953

Atomic Theory of the λ Transition in Helium

R. P. FEYNMAN
California Institute of Technology, Pasadena, California
(Received May 15, 1953)

It is shown from first principles that, in spite of the large interatomic forces, liquid He^4 should exhibit a transition analogous to the transition in an ideal Bose-Einstein gas. The exact partition function is written as an integral over trajectories, using the space-time approach to quantum mechanics. It is next argued that the motion of one atom through the others is not opposed by a potential barrier because the others may move out of the way. This just increases the effective inertia of the moving atom. This permits a simpler form to be written for the partition function. A rough analysis of this form shows the existence of a transition, but of the third order. It is possible that a more complete analysis would show that the transition implied by the simplified partition function is actually like the experimental one.

INTRODUCTION

THE behavior of liquid helium, especially below the λ transition, is very curious.[1] The most successful theoretical interpretations,[2] so far, have been largely phenomenological. In this paper and one or two to follow, the problem will be studied entirely from first principles. We study the quantum-mechanical behavior of strongly interacting atoms of He^4. We shall try to show that the main features of these curious phenomena can, in fact, be understood from this point of view. Because of the enormous geometrical complexity involved, we shall not attempt to obtain useful quantitative results. The quantum mechanics will not supplant the phenomenological theories. It turns out to support them.

In this paper we begin the study of the statistical mechanics of the liquid.[3]

London[4] has proposed that the transition between liquid He I and liquid He II is a result of the same process which causes the condensation of an ideal Bose-Einstein gas. This idea could be criticized on the grounds that the strong forces of interaction between the He atoms might make the ideal gas approximation (in which these forces are neglected) even qualitatively incorrect. We shall argue that London's view is essentially correct. The inclusion of large interatomic forces will not alter the central features of Bose condensation.

The principal point is an argument which shows that in a liquid-like quantum-mechanical system the strong interactions between particles do not prevent these particles from behaving very much as though they move freely among each other.

The exact partition function is first written down as an integral over trajectories, by using the space-time approach to quantum mechanics.[5] The observation that the atoms move very freely among each other is then made. This permits one to write a simpler form [Eq. (7)] for the partition function. This form should be fairly accurate, at least qualitatively. It becomes clear that a transition is to be expected, and that it involves the symmetrical statistics in an essential way.

On the other hand, the geometrical complexity of the problem still prevents us, so far, from giving a very good estimate of the free energy behavior near the transition point and below. A relatively crude approach gives a transition like that of the ideal gas. That is, the specific heat is continuous, contrary to the experimental observation that it appears to be discontinuous. Some of the geometrical problems which might have to be solved to obtain a more satisfactory solution are discussed in an appendix (see also reference 3).

The crude approach should, however, be quite satisfactory a little above the transition point. So there is

[1] W. H. Keesom, *Helium* (Elsevier Publishing Company, Inc., Amsterdam, 1942).
[2] An excellent summary of the theories of helium II is to be found in R. B. Dingle, Advances in Phys. 1, 112 (1952).
[3] A preliminary report on this work has been published. R. P. Feynman, Phys. Rev. 90, 1116 (1953).
[4] F. London, Phys. Rev. 54, 947 (1938).
[5] R. P. Feynman, Revs. Modern Phys. 20, 367 (1948).

no doubt that at least the existence of the rise of specific heat[1] of He I on cooling toward the λ point can be understood from first principles.

At the opposite extreme of very low temperatures (say below 0.5°K), the situation again can be partially analyzed. This is done in the next paper.[6]

EXACT EXPRESSION FOR THE PARTITION FUNCTION

To study the thermodynamic properties we must calculate the partition function

$$Q=\sum_i \exp(-\beta E_i), \tag{1}$$

where $\beta=1/kT$ and E_i are the energy levels of the system. In this form the calculation appears hopelessly difficult because the energies E_i are eigenvalues of such a complex Hamiltonian H. The expression for Q is equivalent to the trace of the operator $\exp(-\beta H)$. In Eq. (1) the trace is written in a representation in which H is diagonal. We shall prefer to use the coordinate representation to describe the trace.

To illustrate how this is done, we take the example of a one-dimensional system, of coordinate x and Hamiltonian $p^2/2m+V(x)=H$. The trace of $\exp(-\beta H)$ is then $Q=\int dz(z|e^{-\beta H}|z)$. The matrix element $(z|e^{-\beta H}|z)$ is similar in form to the matrix element $(z|\exp(-itH/\hbar)|z)$ which represents the amplitude that the system initially at $x=z$, is at time t also at the point $x=z$. This latter is[5] the sum over all paths [signified by $\int\cdots\mathfrak{D}x(t)$] which go from z to z of $\exp(iS/\hbar)$, where S is the action $\int_0^t[\frac{1}{2}m\dot{x}^2-V(x(t))]dt$. If we replace $it/\hbar$ by β, we are lead to expect

$$(z|e^{-\beta H}|z)=\int_{tr}\exp\left\{-\int_0^\beta\left[\frac{m}{2\hbar^2}\left(\frac{dx}{du}\right)^2+V(x(u))\right]du\right\}\mathfrak{D}x(u), \tag{2}$$

the variable $u=it/\hbar$ replacing t, and the various signs adjusted accordingly. The integral $\int_{tr}$ is to be taken on all trajectories such that $x(0)=z$ and $x(\beta)=z$. It is easily verified that Eq. (2) is exactly correct. The normalization of the path integral is to be such that

$$\int_{tr'}\exp\left[-\int_0^\beta\frac{m}{2\hbar^2}\left(\frac{dx}{du}\right)^2du\right]\mathfrak{D}x(u)=(m/2\pi\beta\hbar^2)^{\frac{1}{2}}\exp\left[-\frac{m}{2\beta\hbar^2}(z-z')^2\right], \tag{3}$$

where the trajectory tr' runs from $x(0)=z$ to $x(\beta)=z'$. The integral of (2) with respect to z then gives the partition function.

To apply this to liquid helium two modifications are necessary. First, instead of one variable, we have $3N$ variables which we take as the N three-space coordinate vectors $\mathbf{x}_i$ of each of the N atoms ($i=1$ to N). We designate the entire set of coordinates by $\mathbf{x}^N$ and the integral over them all by $d^N\mathbf{x}_i=d\mathbf{x}_1d\mathbf{x}_2\cdots d\mathbf{x}_N$. The initial and final values of these we call $\mathbf{z}_i$. Secondly, He^4 atoms obey symmetrical statistics. The trace of $\exp(-\beta H)$ is to be taken only over symmetrical wave functions. This means that if the initial coordinates are $\mathbf{x}_i(0)=\mathbf{z}_i$, the final coordinates need not be the same, but may be some permutation of these (signified by $P\mathbf{z}_i$). That is,

$$Q=N!^{-1}\sum_P\int d^N\mathbf{z}_i\int_{trP}\exp\left\{-\int_0^\beta\left[\frac{m}{2\hbar^2}\sum_i\left(\frac{d\mathbf{x}_i}{du}\right)^2+\sum_{ij}V(\mathbf{x}_i-\mathbf{x}_j)\right]du\right\}\mathfrak{D}^N\mathbf{x}_i(u), \tag{5}$$

where the integral $\int_{trP}$ is taken over all trajectories $\mathbf{x}_i(u)$ of all the particles such that $\mathbf{x}_i(0)=\mathbf{z}_i$, $\mathbf{x}_i(\beta)=P\mathbf{z}_i$. That is, the final coordinates $\mathbf{x}_i(\beta)$ may now be some permutation P of the initial coordinates $\mathbf{z}_i$. The sum is taken over all permutations P and the integral over all configurations $\mathbf{z}_i$.

In Eq. (5), m is the mass of a helium atom, and $V(\mathbf{R})$ is the mutual potential of a pair of He atoms separated by $\mathbf{R}$. The forces between He atoms are, very likely, fairly accurately two-body forces. This potential is given by Slater and Kirkwood.[7] There is a weak attraction of maximum depth (of energy equivalent to kT at $T=7$°K) at radius about 3.0A. The atomic volume at the transition is $(3.6\text{A})^3$. At 3.6A, $V(\mathbf{R})$ is about equivalent to kT for $T=3$°K. There is, therefore, a weak attraction at the average atomic distance. There is a violent repulsion if the atoms approach more closely than 2.6A ($V=0$ at 2.6A).

The expression (5) is an exact[8] quantum-mechanical expression for the partition function (even though no imaginary unit i appears). We shall use it to develop a qualitative understanding of this function for liquid helium.

The quality u is of course not the time. However we shall obtain a vivid representation of (5) by imagining that it is the time. We can say that at one time 0 the coordinates $\mathbf{x}_i(0)$ of all the atoms form an initial configuration $\mathbf{z}_i$, and that as time u proceeds the particles move about $[\mathbf{x}_i(u)]$ in such a manner that at the time β the configuration of atoms appears to be the same (although in fact some of the atoms may have been interchanged). Each mode of motion is weighed by the negative exponential of the time integral of the energy required for the motion, and the sum is taken for all such motions. Finally an average (or rather integral) is taken over all possible initial configurations $\mathbf{z}_i$.

We can see immediately that motions which require too large a displacement in the time β have little weight

[6] R. P. Feynman, Phys. Rev. 91, 1301 (1953).

[7] J. C. Slater and J. G. Kirkwood, Phys. Rev. 37, 682 (1931).

[8] In so far as the forces can be represented as two-body forces.

because of the high kinetic energy required. Likewise, motions in which the atoms come so close as to appreciably penetrate the radius of their repulsion are of small importance because of the large potential energy which would result. For this reason, also, initial configurations for which the atoms overlap, that is, have centers so close that they would repel, contribute only a small amount. The method of approximation which we shall apply to (5) is to neglect the contributions from motions $\mathbf{x}_i(u)$ and configurations $\mathbf{z}_i$ which give small contributions, and to study more carefully only those motions which give the larger contributions to the total in (5). That is, we shall have motions in which the atoms do not move too fast or far in the time β and in which the atoms never overlap.

We emphasize again that these "motions" must not be construed as a real description of what the atoms are doing. It is simply a formal description of the expression for the partition function. An expression "the atom does not move too far in the time β" does not refer to a real motion as u is not time, but is $it/\hbar$. The true behavior of the atoms may have some analogy to the description of the formula (5), but such an analogy need not concern us here. Our reason to continue to call u and β "time" is to help to make our arguments as vivid as possible so that intuition will be most effective.

THE CHARACTER OF THE IMPORTANT TRAJECTORIES

Consider a particular motion in (5) in which some atom i moves to the site initially occupied by atom j. Call the displacement $\mathbf{a}=\mathbf{z}_j-\mathbf{z}_i$. The atom j must, of course, move to some other site to leave room for i. The effect of the motion of j we will associate with atom j. We study here the contribution to be expected just from the displacement of the single atom i by a distance $\mathbf{a}$.

Near the transition temperature displacements larger than about d, the atomic spacing (cube root of atomic volume 3.6A) are not very important [$\exp(-md^2/2\beta\hbar^2)=0.3$ at 2.2°K]. Nevertheless we will try to get an idea of the behavior for larger displacements. These will be useful at lower temperatures. Actually our considerations apply to displacements of any size.

Suppose, then, atom i must make a translation $\mathbf{a}$ of length a. We make this, for example, to be nearly along a straight line. Our arguments will apply for any other route.

The central problem is, what is the effect of the potentials of interaction on this translation? As a simple model which retains the essential features imagine the atoms as hard impenetrable spheres. We are, during a time β, to move atom i from $\mathbf{z}_i$ to $\mathbf{z}_j=\mathbf{z}_i+\mathbf{a}$ and at the end to leave all the other atoms in their original positions. The atoms may not overlap at any time.

There may be atoms in the direct line from $\mathbf{z}_i$ to $\mathbf{z}_j$. Nevertheless, a moment's reflection shows that they will not offer a real potential barrier to the translation of atom i.

It is evident that it is possible to place atom i at any position on the route from $\mathbf{z}_i$ to $\mathbf{z}_j$, provided we readjust the positions of the other atoms to make room for it. In the readjusted positions the total potential energy of all the atoms can be made to be very nearly equal to the potential energy of the original configuration. Therefore, atom i can be moved to any intermediate position without violating any repulsive potential, in fact, without any appreciable modifications of potential energy at all. It is only necessary to move the other atoms around out of the way as atom i moves along. When i reaches the final destination $\mathbf{z}_j$, the other atoms (except j of course, as noted above) may all come back to their original positions (or to some permutation thereof[9]).

The readjustment of the other atoms means that their coordinates $\mathbf{x}_k(u)$ change with time. They contribute just kinetic energy in the exponent in Q. Beside the kinetic energy $\frac{1}{2}m(a/\beta)^2$ needed to move atom i a distance a in time β, we have also to add the kinetic energy of the readjusting atoms. This we can expect will also vary directly as the square of the velocity of atom i. The net effect is an energy of the form $\frac{1}{2}m'(a/\beta)^2$, where m' is an effective mass, somewhat larger than the mass of a single atom m. The difference represents the effective inertia of the atoms which are readjusting.

The time integral of the energy needed for readjustment varies with a and β, as a^2/β. This is clear for small displacements of i, for then only a few atoms shift, and they do this with a velocity proportional to that of i. For large displacements ($a\gg d$) the same form, of course, results. To verify this, imagine that as atom i moves at velocity $v(=a/\beta)$ the time it passes near a particular atom is of order d/v. This atom must adjust through distances of order d in this time, or move at speed about v. The time integral of energy needed for this passage is about $mv^2(d/v)=mdv$. The number of such atoms which must be jostled to move a total distance a is of order a/d. Thus the total time integral of energy of readjustment for all these a/d atoms varies as $mdv\ a/d=mva=ma^2/\beta$. It varies with a and b in the required manner.

The effect of the other atoms is not to offer a potential barrier (time integral varying as $a\beta$) but a kind of kinetic energy barrier (time integral varying as a^2/β). The effect of the interactions is taken into account by changing the effective mass of a given moving atom.

To get some idea of the order of m', we recall that the effective mass of a sphere moving through an ideal fluid of the same density is classically $(\frac{3}{2})m$, the extra $\frac{1}{2}$ being the energy of motion of the fluid making way for the sphere. The effect of the attractive forces may in-

[9] The effect of a moving atom in permuting other atoms might have to be considered in more detail if we were to apply these ideas to the case of Fermi-Dirac statistics. It may mean that the m' (discussed further on) is somewhat larger in that case.

crease this somewhat. We may expect m' not to be very much greater than m—perhaps not more than 2 or 3 m.

The effect of the relatively weak attractive potentials may be to alter the motions a bit, in that adjacent atoms tend to stick together a little. Thus the atom i may have a tendency to drag some atoms with it from time to time, possibly increasing m' a little more. On the other hand, in He the zero-point energy[10] is high enough to shake these others loose readily. If the potential were much stronger the group attraction might become accululative, raising m' very much. It is possible that this would result in solidification.

For short displacements a of order d, proportionately less adjustment need be made, so that it is likely that m' may be somewhat less. For high velocities, it may represent less energy to violate the real potential restrictions a little. Thus the readjustments need not be complete, so that m' again may decrease a little with velocity a/β, approaching m as $a/\beta\to\infty$.

As this is meant to be a first approach to the problem, we shall not attempt to calculate the m'. The geometrical complexity is very great. Further, we shall neglect the variation of m' with a and velocity. It is to be expected that this neglect may not alter our conclusions qualitatively. It is always possible, later, to include such finer details. Nor shall we discuss the variation (expected rise) of m' with increasing density of the fluid.

For every trajectory the atom acts like a free particle of effective mass m'. Hence we may take the integral over all paths $\mathbf{x}_i$ for atom i to go a distance a to be proportional to

$$(m'/2\pi\beta\hbar^2)^{\frac{1}{2}}\exp(-m'a^2/2\beta\hbar^2). \tag{6}$$

The normalization factor has been written as $(m'/2\pi\beta\hbar^2)^{\frac{1}{2}}$ for convenience. That it varies as $\beta^{-\frac{1}{2}}$ may be shown by dimensional analysis [compare Eq. (3)]. Actually a change in this factor will just change the partition function by a factor. It will be easiest to discuss the normalization of the entire partition function.

Therefore, we can approximate Q by

$$Q=\frac{K_\beta}{N!}\left(\frac{m'}{2\pi\beta\hbar^2}\right)^{3N/2}\int\sum_P\exp\left[-\frac{m'}{2\beta\hbar^2}\sum_i(\mathbf{z}_i-P\mathbf{z}_i)^2\right]\times\rho(\mathbf{z}_1,\mathbf{z}_2,\cdots\mathbf{z}_N)d^N\mathbf{z}_i. \tag{7}$$

The factor K_β we shall estimate later. The function $\rho(\mathbf{z}_1,\cdots\mathbf{z}_N)=\rho(\mathbf{z}^N)$ represents a density associated with each configuration. It is discussed in the next section. This expression (7) for Q may be rewritten, using Fourier transforms as

$$Q=\int F(\mathbf{k}^N)\exp[-(\beta\hbar^2/2m')\sum_i\mathbf{k}_i^2]d^N\mathbf{k}_i(2\pi)^{-3N}, \tag{7a}$$

where

$$F(\mathbf{k}^N)=K_\beta(N!)^{-1}\int\sum_P\exp[i\sum_i\mathbf{k}_i\cdot(\mathbf{z}_i-P\mathbf{z}_i)]\times\rho(\mathbf{z}^N)d^N\mathbf{z}_i.$$

This form is especially useful near absolute zero, but we will not need it in this paper.

It should be emphasized that the argument which leads to the free particle approximation for the motion of an atom is of greater generality. The argument results simply from a consideration of the limitations to the true trajectories which result from the interatomic potentials. They therefore apply as well to the nondiagonal element of $\exp(-\beta H)$, as to the diagonal element which appears in Q. Likewise, they are applicable to the true quantum-mechanical kernel, which is the nondiagonal element of $\exp(-itH/\hbar)$. The imaginary weights in this case also restrict the atoms not to overlap at low energy, etc.

This principle may have uses in other branches of physics, for example, in nuclear physics. Here there is the puzzling fact that single nucleons often act like independent particles in spite of strong interactions. The arguments we have made for helium may apply to this case also.[9]

THE CONFIGURATIONS OF IMPORTANCE

Not all configurations $\mathbf{z}_N$ are to be weighed equally. If $\mathbf{z}_i$ and $\mathbf{z}_j$ are closer than the distance at which strong repulsion sets in (2.6A) the configuration should be given very little weight [i.e., ρ in (7) is nearly zero if atoms overlap]. We shall discuss this effect first for the case of low temperature (β large).

Suppose initially in (5) two atoms overlap, say by a distance x, and suppose that this results in an extremely high (relative to $1/\beta$) potential V. If they move a distance x further apart, suppose V goes to 0, and take V independent of x for simplicity. During the interval $u=0$ to β, if the atoms remain overlapped for a time τ, the contribution is the negative exponential of $V\tau$. This contribution is extremely small unless τ is very short (if $V\beta\gg1$ then $\tau\ll\beta$). The most important trajectories are then those that release V as quickly as possible. This can only be done by a high kinetic energy $m'x^2/2\tau^2$. Thus the integral of energy has the value $m'x^2/2\tau\hbar^2+V\tau$, which is least if $\tau=(m'x^2/2V\hbar^2)^{\frac{1}{2}}$, in which case it has the value $(2m'V)^{\frac{1}{2}}x/\hbar$. The contribution varies as $\exp[-(2m'V)^{\frac{1}{2}}x/\hbar]$. This is just the quantum-mechanical penetration factor. In ρ it appears twice, for again we must get into the overlapped condition at the end of the interval β. This argument fails if β does not exceed

[10] The zero-point energy referred to appears in the integral for Q in the following guise. Suppose we restrict the motion of a certain atom k so that, for example, it tries to move along close to i to take advantage of some extra attractive potential between k and i. Then the trajectories of k are restricted, and we lose a great deal in the integral over possible paths of k because we are not adding contributions from very many paths. We lose "volume in path-space." This will suffice to offset the attraction if, in the conventional language, the zero-point motion is sufficiently large.

$\lambda\tau$. In that case a larger penetration results. It is due, of course, to the high kinetic energy that such a small β implies. For the large repulsions V at low temperatures involved here this penetration is very small.

In addition, ρ would not be quite uniform even if there is no overlap of the atoms, and even if they are considered as impenetrable spheres. In fact, ρ would be larger if atoms are well spaced than if they are nearly adjacent (for large β). If two atoms are adjacent initially, the available paths are limited to those which move them apart—for they must not come to overlap. This decreases the effective volume of path space for a short time. Actually this effect is offset partly by the actual attractive potential which results upon the closer approach of the two atoms. Thus ρ represents the effects of short-time adjustments (times$\ll\beta$) while the longer-time effects[11] are contained in the exponential factor in (7) [that is, expression (12) below]. For large β, low T, ρ can be taken as nearly temperature independent, and the main temperature dependence comes from the other factor (12). This ρ as $\beta\to\infty$ is the density corresponding to the ground-state wave function. For approximate purposes we can take it to be simply the density function ρ for a classical gas of impenetrable atoms of diameter b. That is, $\rho=0$ if any two $\mathbf{z}_i$'s are closer to each other than b, and is 1 otherwise. This neglects the variations with distance due to the quantum-mechanical effect discussed above of restricted path-space volume, and due to the attractive part of the potential.

For high temperatures, the exponential terms in (7), representing diffusion, are unimportant, and ρ should approach the classical distribution function. Now the attractive forces are weak and unimportant so again can be roughly represented by an impenetrable sphere model. The radius b should be somewhat smaller. We shall neglect this variation of b with temperature.

To summarize, ρ is qualitatively similar to the density distribution of a classical gas. It changes somewhat with temperature.

PROPERTIES OF THE PARTITION FUNCTION

A partition function has several formal properties, and we may test our approximate expression (7) to see how well it satisfies these conditions. Another important function is the nondiagonal element of $\exp(-\beta H)$. That is, $G_\beta(z', z)=(z'|\exp(-\beta H)|z)$, or

$$G_\beta(\mathbf{z}'^N, \mathbf{z}^N)=(N!)^{-1}\sum_P\int_{tr_{P'}}\exp\Big[-\int_0^\beta\Big(\frac{m}{2\hbar^2}\sum_i\Big(\frac{d\mathbf{x}_i}{du}\Big)^2+\sum_{ij}V(\mathbf{x}_i-\mathbf{x}_j)\Big)du\Big]\mathfrak{D}^N\mathbf{x}_i(u), \quad (8)$$

the integral being taken over all trajectories $tr_{P'}$ such that $\mathbf{x}_i(0)=\mathbf{z}_i$, $\mathbf{x}_i(\beta)=P\mathbf{z}_i'$. The final configuration $\mathbf{z}_i'$ may differ from the initial configuration. Of course, $Q=\int G_\beta(\mathbf{z}^N, \mathbf{z}^N)d^N\mathbf{z}_i$.

For large β we have given an argument for behavior of the function $\rho(\mathbf{z}^N)$, which represented it as the square of a function, say $\phi(\mathbf{z}^N)$. One factor was for leaving an unfavorable (say overlappping) configuration. The second was for entering it again. At low temperature $\phi(\mathbf{z}^N)$ is the ground-state wave function. The same arguments give for $G_\beta(\mathbf{z}'^N, \mathbf{z}^N)$ the approximate expression,

$$G_\beta(\mathbf{z}'^N, \mathbf{z}^N)=K_\beta(N!)^{-1}\sum_P\phi(\mathbf{z}'^N)\phi(\mathbf{z}^N)\times\exp\Big[-\frac{m'}{2\beta\hbar^2}\sum_i(\mathbf{z}_i-P\mathbf{z}_i')^2\Big]\Big(\frac{m'}{2\pi\beta\hbar^2}\Big)^{3N/2}. \quad (9)$$

Since $\exp(-\beta_1 H)\exp(-\beta_2 H)=\exp[-(\beta_1+\beta_2)H]$, we have the condition of matrix multiplication, namely, that

$$G_{\beta_1+\beta_2}(\mathbf{z}'^N, \mathbf{z}^N)=\int G_{\beta_1}(\mathbf{z}'^N, \mathbf{z}''^N)G_{\beta_2}(\mathbf{z}''^N, \mathbf{z}^N)d^N\mathbf{z}_i''.$$

This requires that $(\beta=\beta_1+\beta_2)$

$$K_\beta\Big(\frac{m'}{2\pi\beta\hbar^2}\Big)^{3N/2}\sum_P\exp\Big[-\frac{m'}{2\beta\hbar^2}\sum_i(\mathbf{z}_i-P\mathbf{z}_i')^2\Big]$$
$$=K_{\beta_1}K_{\beta_2}\Big(\frac{m'}{2\pi\beta_1\hbar^2}\Big)^{3N/2}\Big(\frac{m'}{2\pi\beta_2\hbar^2}\Big)^{3N/2}\int N!^{-1}\sum_P\sum_{P'}$$
$$\times\exp\Big[-\frac{m'}{2\beta_1\hbar^2}\sum_i(P\mathbf{z}_i'-\mathbf{z}_i'')^2-\frac{m'}{2\beta_2\hbar^2}\sum_i(P'\mathbf{z}_j''-\mathbf{z}_j)^2\Big]\times[\phi(\mathbf{z}''^N)]^2d^N\mathbf{z}_i''.$$

Now, in the exponent we can relable i (by permuting the names) as $P'i$, since the sum in on all i. That is, $\sum_i(P\mathbf{z}_i'-\mathbf{z}_i'')^2=\sum_i(PP'\mathbf{z}_i-P'\mathbf{z}_i'')^2$. Now call $PP'=P''$, and the sum on all P is equivalent to a sum on all P''. Finally, since $\mathbf{z}_i''$ are variables of integration in a symmetrical factor $[(\phi(\mathbf{z}''^N)]^2=\rho(\mathbf{z}''^N)=\rho(P'\mathbf{z}''^N)$ for any P', the $\mathbf{z}''^N$ integral does not depend on P', and $\sum_{P'}$ just gives a factor $N!$.

Now the weight function $\rho(\mathbf{z}''^N)$ prevents various $\mathbf{z}_i''$ from being too close together. If, however, β_1 and β_2 are so low that $\exp(-m'd^2/2\beta_1\hbar^2)$ is fairly close to 1, the variations in the exponents due to this restriction

[11] There is a kind of distortion that takes a long time τ to release, namely, a general increase in density over a large area. This restricts the path-space volume for each atom in the area and results in a factor $e^{-\beta E}$ for each particle, where E is the excess energy per particle induced by compression. The energy E can only be released by moving many particles, distributed over the area. These density fluctuations are sound waves. If the wavelength of the fluctuation is $\lambda=2\pi/K$, the time needed to release it is $\tau=1/\hbar\omega=1/\hbar cK$, where c is the speed of sound. This exceeds β if $\lambda>2\pi\hbar T/\hbar c$ or $\lambda>2\pi\cdot 8$A for 2.2°K. (This exceeds the diffusion distance $d\approx 3.6$A even at 2.2°K, so will not be of concern to us near the transition.) Thus (7) is incomplete in that it does not correctly describe long wave-sound fluctuations. This matter is discussed in a subsequent paper.

is very small. [That is, in the integrand, $\rho(\mathbf{z}''^N)$ varies rapidly while the other factor is smooth.] Therefore, we replace $\rho(\mathbf{z}''^N)$ by an average value and integrate over all $\mathbf{z}''^N$. The effect of $\rho(\mathbf{z}''^N)$ just restricts the volume available to the configuration variables. Let us call

$$V_N=\int \rho(\mathbf{z}^N)d^N\mathbf{z}_i, \tag{10}$$

so that the average value of ρ is V_N/V^N. The integral on $\mathbf{z}_i''$ is now easy, and we find

$$K_{\beta_1+\beta_2}=K_{\beta_1}K_{\beta_2}V_N/V^N.$$

[This verifies our choice of the $\beta^{-\frac{3}{2}}$ dependence in (6).] This means that K_β must have the form

$$K_\beta=V^N V_N^{-1}e^{-\beta E_0}, \tag{11}$$

where E_0 is a constant. Such a constant means, in Q, a constant energy (the energy at absolute zero). We will not try to determine this energy. Let us measure energies above this as a zero level. Then it can be ignored. Our final partition function is then (17) with $K_\beta=V^N/V_N$.

For extremely large β, Q should approach 1 from its original definition as $\sum_i e^{-\beta E_i}$ and the choice $E_0=0$. For such large β, the sum on permutations P means that $P\mathbf{z}_i$ goes successively over every site, while the exponential $\exp[-(m'/2\beta\hbar^2)(\mathbf{z}_i-P\mathbf{z}_i)^2]$ varies smoothly. It is approximated by writing it as $\exp[-(m'/2\beta\hbar^2)\times(\mathbf{z}_i-\mathbf{z}_i')^2]$ and integrating over all $\mathbf{z}_i'$ but dividing by the atomic volume $V_A=V/N$ for each $\mathbf{z}_i'$. However, this does not take into account the restriction that all the $\mathbf{z}_i'$ are on different sites. So an additional $N!/N^N$ is needed. Thus the total factor is $N!/(NV_A)^N$ or $N!/V^N$. The integrals give $(2\pi\beta\hbar^2/m')^{\frac{3}{2}}$ per degree of freedom, so we see that Q approaches

$$Q\sim(K_\beta/V^N)\int\rho(\mathbf{z}^N)d^N\mathbf{z}_i=1$$

as $\beta\to\infty$, as required.

This value of K_β was obtained by an argument involving large β. Let us study its behavior for small values of β. For small β (high T) no permutation is important in (17) except the identity. For no atoms are closer than $b=2.6$A, and $(m'/2\beta\hbar^2)(\mathbf{z}_i-P\mathbf{z}_i)^2$ would be at least $b^2m'/2\beta\hbar^2$, if $P\mathbf{z}_i\neq\mathbf{z}_i$. For small β this results in a large negative exponent. Thus Q approaches the value

$$Q=K_\beta(N!)^{-1}(m'/2\pi\beta\hbar^2)^{3N/2}\int\rho(\mathbf{z}^N)d^N\mathbf{z}_i.$$

The correct limit, according to the classical theory, should be

$$Q=N!^{-1}(m/2\pi\beta\hbar^2)^{3N/2}\int\rho(\mathbf{z}^N)d^N\mathbf{z}_i.$$

Since m' approaches m, K_β must approach 1 as $\beta\to 0$.

This means that K_β must be a function of β which varies from V^N/V_N to 1 as β varies from large to small values.

An accurate quantitative analysis of this problem would require close attention to the variation with temperature and density of m', of K and of b (or, more completely, of ρ).

EXISTENCE OF THE TRANSITION

We can use this partition function (7) and the ideas associated with it to understand many of the properties of liquid helium. The behavior of the liquid at very low temperature (below 0.5°K) will concern us in the following paper.[6] Here we will study the behavior in the region of a few degrees and shall show that a transition should occur.

In the qualitative study of such a transition we need not concern ourselves with the continuous variations in the effective constants m', K_β, b. It might be well to remark, however, that 2.2°K the expression $mx^2/2\beta\hbar^2$ is unity for $x=3.4$A. This is just the order of the average spacing of the atoms. Therefore, we are not going to be involved in very long displacements, and it may be that m' does not differ too much from m.

It is not hard to understand that (17) gives a transition. If ρ were a constant it would be the same as the partition function for an ideal gas. The fact that ρ is not perfectly uniform cannot change this much.

To see in more detail how this transition arises, consider the factor

$$\sum_P\exp[-(m'/2\beta\hbar^2)\textstyle\sum_i(\mathbf{z}_i-P\mathbf{z}_i)^2] \tag{12}$$

in the partition function (7).

Each permutation may be divided into cycles. A cycle of length s is a chain of permutations, such as 1 goes to 2, 2 goes to 3, 3 goes to 4, etc. until $s-1$ goes to s and finally s goes to 1. Such a cycle contributes to a term in (12) the factor

$$\exp[-(m'/2\beta\hbar^2)(\mathbf{z}_{12}^2+\mathbf{z}_{23}^2+\cdots+\mathbf{z}_{s1}^2)], \tag{13}$$

where $\mathbf{z}_{ij}=\mathbf{z}_i-\mathbf{z}_j$ and $\mathbf{z}_1$, $\mathbf{z}_2$, etc., are the positions of the particular atoms in the cycle. The total contribution from a given permutation is the product of all these factors, one from each of its cycles. For a given configuration we are to sum such a product over all permutations, that is, over all possible ways of laying out cycles on the configuration.

Consider a permutation of a certain "type," that is, having a certain number of each kind of cycle. That is, P has n_1 cycles of length 1 (i.e., n_1 atoms are not permuted), n_2 cycles of length 2, $\cdots n_s$ of length s, etc. The total number of atoms is N, so that

$$N=\textstyle\sum_s sn_s. \tag{14}$$

To these cycles there correspond n_1 atoms, n_2 polygons of 2 sides, n_3 triangles, $\cdots$, etc. drawn on the configuration, and each contributes its factor (13).

Next we may sum over all permutations of the same type. This means that each polygon will change its shape and location (but not its number of sides) as we go from one permutation to another in the sum. Eventually a given polygon can be considered as taking up all possible forms—that is, a polygon of s sides will have had every possible set of s atom sites for its vertices.[12] As a given polygon changes, of course, the others must change too, for no atom may be a member of more than one polygon in any given permutation P. This presents an enormously complicated mathematical problem.

We shall try to simplify it by an assumption that the various geometrical forms that a given polygon can take are roughly independent of the shapes of the other polygons. That is, we shall assume that the contribution of a polygon of a given size is the average for such a polygon over all possible forms the polygons can take without restriction. That is, *we assume the average factor contributed by a given polygon does not depend on what type the other polygons are.* This assumption is probably not sufficiently accurate to give an exact description of the order of the transition.

We shall actually use, for the contribution of a polygon, the total effect it would have if it were alone. In the various integrations over many polygons, the fact that no atom may be used twice actually restricts the volume of configuration space. It is V_N [Eq. (10)]. To include this effect we will have an additional factor V_N/V^N.

Making these assumptions the total contribution to (12) of all permutations of a given type is

$$V_N V^{-N} C(n_1, n_2 \cdots) f_1^{n_1} f_2^{n_2} \cdots f_s^{n_s} \cdots, \tag{15}$$

where $C(n_1, n_2 \cdots) = N!/\prod_s n_s! s^{n_s}$ is the total number of permutations of a given type, and f_s is the contribution of a polygon of type s, for a given configuration calculated as though it were alone. We may average this over the possible configurations also. That is,

$$f_s = V \int \exp\left[-\frac{m'}{2\beta\hbar^2}(\mathbf{z}_{12}^2 + \mathbf{z}_{23}^2 + \cdots + \mathbf{z}_{s1}^2)\right] \times \rho^{(s)}(\mathbf{z}_1, \mathbf{z}_2 \cdots \mathbf{z}_s) d\mathbf{z}_2 \cdots d\mathbf{z}_s, \tag{16}$$

where $\rho^{(s)}(\mathbf{z}_1 \cdots \mathbf{z}_s)$ is the chance of finding s atoms with their centers at $\mathbf{z}_1, \mathbf{z}_2, \cdots \mathbf{z}_s$. That is,

$$\rho^{(s)}(\mathbf{z}_1 \cdots \mathbf{z}_s) = \int \rho(\mathbf{z}_1, \mathbf{z}_2, \cdots \mathbf{z}_s, \mathbf{z}_{s+1} \cdots \mathbf{z}_N) d\mathbf{z}_{s+1} \cdots d\mathbf{z}_N,$$

where $\rho(\mathbf{z}^N)$ is the configuration density of (7). The factor V comes in (16) from the fact that $\mathbf{z}_1$ can be anywhere and has been integrated out.

[12] Actually the only contributions (near 2°K) come from polygons formed from nearly adjacent atoms. The factor (13) is very small if any of the sides are very long. The polygons of importance may be of any total perimeter (for large s), of course. It is only their individual sides which are limited.

Next we must sum over all permutation types. That is, over all values of $n_1, n_2, \cdots$ subject to (14). Signifying this by $\sum'$ and substituting into (7) we find

$$Q = (K_\beta V_N / V^N)(m'/2\pi\beta\hbar^2)^{3N/2} \sum' \prod_s (f_s^{n_s}/n_s! s^{n_s}).$$

The factor $K_\beta V_N/V^N$ presumably varies exponentially with N. We could write it as $\exp(N\alpha)$ where α is independent of N and varies slowly with temperature, vanishing as $T \to 0$. It will make no essential difference in our study of the transition (it just adds $NkT\alpha$ to the free energy A) so we will not bother to carry it along.

The sum is very difficult because of the restriction (14). However, we may use the usual methods of steepest descents. We multiply Q by a factor of the form $\exp(\mu N/kT)$ (μ is the chemical potential) and sum over N. If we then put $N = \sum_s s n_s$, we can sum on all n_s without restriction. Further, if the free energy A is

$$A = -kT \ln Q$$

and the sum is written $\exp(-B/kT)$, we can determine A from

$$A = B + \mu N \tag{18}$$

and

$$\bar{N} = -\partial B/\partial \mu \tag{19}$$

in the usual way. ($\bar{N}$ is the mean number of atoms.) Hence, putting

$$x = (m'/2\pi\beta\hbar^2)^{\frac{1}{2}} \exp(\mu/kT), \tag{20}$$

we can write

$$\exp(-B/kT) = \sum_{\text{all } n_s\text{'s}} x^N \prod_s (f_s^{n_s}/n_s! s^{n_s}) = \prod_s \sum_{n_s} (f_s^{n_s} x^{s n_s}/n_s! s^{n_s}) = \prod_s \exp(f_s x^s/s),$$

or

$$-B = kT \sum_s f_s x^s/s; \tag{21}$$

and (19) gives

$$N = \sum_s f_s x^s. \tag{22}$$

This pair of equations together with (20), (18) determines A and thereby all the thermodynamic functions. The x is determined from the second Eq. (22), by the condition that $\bar{N}$ equal N, the actual number of atoms.

To proceed further we shall have to evaluate f_s. This we do approximately, for the calculation f_s from (16) is difficult. The distribution $\rho^{(s)}$ does not permit atoms to be too close together. This is important for atoms adjacent in the polygon, such as 1 and 2. On the other hand, it is not of great geometrical importance for links much further apart (like 1 and 5). The important polygons correspond to random walks of s steps from each atom to a neighbor, finally returning to the origin. In three dimensions the chance, after a few steps, of coming back to the origin before the final step is not large. In averaging (16) over the polygons, if we include self-crossing polygons in the average we may not be

far off. There are not many of them so they probably do not alter the average very much. This is a second assumption. It is similar to the first. To be more explicit we shall approximate $\rho^{(s)}$ in (16) by

$$\rho^{(s)}(\mathbf{z}_1, \mathbf{z}_2, \cdots \mathbf{z}_s)=p(\mathbf{z}_{12})p(\mathbf{z}_{23})\cdots p(\mathbf{z}_{s1}), \quad (23)$$

where $p(\mathbf{z}_{12})$ is the probability per unit atomic volume, of finding an atom located at $\mathbf{z}_2$ if one is known to be at $\mathbf{z}_1$ (and $\mathbf{z}_{12}=\mathbf{z}_1-\mathbf{z}_2$). It is a function only of the radial distance $p(r)$, $r^2=\mathbf{z}_{12}^2$, which approaches unity as r gets beyond a few times the atomic spacing d [$p(\mathbf{z}_{12})$ is proportional to $\rho^{(2)}(\mathbf{z}_1, \mathbf{z}_2)$.]

Thus, approximately,

$$f_s=V\int \exp[-(m'/2\beta\hbar^2)(\mathbf{z}_{12}^2+\mathbf{z}_{23}^2+\cdots+\mathbf{z}_{s1}^2)] \times p(\mathbf{z}_{12})p(\mathbf{z}_{23})\cdots p(\mathbf{z}_{s1})d\mathbf{z}_2\cdots d\mathbf{z}_s. \quad (24)$$

This formula is wrong for $s=1$, for, of course, $f_1=V$. For $s=2$, (24) is not very good, for (24) averages with weight $p(\mathbf{z}_{12})^2$ while the correct weight in (16) should be $p(\mathbf{z}_{12})$. Short rings are not important in determining the existence nor order of the transition, however.

The expression (24) is nearly in the form of a convolution and can therefore easily be simplified. If the last point of the polygon were not 1 but some other location, say $\mathbf{z}_0$ [i.e., replace $\mathbf{z}_{s1}$ by $\mathbf{z}_{s0}$ in (24)], the expression ($\div V$) would depend on $\mathbf{z}_1-\mathbf{z}_0$ or $\mathbf{z}_{10}$. Call it $g_s(\mathbf{z}_{10})$. Its Fourier transform, $\int g_s(\mathbf{z}_{10})\exp(i\mathbf{K}\cdot\mathbf{z}_{10}) \times d\mathbf{z}_{10}$, is the s power of the Fourier transform,

$$\Gamma(\mathbf{K})=\int \exp(-m'\mathbf{z}^2/2\beta\hbar^2)p(\mathbf{z})\exp(i\mathbf{K}\cdot\mathbf{z})d\mathbf{z}. \quad (25)$$

Therefore $g_s(\mathbf{z})=\int \exp(-i\mathbf{K}\cdot\mathbf{z})(\Gamma(\mathbf{K}))^s d\mathbf{K}(2\pi)^{-3}$, and since $f_s=g_s(0)$, we find

$$f_s=V\int \Gamma(\mathbf{K})^s d\mathbf{K}(2\pi)^{-3}. \quad (26)$$

This is not true for $s=1$. For $s=1$ the true f_1 is V, while this gives $f_1=V\int\Gamma(\mathbf{K})d\mathbf{K}(2\pi)^{-3}=Vp(0)$ from (25). This $p(0)$ should be practically 0, but for generality we retain it. Substitution of this into (21) we get ($f_1=V$)

$$-B=kTV\sum_{s=2}\int (x^s\Gamma(\mathbf{K})^s/s)d\mathbf{K}(2\pi)^{-3}+kTVx,$$

or

$$-B=kTV\int \ln(1-x\Gamma(\mathbf{K}))d\mathbf{K}(2\pi)^{-3} +kTV[1-p(0)]x; \quad (27)$$

and, similarly,

$$\bar{N}/V=\int[1-x\Gamma(\mathbf{K})]^{-1}d\mathbf{K}(2\pi)^{-3}+[1-p(0)]x. \quad (28)$$

To study the transition more closely, we study the effects of the longer cycles. We need f_s for very large s. Since $\Gamma(\mathbf{K})$ in (25) falls as $\mathbf{K}$ rises and is maximum for $\mathbf{K}=0$, we expand $\ln\Gamma(\mathbf{K})$ in powers of $\mathbf{K}^2$ and carry only the first two terms in the form

$$\ln\Gamma(\mathbf{K})\sim\ln\delta-\tfrac{1}{2}w^2K^2. \quad (29)$$

Here

$$\delta=\int \exp(-m'r^2/2\beta\hbar^2)p(r)4\pi r^2dr, \quad (30)$$

and

$$w^2\delta=\frac{1}{3}\int r^2\exp(-m'r^2/2\beta\hbar^2)p(r)4\pi r^2dr. \quad (31)$$

Then for large s, asymptotically,

$$f_s\sim V\delta^s\int \exp(-\tfrac{1}{2}sw^2\mathbf{K}^2)d\mathbf{K}(2\pi)^{-3}=V\Delta\delta^s/s^{\frac{3}{2}}, \quad (32)$$

with $\Delta=(2\pi w^2)^{-\frac{3}{2}}$. If we use this asymptotic form for all $s>1$, we make errors for the first few terms. But the transition occurs because of the character of the convergence of the series for large s. Therefore the predicted character of the transition may be found by studying the sums (21), (22) with the asymptotic form (32) for f_s. For example, the $\bar{N}/V$ sum is [putting $p(0)=0$]

$$\bar{N}/V=\Delta\sum_{s=1}(\delta x)^s/s^{\frac{3}{2}}+(1-\Delta\delta)x, \quad (33)$$

and the expression for B is

$$-B=kTV[\Delta\sum_{s=1}(\delta x)^s/s^{5/2}+(1-\Delta\delta)x]. \quad (34)$$

The same situation exists here as for the ideal gas case. The sum in (33) cannot exceed 2.612 (for $x=1/\delta$) and $2.612\Delta+(1-\Delta\delta)/\delta$ may be less than the actual N/V desired. This will not occur for high temperature (δ small), but on lowering T the difficulty suddenly sets in.

What one must do, as is well known,[4] is to note that the $d\mathbf{K}$ integral in (32) is really a sum over the values of $\mathbf{K}$ which fit in the box of volume V. The lowest state ($\mathbf{K}=0$, using running waves) is distant $d\mathbf{K}=(2\pi)^3/V$ from the next. It suffices to sum on this one and integrate the others. Thus we should add a factor $1+(2\pi)^3V^{-1}\delta(\mathbf{K})$ to the integrand of (32),[13] so that the sum is more closely

$$f_s=V\Delta\delta^s s^{-\frac{3}{2}}+\delta^s, \quad (32a)$$

and (33) becomes

$$\bar{N}/V=\Delta\sum_{s=1}(\delta x)^s s^{-\frac{3}{2}}+(1-\Delta\delta)x+V^{-1}(1-\delta x)^{-1}. \quad (33a)$$

Now δx can become very nearly 1, to order $1/V$ (e.g., put $\delta x=1-1/gV$) and the sum in this region is $\bar{N}/V=2.612\Delta+(1-\Delta\delta)\delta^{-1}+g$ which can be satisfied for proper choice of g. For higher temperatures (smaller δ) we can use the original expansion (33). This change in

[13] The validity of this procedure has sometimes been questioned. A method of arriving at the result (32′) which avoids the use of this procedure is given in the Appendix.

behavior for the two regions of T reflects in B (34) as a phase transition. This shows the existence of the transition. As temperature falls δ in (30) rises without limit. Since $p(r)\sim 1$ for large r, as β rises δ eventually behaves as $(2\pi\beta\hbar^2/m')^{\frac{3}{2}}$. Likewise, eventually w^2 becomes $\beta\hbar^2/m'$, so that $\Delta=(2\pi w^2)^{-\frac{3}{2}}$ approaches δ^{-1}. Therefore $2.612\Delta+(1-\Delta\delta)\delta^{-1}$ tends toward zero (as $T^{\frac{3}{2}}$) and must eventually fall below N/V. Actually by putting in very reasonable values for the parameters, it is easy to obtain a transition at about the right place.

If we do not use the asymptotic form (29) for Γ, nothing is fundamentally changed. In (28) the integral on $\mathbf{K}$ should have its factor $1+(2\pi)^3V^{-1}\delta(\mathbf{K})$ as explained, and the equation to determine x becomes (calling $\Gamma(0)=\delta$)

$$\bar{N}/V=\int[1-x\Gamma(\mathbf{K})]^{-1}d\mathbf{K}(2\pi)^{-3}+[1-p(0)]x+V^{-1}(1-\delta x)^{-1}. \quad (36)$$

Above the transition the last term is not required. Below, $\delta x=1-1/gV$ and $\bar{N}/V=\int[1-\delta^{-1}\Gamma(\mathbf{K})]^{-1}d\mathbf{K}\times(2\pi)^{-3}+[1-p(0)]\delta^{-1}+g$, with analogous expressions for B.

RELATION TO EXPERIMENT

It would not be worth while to substitute numbers in these expressions as too many small approximations have been made. In addition, it is difficult to estimate m'. The function $p(r)$ might be calculated roughly for the smaller r by using the corresponding function required for the quantum-mechanical second virial coefficient. This assumed that any two colliding atoms are independent of the others. Alternatively, $p(r)$ could be taken experimentally from x-ray or neutron scattering data.

On the other hand, the formulas (27), (28) have even qualitative faults when compared to experiment. They predict that helium, like the ideal Bose gas, would show a third-order transition (specific heat continuous but discontinuous slope). The experimental data[1] do not agree (apparently the specific heat is discontinuous). This disagreement probably stems from the neglected geometrical correlations among the rings.[14]

In order to study this in greater detail, it was thought that a careful study of the situation at extremely low temperature would be of value. The character of the transition must depend on an accurate description of the phase into which the liquid changes as it cools past the λ point. This phase is represented in an extreme form near absolute zero. In this region Eqs. (27), (28) fail very badly. They predict a specific heat varying as $T^{\frac{3}{2}}$ while experimentally it varies as T^3. In the following paper[6] we shall see that this discrepancy is a result of the error produced by the geometrical approximations made in passing from (7) to (27), (28). The approximations here permit much larger fluctuations in density than is available to the true liquid, and this qualitatively alters the behavior of the specific heat at low temperature.

The experimental specific heat curve shows[1] a slight rise in the He I region as the λ point is approached from above. This is also a property of our expression (34). In this region only a few chains are starting to form. The restriction that no atom may be in more than one chain is not yet of importance. Therefore, in this region our geometrical approximations should be valid. The very least we can say, then, is that the rapid rise in specific heat of He I with falling temperature is completely explained.

SUMMARY

Starting with an exact quantum-mechanical partition function, we have derived an approximate expression [Eq. (7)] which should be qualitatively accurate. It has been shown to be in agreement with experiment in predicting a transition which depends in an essential manner on the statistics.

Further mathematical approximations have not been accurate enough to show whether (7) will correctly predict the order of the transition and the temperature dependence of the specific heat near absolute zero. They do suffice at high temperatures to show the rise in specific heat of He I as the transition is approached.

It is proposed that a more careful analysis of (7) would show more complete agreement with the experimental facts.[1] In the next paper[6] the situation near absolute zero is studied in detail, and it is found that (7) (corrected for the effect mentioned in footnote 11) very likely does predict the correct behavior in this region.

The physical idea which plays a central role is that in a quantum-mechanical Bose liquid the atoms behave in some respects like free particles.

The author appreciates conversations with Edward Kerner and with M. Kac, as a result of which he became interested in the problem. He also is grateful for discussions with E. Wigner, H. Bethe, and R. F. Christy.

APPENDIX

We give here another derivation of the approximate partition function (33a), (34). It has the advantage of showing more clearly the origin of the transition. We will treat it in a very approximate manner as we have already given more complete formulas. (It is the first derivation that the author made.)

Near the transition only permutations involving shifts of each atom to the position of its neighbor, at some mean distance d, are important. The exponential factor from such a shift is $y=\exp(-m'd^2/2\beta\hbar^2)$.

Each permutation can be broken into cycles. We now only count those cycles for which all atoms are adjacent,

[14] Assumptions about the temperature variation of the parameters m, $p(r)$, K_β cannot alter the order of the predicted transition. The effect of modifying (7) in the manner indicated in footnote 11 is discussed in the paper to follow. The change, if anything, is in the wrong direction.

forming a closed chain or ring. If a ring contains s atoms, its contribution is y^s. If we have n_2 rings of 2 atoms, n_3 of 3 atoms, etc., the contribution is $y^{2n_2+3n_3+\cdots}$. The part of the partition function which determines the transition is then

$$q=\sum G(n_2, n_3, \cdots)y^{\sum_s (s-2)n_s}, \tag{1-A}$$

where $G(n_2, n_3, \cdots)$ is the number of ways we can lay out polygons, n_2 of 2 sides, n_3 triangles, etc., on the configuration—the sides of the polygons consisting of lines joining nearest neighbors (length d). The sum is restricted for the number of single atoms $n_1=N-\sum_{s=2} sn_s$ must not be negative.

Here again we shall make the error of neglecting the geometrical interference of polygons due to the restriction that each atom be a vertex of only one polygon. We shall include the competition among the polygons for the total number of available atoms by saying that there is an average probability t that any site is unoccupied. This t will later be determined so that the average number of atoms occupied is $\bar{N}=N$. (This can be done in detail by steepest descents, but it amounts to the same thing.) Therefore, instead of q we calculate $t^N q$ and call it $\exp(-B/kT)$.

The polygons can now be considered as independent. If R_s is the total number of s gons that can be drawn on the configuration, each s gon can be chosen in R_s ways, and all n_s of them in $R_s^{n_s}/n_s!$ ways. The n_1 single atoms can be chosen in $N^{n_1}/n_1!$ ways. Thus

$$\exp(-B/kT)=\sum_{n_1\, n_2 \cdots} \prod_{s=2} R_s^{n_s}(n_s!)^{-1}y^{sn_s}t^{sn_s}t^{n_1}N^{n_1}(n_1!)^{-1}$$

$$=\exp[Nt+\sum_{s=2} R_s(ty)^s],$$

or

$$-B=kT[Nt+\sum_{s=2} R_s(ty)^s]. \tag{2-A}$$

The average number of atoms $\bar{N}$ used in sites is $t(\partial q/\partial t)q^{-1}$, so we have

$$\bar{N}=Nt+\sum_{s=2} sR_s(ty)^s. \tag{3-A}$$

Now we calculate R_s. We call h_s the total number of ways that we can, starting at an atom and making successive steps to adjacent atoms, return after exactly s steps. This forms a closed s gon. We may start at any of the N atoms. The total Nh_s measure the total number of s gons, but counts each s times (for you could start at any of the s atoms as the "first"). Hence we have, $R_s=Nh_s/s$, and

$$-B=kTN[t+\sum_{s=2} h_s(ty)^s/s], \tag{4-A}$$

where t is a parameter determined from $\bar{N}=N$, in (3-A), or

$$1=t+\sum_{s=2} h_s(ty)^s. \tag{5-A}$$

We shall bother to determine the form of h_s for large s only. In the random walk, each step is length d, and may be made to any of the adjacent atoms, which we shall say are l in number, on the average. Each step can be made in l ways, the entire s steps in l^s ways. There are l^s walks in total but only a certain number return to the origin. As is well known, the probability of being at a given point, per unit volume, radius r from the starting point is

$$(2\pi d^2 s/3)^{-\frac{3}{2}}\exp(-3r^2/2sd^2)=\Delta s^{-\frac{3}{2}}\exp(-r^2/2sw^2),$$

putting $w^2=d^2/3$ and $\Delta=(2\pi w^2)^{-\frac{3}{2}}$. The chance we are back at the original atom (that is, within a space of one atomic volume $V_A=V/N$ near the origin) is $V_A\Delta s^{-\frac{3}{2}}$, so that

$$h_s=V_A\Delta s^{-\frac{3}{2}}l^s.$$

The reason for the dependence on the $s^{-\frac{3}{2}}$ is easy to see. After s steps we have wandered out to a mean radius of order $s^{\frac{1}{2}}$, or over a volume $s^{\frac{3}{2}}$. Hence, the chance that in this volume we are back at the original atom varies as $s^{-\frac{3}{2}}$. This is correct except for enormous chains $s\gtrsim N^{\frac{2}{3}}$, which are long enough to wander all over the liquid. Then the available volume no longer increases. The chance that one is back at the original atom instead of one of the other roughly equally likely atoms is $1/N$. For such large s, then, $h_s=l^s/N$. The formula,

$$h_s=(V_A\Delta s^{-\frac{3}{2}}+1/N)l^s, \tag{6-A}$$

takes care of both cases, because for small s the first term dominates, and for large s the second takes over, as it should.[15] Substitution into (3-A) gives

$$1=t+V_A\Delta\sum_{s=2}(lyt)^s s^{-\frac{3}{2}}+N^{-1}(1-lyt)^{-1}. \tag{7-A}$$

This may be compared to (33a). To do so, note that our interpretation of $3w^2$ as the mean square length of a step agrees with (31). Further ly is the number of atoms available per step, l, multiplied by the weight $y=\exp(-m'd^2/2\beta\hbar^2)$. In the more general case it becomes δ/V_A as an inspection of (30), the expression for δ, shows. Finally, the parameter t can as well be called V_Ax, and Eq. (7-A) is seen to be identical to (33a) (times V_A). Likewise, substitution of (6-A) into (4-A) gives the corresponding equation (34) for B.

This derivation throws some light on the mathe-

[15] According to (6-A) the chance to return is higher than for an infinite medium. It might be objected that there should be fewer paths available when there are a finite number of atoms in an enclosed space. What we have done corresponds to working with a periodic boundary condition, and the excess arises from the chance to return to one of the images of the origin, instead of to the origin itself. With the more physical boundary condition—that paths cannot cross the liquid surface—the total number of paths for high s is not l^s, but is reduced. It becomes eventually proportional to $e^{-\epsilon s}l^s$, where ϵ is a very small number of order N^{-1}. It makes no essential difference in the result. In the momentum representation of the text it corresponds to taking the lowest state to have a wavelength controlled by the size of the box. I am indebted to Herman Kahn for pointing out this possible objection to (6-A).

matical "cause" of the transition. As T falls y rises, until the enormous number of possible orientations of a very long ring more than compensates for the small contribution of each ($y^l \ll 1$, for l large).

There is no doubt of the geometrical fact of large numbers of orientations for long rings, even if these rings may never use the same atom twice (i.e., cannot cross themselves).[16] Therefore there can be no doubt that (1–A), and its more complete expression (17), will show a transition from this cause. But the order of the transition need not be the same as that of the approximate evaluations we have made. They neglect the geometrical correlations. For example, if a large chain of K atoms is already formed, are the remaining $N-K$ atoms more (or perhaps less) likely to be contiguous and therefore more easily able to make other chains, than if these $N-K$ atoms were chosen at random from among the N? Our assumption in deriving (5–A) was that it was equally likely either way.

[16] For example, the number of ways in which a single polygon which does not cross itself can still be oriented in an infinite medium = constant $\times l^s s^{-\frac{1}{2}}$, but the value of l is reduced.

PHYSICAL REVIEW VOLUME 91, NUMBER 6 SEPTEMBER 15, 1953

Atomic Theory of Liquid Helium Near Absolute Zero

R. P. FEYNMAN
California Institute of Technology, Pasadena, California

(Received June 1, 1953)

The properties of liquid helium at very low temperatures (below 0.5°K) are discussed from the atomic point of view. It is argued that the lowest states are compressional waves (phonons). Long-range motions which leave density unaltered (stirrings) are impossible for Bose statistics since they simply permute the atoms. Motions on an atomic scale are possible, but require a minimum energy of excitation. Therefore at low temperature the specific heat varies as T^3 and the flow resistance of the fluid is small. The arguments are entirely qualitative—no calculation of the energy of excitation nor of the low-temperature viscosity is given. In an appendix an expression, previously given, for the partition function is modified to include the effects of phonons.

INTRODUCTION

TISZA[1] has suggested the very fruitful concept that He II might be thought of as a mixture of two fluids, "superfluid" and "normal." At zero temperature the helium is pure superfluid. With rising temperature some sort of "excited molecules" form. These constitute the "normal fluid" which behaves very much like a gas. The proportion of normal fluid increases at first slowly, and then rapidly, with temperature until at the transition temperature of 2.19°K (λ point) the liquid, now He I, contains no more superfluid.

Landau[2] has made even more detailed suggestions. He suggests that there are two kinds of "excited molecules," phonons or quanta of longitudinal compressional waves (sound) and "rotons." The latter are not well understood. It is suggested that they have a minimum energy Δ needed to excite them. For this reason below 0.5°K there are practically only phonons. The rotons can become excited when more energy is available; i.e., at higher temperature. This idea is in agreement with the fact that below 0.5°K the specific heat varies as T^3 in just the manner (and with the correct coefficient) to be expected if only longitudinal sound waves could be excited.

Tisza's view is frankly phenomenological. No serious attempt is made to justify the description from first principles. Landau has made such an attempt by studying the quantum mechanics of a continuous liquid medium. The role of the statistics is not clear in his arguments, however. Furthermore, the magnitudes of energy and inertia that the "rotons" appear to have correspond to a few atoms. A complete understanding of the "roton" state can therefore only be achieved by way of an atomic viewpoint.

A more complete study of liquid helium from first principles might attempt to answer at least three important questions:

(a) Why does the liquid make a transition between two forms, He I and He II?

(b) Why are there no states of very low energy, other than phonons, which can be excited in helium II (i.e., below 0.5°K)?

(c) What is the nature of the excitations which constitute the "normal fluid component" at higher temperatures, say from 1 to 2.2°K?

The first question was answered in a preceding paper.[3] We showed that London's suggestion, that it is the analog of the transition in an ideal Bose gas, is correct.

In this note we hope to make a qualitative argument from first principles to answer the second question.

[1] L. Tisza, Phys. Rev. **72**, 838 (1947). An excellent summary of the theories of helium II is to be found in R. B. Dingle, Supplement to Phil. Mag. **1**, 112 (1952).

[2] L. Landau, J. Phys. U.S.S.R. **5**, 71 (1941).

[3] R. P. Feynman, Phys. Rev. **91**, 1291 (1953), hereafter called I.

(We have not yet found the answer to the third.*) It is to be understood, therefore, that we are aiming here to explain the properties of the liquid only at extremely low temperatures.

We take, as a model, helium atoms obeying the Schrödinger equation and the symmetrical Einstein-Bose statistics with forces between pairs similar to that worked out by Slater and Kirkwood,[4] an attraction at large distances and a strong repulsion at small.[5] (It is sufficient for qualitative purposes, if one wishes, to imagine impenetrable spheres of radius about 2.7A packed into a space so the mean spacing (cube root of atomic volume) is about 3.6A.

At absolute zero the system is in the ground state. Why this is not a solid has been explained by London.[6] The large zero-point motions of the atoms are capable of "melting" any ordered crystalline arrangement that may be temporarily set up. We begin by describing the wave function for this ground state.

DESCRIPTION OF THE GROUND-STATE WAVE FUNCTION

Wave functions can be described qualitatively in words by giving the amplitude for every configuration of the atoms. The ground state has a positive amplitude for any configuration since the lowest state has no nodes. The amplitude is negligible if any two atoms are so close together that they overlap—that is, that a large negative potential has set in. Thus for any atom surrounded by neighbors, considered for a moment fixed, the amplitude falls to zero when the atom moves over to touch any of the neighbors and is probably bound in such a way that it is maximum when the atom is near the center of its "cage." (This curvature makes a strong kinetic energy tending to blow apart the cage, an energy effect canceled by the long-range attraction of the atoms.) Compressing the atom into a smaller cage requires more kinetic energy, and expansion works against the attractive potential so there is some mean density of equilibrium. Fluctuations away from that mean density are of amplitude distributed in a Gaussian manner. For wavelengths exceeding the atomic spacing they are analogous to zero-point fluctuations of the vacuum electromagnetic field (but are wholly longitudinal, scalar waves, of course).

There are a large number of configurations with densities near the most likely density, in all of which the atoms tend to keep separate from one another. They differ from one another mainly in the location of the atoms. The various configurations differ only in that one may be "stirred" into another with a little reshuffling or stirring of the atoms. We may take it that all configurations which have nearly the same type of density fluctuations and which can be essentially just stirred from one to another have the same amplitude in the ground state.

THE CHARACTER OF LOW-ENERGY STATES

We must next determine the character of the low-energy states near the ground state. We aim to show that the only states which differ from the ground state by an infinitesimal energy are the phonons.

First we can take it that the lowest states are those involving large numbers of atoms or large distances. Consider, for example, a tiny region of the liquid, say a cube 3 or 4 atomic spacings on a side. If the atoms in this region are confined in this region (so we have a submicroscopic sample of liquid He) the excitations above the ground state will all involve one node somewhere among the configurations and hence a wavelength of order of a. This must mean an excess kinetic energy of the order $\hbar^2/a^2m$, or at least of order $\hbar^2/Na^2m$, where N is the number of atoms and m is the mass of each—leaving an appreciable gap from the ground state.

This argument fails if in the ground state there are two (or more) regions of configuration space in both of which the amplitude is large and which are completely separated by a region of very small amplitude. (Analogous to a particle in a potential with two wells separated by a barrier.) If the nodal surface is passed through the region of small amplitude (the barrier) very little change in energy results. But we have seen that the states of large amplitude are just all those in which the atoms are reasonably well separated. We can assume that we can get from one to any other without crossing any high potential barrier. We suppose all possible rearrangements may be achieved without the atoms coming too close together at any time. That this is reasonable can be seen by comparing the size of the atoms to their spacing. For example, if at some point they are locally roughly on a cubic close-packed lattice, the nearest neighbors are 4.0A apart (corresponding to the observed atomic volume[5] of 45A³). The diameter of the atoms is 2.7A, the radius at which Slater and Kirkwood's potential passes from minus to plus. The cube edge of the lattice is 5.6A, so a face-centered atom could even pass between those at the corners of the cube! Clearly, if they are allowed to vary their mutual distances a little, all kinds of rearrangements can be made. It is likely that the condition that there be no effective barrier between configurations is equivalent to the condition that the He II is liquid in the lowest state. We assume it valid for He II.

If we are to find extremely low-energy states we must therefore look to excitations involving large groups of atoms or long wavelengths. One possibility is in the compression waves. Suppose the atoms are compressed to a small excess density over a large volume and a rarefaction left adjacently. The only way this fluctuation could even out is for a considerable number of atoms to move, each a little bit. This involves, effectively, the motion of a large mass and can have low kinetic energy. There is, therefore, little doubt that such compressional waves represent a true mode of excitation in the helium, and a mode of very low energy. If the speed of sound is called c and the wave number of the waves K, the frequency $\omega=cK$ and quantum

* *Note added in proof:*—This problem has now been solved. Its solution will appear in a forthcoming publication.

[4] J. C. Slater and J. G. Kirkwood, Phys. Rev. **37**, 682 (1931).

[5] For a detailed account of the properties of helium see W. H. Keeson, *Helium* (Elsevier Publishing Company, Inc., Amsterdam, 1942).

[6] F. London, Nature **141**, 643 (1938).

energy $\hbar cK$ can be as small as desired. Thus there is no reason why we should not expect the specific heat, varying as T^3, from these modes.[7]

We have assumed that the way to release a density fluctuation in a given time which requires the least kinetic energy is to move many atoms a short distance. If only a fraction f of the atoms move, the velocity required is $1/f$ times higher. This means a higher kinetic energy if the kinetic energy varies as the velocity squared.

There are cases, however, in which the kinetic energy does not vary in this way. In a degenerate Fermi gas a single atom excited by a small excess momentum p above the Fermi surface of momentum p_0 has an energy $((p_0+p)^2-p_0^2)/2m$ or $(p_0/m)p$. This linear dependence of energy on momentum means that the group velocity of such waves is energy independent. It is p_0/m. The speed of sound calculated from the compressibility and density of the ideal gas is $3^{-\frac{1}{2}}p_0/m$. The single atoms will run ahead of the sound. Fluctuations are reduced by a process more like diffusion than sound. The specific heat near $T=0$ varies as T instead of T^3 because the density of states for exciting single atoms exceeds that for sound. As we shall see, for Bose helium there are no states, except phonons, whose energy approaches zero as their momentum approaches zero. The sound has no competitor capable of discharging pressure.[8]

LOW-ENERGY STATES DISREGARDING STATISTICS

The Bose statistics play an essential part in the discussion of other possible states of very low-excitation energy. To make this role clear by contrast, we shall first analyze the situation, disregarding the statistics. More precisely we consider in this section an imaginary quantum liquid made of atoms which are, in principle, distinguishable ("Bolzman" statistics). The ground state for Bolzman statistics is the same as for Bose statistics, since the lowest state is, in either case, symmetrical.

We must try to find modes of excitation involving long wavelengths which do not involve changes in mean density. The density fluctuation modes have already been considered. To simplify the argument consider first the following crude model. We consider a set of cells, each of which contains one atom. Each atom is free to wander in its cell and may occupy therefore some ground-state wave function, say constant amplitude in the cell. We are to consider states which can be made solely by rearranging the atoms among the cells. No two atoms may go into the same cell, for that corresponds to a density fluctuation, and we do not wish to consider those.

In the ground state, the amplitude is the same for any rearrangement of the atoms among the cells. The lowest states can now be analyzed as follows. We neglect the statistics, that is, we assume Boltzmann statistics to apply. We can describe a wave function which corresponds to a very low excitation as follows: Pick out a certain atom, A, say. Put it in a given cell. Then all rearrangements of the other atoms among the other cells can be taken to have the same amplitude. If atom A is in a different cell, the amplitude may be different, but again independent of the arrangement of the other atoms. We thus can specify this wave function by giving just the amplitude for various positions of atom A. The function is independent of the position of the others. We may take this as $\exp(iK\cdot R_A)$, where R_A is the position of the center of the cell in which atom A is. We may, with enough accuracy, let R_A be just the position of atom A. The K can be a long wave fitting into the volume V in which the helium is contained. The energy of this state is $\hbar^2K^2/2m'$ where m' is the effective mass needed to move atom A. This energy can be very small, for K can be small.

This effective mass is not far from the mass of one helium atom. It is discussed in a previous paper.[3] We summarize the argument here. To push a single atom along, we need not go over any potential barriers. The other atoms may move out of the way. No matter where atom A is located the other atoms can arrange themselves into a state of minimum energy and this minimum energy is independent of the location of atom A. However, as A moves, the others must readjust themselves into the state of minimum energy for the new position of A. That is, in addition to the kinetic energy of A there is a kinetic energy of the other atoms which must move away to make room for A.

Thus there would be low-lying states, of energy $\hbar K^2/2m'$. The number of such states would be very large, for the wave function could depend in similar ways on the coordinates of other atoms also. [For example, we could choose two atoms A, B and have the wave function vary with their location as $\exp(iK_1\cdot R_A)\exp(iK_2\cdot R_B)$ and be otherwise independent of the distribution of the others in the cells, etc.] The large density of low-lying states would result in a large specific heat near absolute zero.

LOW STATES WITH SYMMETRICAL STATISTICS

However, if the atoms obey Bose statistics none of these states can exist. For in our model, whether atom A is at one location or another is merely an interchange of which atom is which. This cannot change the wave function. In fact, for the model of one atom in a cell no excited state at all can exist for Bose particles without excitation of the atom within the cell.

For the real liquid a similar situation holds—aside from the phonons, there can be no low-lying state. The wave function must have the property that any change, that just means an interchange among the atoms, must not alter the wave function. The excited state must be orthogonal to the ground state, of course. Starting at any configuration and supposing the amplitude is

[7] The partition function discussed previously (reference 3) is extended to include a description of the phonons in the Appendix to this paper.

[8] For an *ideal* Bose gas, as $T\to 0$, the sound velocity approaches zero, so that the expected T^3 specific heat does not appear. It is replaced by $T^{\frac{3}{2}}$. Density fluctuations are much more restricted in liquid helium than they are in the ideal gas. This is the origin of many differences in behavior for the two cases.

the same as for the ground state, we must find a new configuration that represents a kind of stirring of the old configuration (to omit phonon states) such that the amplitude is now reversed in sign. It is clear that *every* configuration is close to the original one, albeit with some atoms interchanged. So it is hard to find a configuration to give the minus amplitude which is sufficiently far (in configuration space) from the original positive amplitude configuration to have a slow rate of decay and thus a low energy.

There are some possibilities of this kind which might at first sight seem allowable. Consider among the atoms a set of adjacent ones forming a large ring. Suppose l atoms are in the ring and let us imagine the wave function is such that, if all move together half an atomic spacing a, the phase changes by π, so when they turn about another $a/2$ the phase changes by 2π, as required by the Bose statistics. (For a shift of a just changes each atom for the one behind.) The mass moving is $m'l$, and the momentum is $\hbar/a$ (for the wavelength is a) so that the energy is $\frac{1}{2}(m'l)^{-1}(\hbar/a)^2$. This may be made low by choosing l very large. (If it were not for the Bose statistics we could have taken the wavelength to be la and the energy would be even lower, varying as $1/l^3$.)

The argument for calculating this energy is incorrect, however. By assuming that all the atoms must move together, degrees of freedom (in which parts of the ring turn by themselves) have been restricted. This tacitly adds a considerable energy by the uncertainty principle.

To understand better the failure of the argument, consider by the same reasoning the case that one had two equal rings parallel to each other. We may argue as before that now the entire mass $2m'l$ moves together with momentum $\hbar/a$ and thus expect an energy $\frac{1}{2}(2m'l)^{-1}(\hbar/a)^2$ for the energy of the lowest state. But this is certainly wrong, for if we consider the rings as independent, the lowest state but one is surely that one for which one of the rings is excited and the other not (momentum O), an energy $\frac{1}{2}(m'l)^{-1}(\hbar/a)^2$, larger than our previous lowest estimate! The error for the first figure of $\frac{1}{2}(2m'l)^{-1}(\hbar/a)^2$ consisted in this. In describing the wave function, the possibility that the two rings could turn independently was omitted. To force the rings to move together would be to force the difference between their displacements to be fixed at zero. Thus the momentum conjugate to this difference coordinate would be very high, and the energy associated with this coordinate very large. Thus the system does not have just the energy estimated but, in fact, a very much higher one. We have not completely specified the wave function. We have not said what the amplitude is to be for configurations in which only one of the rings moves.

In an analogous way, the estimate of energy for the single ring, $\frac{1}{2}(m'l)^{-1}(\hbar/a)^2$, is incorrect. We have specified the amplitude for the case that *all* of the atoms are simultaneously in the mid positions. What is to be the amplitude if only a few move to mid positions? (This may be accomplished without doing violence to the potentials by using the atoms adjacent to the ring and by turning on smaller rings of 4 or 5 atoms.) If the motion (all to $\frac{1}{2}$ position) can be made up of smaller parts moving in concert, and if these smaller parts could also have moved independently, the energy cannot be lower than that corresponding to just one of independent parts being excited.

Since we may well imagine that any motion of the atoms could be made up of combinations of motions of small groups (say 3 or 4 revolving about each other) the lowest energy is the excitation of one of these small groups. These we can identify with Tisza's excited molecules or Landau's rotons (but see next paragraph). Any such small ring of r atoms must, of course, have its first state of excitation of angular momentum $r\hbar$ (or $p=\hbar/a$) because of the statistics. Landau's arguments[2] that such angular momentum must have an excitation energy was made in a way that does not involve the statistics of the atoms. It is possibly equivalent to our argument that large rings need not be considered if their motion is analyzable in smaller parts. The central importance of the statistics would not seem to be here, but rather in the previous argument which shows that no states corresponding to the slow linear motion of a single atom are permitted.

It is not obvious whether the lowest excited state, excluding the phonons, is actually a small ring of atoms turning. The arguments do not exclude the possibility that these are all higher than another type of mode; namely, the rapid motion of a single atom. In our cell model a state which depends on atom A as $\exp iK\cdot R_A$ with $Ka=2\pi$ is, of course, possible. Another possibility is the analog of the excitation of a single atom in a cell. (This may be the same as the single atom motion.) All of these states differ from the ground state by a finite energy. But which is lowest is hard to determine.

Any such excitation can, of course, move through the fluid. (In fact, the lowest state is that in which it has equal amplitude of being anywhere in the fluid.) That is, the wave function could vary as $\exp(iK\cdot R)$, where R is the location of the center of excitation. Then the energy of these excitations might have the form, suggested by Landau,[2] $\Delta+K^2/2\mu$, where Δ is the energy needed to excite the ring or other excitation, and μ is a sort of effective mass.

Our primary purpose was to show that no states close to the ground state exist, exclusive of the phonons. We are not yet able to calculate the energy, nor to give a clear picture of the other modes of excitation.

DESCRIPTION OF SOME PROPERTIES OF THE LIQUID

In concluding that only phonons exist at low temperature, we concur with the opinion of the phenomenological theories. Therefore, a description of how some of the properties of helium arise, according to this model, will repeat much that has already been pointed out by others.[1,2] We limit ourselves, therefore, to a very brief summary from a kinetic theory point of view.

First, consider the motion of an object, such as a small sphere through the liquid. If the object is stationary at a fixed position R, the liquid may get into a certain state which, omitting phonons for a moment, is like the ground state (except that now part of the space occupied by the object is not available to the helium). Let the wave function of the helium be ψ_R. It is a function of all the helium atoms and depends, say parametrically, on R. If R is changed, ψ is also, but the energy of the fluid is not changed. For, rearranging the fluid to a new shape at the same density does not alter the energy. Now if we alter the R from R_1 to R_2 nearby, we will only need to add a little kinetic energy to push the helium atoms out of the

way. The overlap of ψ_{R_1} and ψ_{R_2} will be nearly perfect (except near the surface of the object which is a different volume of space in the two cases, $R=R_1$ and $R=R_2$). The object can move therefore with an *energy* equal to just the kinetic energy of itself and the liquid which flows around it. The fluid will move so that the curl of the velocity is zero, because circulation corresponds to permutation of atoms, and the Bose statistics will not permit such motions, as we have seen.[9]

What will be the losses of energy suffered by such an object? If it loses energy it can do so only by exciting the helium. First, can it excite the molecular excitations? These take a certain energy Δ to excite, but a massive object even moving very slowly may have sufficient energy. On the other hand, for such an object to change energy by Δ its momentum must change by an enormous amount. To create an excitation of very high momentum may take much more energy than Δ. As Landau has shown,[2] if the energy to excite a roton of momentum p is taken to be of the form $\Delta+p^2/2\mu$, the laws of conservation of energy and momentum show that slowly moving objects [velocity less than $(2\mu\Delta)^{\frac{1}{2}}$] can produce no excitations. Likewise, objects moving at velocities below that of sound cannot lose energy by creating phonons. Therefore, at absolute zero and for not too high velocity, a moving object will suffer no viscous drag.

At low temperatures, there are, however, some phonons already existing in the liquid. They can scatter off of the object (changing their energy by the Doppler effect) and in this way the object can lose a little energy. A phonon of energy $\hbar\omega$ carries momentum $\hbar\omega/c$ and behaves very much like a particle of mass $\hbar\omega/c^2$ moving at velocity c. The phonons act in most respects like a gas of such particles, and the resistance suffered by our object is just like the viscosity that would be suffered by an object moving through such a gas.

The actual calculation of this energy loss means a calculation of the viscous drag of such a gas. This requires a knowledge of the mean free path for collision among the phonons. The phonons scatter from one another because the medium is not linear. The speed of sound depends a little on the density. Therefore speaking classically, if a wave is present, another wave impinging finds the index of refraction varying sinusoidally and is thereby scattered. The same thing happens in the quantum mechanical system.

The mean free path should rise rapidly as the temperature falls, because the density of phonons decreases (the scattering cross section also decreases). Interesting phenomena should result when this free path becomes comparable to the dimensions of the apparatus.

If the viscosity is measured by connecting two vessels with a capillary, a different situation arises. The phonons cannot readily work their way through the long capillary, but the bulk liquid can move through it. There is work done on the phonons as the piston moves down in one of the vessels, again by the Doppler effect on the phonons bouncing off of the moving piston. But this work just goes into increasing the phonon energy—that is, the liquid in one vessel is heated, in the other cooled. If isothermal conditions are maintained it goes as heat to the walls of the first container and from the walls of the second. No net work is done in this case and the viscosity appears to vanish. There is essentially no energy loss because there is no real viscous flow of our phonon "gas" through the capillary. The reason that, experimentally, resistance appears[5] if the flow velocity exceeds a certain critical velocity is not clear. Perhaps in passing sharp protuberances in the capillary wall the velocity locally exceeds that needed to create excitations or new phonons. It cannot very well be a kind of turbulence because presumably the velocity field should be always free of circulation.

In two volumes of liquid helium connected by a capillary, the hotter one will exert the higher pressure (fountain effect). The larger number and higher average momentum of phonons in the hotter region results, from wall bombardment, in a higher pressure there. The pressure can only be released slowly by phonons passing through the long capillary. This would be the mechanism of heat conductivity through capillaries. The rate of such conduction would depend on the relative size of the capillaries and the phonon mean free path.

If temperature varies from one point to another in the bulk liquid, then the phonon density varies. What happens depends on the mean free path. If it is long compared to the distances over which the variations occur, the variations are almost immediately evened out by the diffusion of phonons rushing from one place to another (at the speed c). If the mean free path is shorter than the distances involved in the variation, no single photon can go directly from a high- to low-density region. Instead, a cooperative movement sets in. If we consider the analogy to a gas of phonon "particles," a pressure variation is released by body motion—that is, by sound waves. The speed of this sound is $3^{-\frac{1}{2}}$ times the individual particle velocities. In our case, this "second sound" representing waves of phonon density (i.e., temperature) should travel at a velocity $3^{-\frac{1}{2}}c$. At low temperatures it will be experimentally hard to keep the mean free path very small compared to the wave-

[9] For a liquid contained in a simply connected region, the circulation vanishes everywhere if it vanishes locally. But in a region of connectivity like the inside of a torus, although the curl is everywhere zero, the circulation around the ring may not be zero. Such a circulation cannot be compounded of smaller independent units. Therefore, it should be possible to demonstrate circulatory motions in such a vessel which will maintain themselves for a long time. Circulation may be created by rotating the vessel containing He I and cooling to the temperature desired below the λ point. Stopping the rotating vessel should leave the interior liquid with a nearly permanent angular momentum, which could be demonstrated, for example, by its gyroscopic effects. The liquid must be completely confined with no free surface because the exchange of atoms between the rotating liquid and the stationary gas above it might cause a rapid damping of the angular momentum.

length, so that second sound would show appreciable damping and dispersion. At extremely low temperatures the free path may be larger than the apparatus. Then, if a pulse of heat at one point creates extra phonons, these will rush away at speed c so that temperature rise will begin at a distant point delayed only by the time required for *first* sound to traverse the apparatus.[10]

EFFECTS OF He^3 ATOMS AT LOW CONCENTRATIONS

If a foreign atom, say an atom of He^3, is in the liquid, our arguments indicate that it will move about essentially as a free particle, albeit with an effective mass m'' larger than its true mass. (This m'' should be about one atomic mass unit less than the effective mass m' of a He^4 atom.) Such He^3 atoms put into He^4 at low concentration should behave as a perfect gas. Consider, as an example, a concentration of 0.1 percent. The mean spacing of the atoms is so large that the gas is not degenerate, except at a few hundredths of a degree. (The statistics can only be of importance if the atoms can permute. This occurs only if $\exp(-m''D^2kT/2\hbar^2)$ is not too small. Here, D is the mean spacing of the atoms, which is 36A in our example.) The specific heat contributed by the He^3 is then k per atom. This can exceed the specific heat of the phonons (below 0.4° in our example). A temperature pulse would then go mainly into increasing the energy of the He^3 gas. The speed of sound in this gas is of the order of the He^3 atom velocity, and therefore the observed second sound velocity should vary as $(kT/m'')^{\frac{1}{2}}$ in this region.[10] A more detailed analysis of the intermediate region in which both He^3 and phonons contribute requires a study of the collision cross sections for phonon-phonon, phonon-He^3, and He^3-He^3 collisions. Higher concentrations of He^3 require a study of the degenerate Fermi gas. The entire analysis of this paper fails to apply to pure He^3, because the ground state from which we begin is different.

An atom of He^3 should show an appreciably higher free energy when dissolved in He^4, than if that atom is replaced by He^4. This is because, as discussed in a previous paper,[3] for pure He^4 the partition function is the sum on all trajectories which start at some configuration of the atoms $\mathbf{z}_i$ and return to *any* permutation of the original configuration $P\mathbf{z}_i$. With a He^3 atom at $\mathbf{z}_1$ say (and no others nearby), the final configuration is limited to only those permutations for which this atom returns to $\mathbf{z}_1$. We can estimate the effect as follows: Neglect the mass difference of He^3 and He^4. Consider the nondiagonal matrix element $(z'|e^{-\beta H}|z)$ in which the final state differs from the initial state only in that the atom $\mathbf{z}_1$ is moved to another site $\mathbf{z}'_1$. All the other atoms may go to some permutation of the original positions. From what we have said in I, this should depend upon $\mathbf{z}_1$ and $\mathbf{z}'_1$ approximately through a factor $(\beta=1/kT)$,

$$(m'/2\pi\beta\hbar^2)^{\frac{1}{2}}\exp[-m'(\mathbf{z}_1-\mathbf{z}'_1)^2/2\beta\hbar^2], \qquad (1)$$

since the atom acts essentially as a free particle of mass m'. To get the partition function if the atom 1 is He^4, we must sum this over all possible sites $\mathbf{z}'_1$. At low temperatures, where the diffusion distance $(2\beta\hbar^2/m')^{\frac{1}{2}}$ exceeds the atomic spacing $d=(V_A)^{\frac{1}{3}}$, this is approximately the integral of (1) over all $\mathbf{z}'_1$ divided by V_A, the atomic volume. This gives V_A^{-1} for He^4. For He^3, $\mathbf{z}'_1$ must coincide with $\mathbf{z}_1$ so (1) gives $(m'/2\pi\beta\hbar^2)^{\frac{1}{2}}$. The ratio of the partition function for He^4 to that in which a He^4 atom is replaced by He^3 is therefore $(m'kT/2\pi\hbar^2)^{-\frac{1}{2}}V_A^{-1}$ (at low concentration). The extra free energy per atom of He^3 is therefore $(3/2)kT\ln(2\pi\hbar^2/m''kTV_A^{\frac{2}{3}})$ at low temperatures. (Since He^3 is lighter than He^4, m' is replaced by m''.)[11]

DISCUSSION

A number of problems are suggested by this work.

First, a detailed quantitative analysis of all of the properties of liquid He^4 below 0.5°K should be undertaken with the confidence that the problem is relatively simple. Only the phonons should be involved. Their wavelengths are long compared to atomic dimensions, and we have to do essentially with a continuous medium. The statistical mechanical aspects are considered in the appendix. The mean free path for phonon collisions could be computed if the nonlinearity of the medium is included. Work in this direction has been done by Landau and Khalatnikov.[12] An extension could be made to include the effects of small concentrations of He^3.

A more difficult class of problem, and one which we have left completely untouched, is the answer to question (c) of the introduction. Namely, what is the detailed nature of the excitations involved at the higher temperatures of 1 to 2.2°K? Most of the experimental work has been done in this region. The atomic viewpoint cannot claim a real understanding of the situation in liquid He II until this problem is solved. (See note added in proof.)*

There is a third group of problems which has not been touched upon. They involve the question,

(d) What is the mechanism of the Rollin film?[5]

This seems to be a problem of the very low-temperature behavior and should properly have been discussed in this paper. A suggestion of Bijl, de Boer, and Michels[13] involves the idea that the energy of a layer of the

[10] I am indebted to F. G. Brickwedde for calling my attention to this phenomenon.

[11] To the effect considered in the text, there must be added the large free-energy difference at absolute zero, which arises from the difference in zero-point energy occasioned by the difference in atomic mass. In first approximation, the wave function is unaltered but the kinetic energy, $-(\hbar^2/2m)\nabla^2$, is higher, for m is 3 instead of 4. This difference is, therefore, close to $\frac{1}{3}$ of the mean kinetic energy, per He^4 atom, in the ground state.

[12] L. D. Landau and I. M. Khalatnikov, J. Exptl. Theoret. Physik (U.S.S.R.) **19**, 637, 709 (1949).

[13] Bijl, de Boer, and Michels, Physica **8**, 655 (1941).

liquid depends strongly on the thickness of the layer, decreasing for thicker layers, even up to 100 atoms thick. If this view is correct, we should look for the answer by studying the energy of the ground state, to see if it is dependent on the shape of the container. It is possible that such a dependence exists, even for thick layers, because of the very long permutation rings involved at low temperatures in the condensed phase. These rings are long enough to wander over the entire volume of the liquid, so that the energy may be sensitive to the shape. We have not yet been able to verify quantitatively the correctness of this idea.

Finally, the problem of critical flow velocities, and the resistance to high-speed motions, remains unsolved.

The analysis of pure liquid He^3 requires a new start because our physical arguments so far have depended so strongly on the Bose statistics.

The author has profited from conversations with E. Wigner, H. A. Bethe, and R. F. Christy.

APPENDIX

In a previous paper,[3] I, an approximate partition function was proposed for liquid helium. Without modification it will not describe the phonon states correctly. The necessary modifications are discussed here.

In I, it was noted that the partition function of helium is the integral over all configurations $\mathbf{z}_i$ of the quantity

$$W(\mathbf{z}^N)=N!^{-1}\sum_P \int_{\mathrm{tr}P} \exp\left\{-\int_0^\beta \left[\frac{m}{2\hbar^2}\sum_i\left(\frac{d\mathbf{x}_i}{du}\right)^2 + \sum_{ij} V(\mathbf{x}_i-\mathbf{x}_j)\right]du\right\}\mathfrak{D}^N\mathbf{x}_i(u), \quad (2)$$

using the notation of that paper [I, Eq. (5)]. The integral $\int_{\mathrm{tr}P}$ is taken over all trajectories $\mathbf{x}_i(u)$ of the atoms which start from the positions $\mathbf{x}_i(0)=\mathbf{z}_i$ and end up at some permutation $\mathbf{x}_i(\beta)=P\mathbf{z}_i$ of $\mathbf{z}_i$. The sum is taken on all permutations P. It was pointed out that if a configuration $\mathbf{z}_i$ contained atoms nearly overlapping, or in some other unfavorable arrangement, the important trajectories $\mathbf{x}_i(u)$ would almost immediately move to release the energy of the unfavorable arrangement (for example, overlapping atoms would spring apart). The time for this was generally much less than β. Thus the various configurations could be given a weight $\rho(\mathbf{z}_1, \mathbf{z}_2\cdots\mathbf{z}_N)=\rho(\mathbf{z}^N)$. The slower motions of atomic diffusion contributed an additional exponential factor, so an approximate expression [I Eq. (7)],

$$W(\mathbf{z}^N)=N!^{-1}K_\beta\left(\frac{m'}{2\pi\beta\hbar^2}\right)^{3N/2} \times\sum_P \exp\left[-\frac{m'}{2\beta\hbar^2}\sum_i(\mathbf{z}_i-P\mathbf{z}_i)^2\right]\rho(\mathbf{z}^N), \quad (3)$$

was proposed. The factor ρ was to be (for low temperatures, large β) nearly independent of β, as it represents the effect of rapid local motions. It is $\phi(\mathbf{z}^N)^2$, where ϕ is the ground-state wave function.

On the other hand, it was remarked (footnote 11 of I) that general variations in density over large distances would not be so rapidly released. We now consider in detail the effect of these compressional waves.

To a sound wave of wave number K, and frequency $\omega=cK$, where c is the speed of sound, correspond phonons of energy $\hbar\omega$. A density fluctuation of this wavelength $\lambda=2\pi K^{-1}$ would take a time τ of order $1/\hbar\omega$ to decay (calling u "time" as in I). This time will exceed β, for wavelengths $\lambda>2\pi\hbar c/kT=2\pi\cdot 17\mathrm{A}/T^0K$. Since this distance exceeds the atomic spacing 3.6A and the diffusion distance $(2\beta\hbar^2/m)^{\frac{1}{2}}=4.8\mathrm{A}/(T^0K)^{\frac{1}{2}}$, there is clear separation of these waves from local atomic motions. Therefore, an expression like (3) is correct locally, for a density fluctuation over a small region is rapidly released so that its effect can be contained in ρ (by having ρ smaller for such fluctuations). But a fluctuation over long distances will not even out in a time β and is not correctly described in (3).

Choose a length $1/K_0$ exceeding the atomic spacing, but below the wavelength of sound excited at the given temperature ($\hbar cK_0\beta\gg 1$). For distances inside $1/K_0$ no new considerations are necessary, and (3) is locally correct. For long distances it must be altered. Let $n(\mathbf{R})$ be the average number density at $\mathbf{R}$ in the configuration $\mathbf{z}_i$—the average being taken over a region of volume $1/K_0^3$. We shall determine how the probability of this configuration depends on $n(\mathbf{R})$ [the result is (9) below].

We may describe the motions $\mathbf{x}_i(u)$ as local atomic movements (within distances K_0^{-1}) and general drift motions of the center of gravity of the atoms in a volume K_0^{-3}. It is convenient to describe the initial density distribution by imagining that it arose from an initially uniform distribution of density n_0 by a displacement. If the atoms originally at $\mathbf{R}$ were displaced by $\mathbf{D}_0(\mathbf{R})$, the density is $n(\mathbf{R})=n_0(1+\nabla\cdot\mathbf{D}_0)$, where $n_0=V_A^{-1}$ is the density averaged over the entire fluid. As the trajectories in (2) move, there is a general drift which we will describe by giving the displacement $\mathbf{D}(\mathbf{R}, u)$ as a function of u. What is the energy associated with this drift? First, to the kinetic energy of the atoms due to local motion there is an extra contribution from the general drift $(m/2V_A)(\partial\mathbf{D}/\partial u)^2$ per unit volume. Further, suppose a region temporarily has an extra high local density. This will limit the path-space volume available (and also change the average mutual potential energies). Therefore there will be an extra factor accumulated in each little interval of time. This we can write as a factor e^{-EdudV} for the volume dV in time du. This E (the energy resulting from the compression) depends just on the density, hence on $\nabla\cdot\mathbf{D}$. We expand it in powers of $\nabla\cdot\mathbf{D}$. The constant term may be omitted by changing the zero from which we

measure energy. The linear term gives nothing for its integral over all the liquid vanishes $\int \nabla\cdot\mathbf{D}dV=0$ from the conservation of mass. The quadratic term is written conveniently as $(1/2)mc^2V_A^{-1}(\nabla\cdot\mathbf{D})^2$, for c defined this way becomes the speed of sound. Higher-order terms produce phonon-phonon scattering, but we neglect them here. Thus, in addition to the features which go to make up (3) locally, there is a factor controlling the large scale motions,

$$\exp\left[-\frac{m}{2V_A\hbar^2}\int_0^\beta\int\left(\left(\frac{\partial\mathbf{D}}{\partial u}\right)^2+\hbar^2c^2(\nabla\cdot\mathbf{D})^2\right)dVdu\right]\mathfrak{D}\mathbf{D}(\mathbf{R},u), \quad (4)$$

where the $\mathbf{D}$ are to be summed over all displacement fields such that[14] $\mathbf{D}(\mathbf{R},0)=\mathbf{D}_0=\mathbf{D}(\mathbf{R},\beta)$, with

$$\nabla\cdot\mathbf{D}_0=V_An(\mathbf{R})-1.$$

This is best analyzed in momentum space. We put

$$\mathbf{D}(\mathbf{R},u)=\int\mathbf{A}(\mathbf{K},u)\exp(i\mathbf{K}\cdot\mathbf{R})d^3\mathbf{K}(2\pi)^{-3}.$$

The $\mathbf{A}$ field can be separated into transverse components $\mathbf{A}_1$ and $\mathbf{A}_2$ and a longitudinal component $B=\mathbf{A}\cdot\mathbf{K}/K$. Thus

$$\nabla\cdot\mathbf{D}=\int KB(\mathbf{K},u)e^{i\mathbf{K}\cdot\mathbf{R}}d^3\mathbf{K}(2\pi)^{-3}, \quad (5)$$

and (4) becomes

$$\exp\left[-\frac{m}{2V_A\hbar^2}\int_0^\beta\int\left[\left(\frac{\partial B}{\partial u}\right)^2+\hbar^2c^2\mathbf{K}^2B^2+\left(\frac{\partial\mathbf{A}_1}{\partial u}\right)^2+\left(\frac{\partial\mathbf{A}_2}{\partial u}\right)^2\right]d^3\mathbf{K}(2\pi)^{-3}du\right]\mathfrak{D}\mathbf{A}_1(\mathbf{K},u)\mathfrak{D}\mathbf{A}_2\mathfrak{D}B. \quad (6)$$

The transverse displacements act as free particle motions, for there is no restoring force. They contribute a factor $(m/2\pi\beta\hbar)^{\frac{1}{2}}$ per mode. This is already contained in (3) (but with m' for m, which makes little difference for the small fraction $V_AK_0^3$ of modes involved). We need not count them again in (6).

For each longitudinal mode, we must integrate an expression of the form

$$\exp\left\{-\frac{m\delta}{2V_A\hbar^2}\int_0^\beta\left[\left(\frac{dB}{du}\right)^2+\hbar^2c^2K^2B^2\right]du\right\}\mathfrak{D}B(u),$$

where δ is the K-space volume per mode for all $B(u)$ which begin and end at B_0. This may be easily done by a method explained in another connection by the author.[15] There results ($\omega=cK$),

$$\left(\frac{m\delta\omega}{2\pi V_A\hbar\sinh(\omega\beta\hbar)}\right)^{\frac{1}{2}}\exp\left[-\frac{m\delta\omega}{V_A\hbar}B_0^2\tanh\left(\frac{\omega\beta\hbar}{2}\right)\right]. \quad (7)$$

The complete partition function is the integral (2) over all initial configurations. We therefore integrate (7) over B_0 to get the contribution from this mode. The integral yields

$$[2\sinh(\omega\beta\hbar)\tanh(\tfrac{1}{2}\omega\beta\hbar)]^{-\frac{1}{2}}=[2\sinh(\tfrac{1}{2}\omega\beta\hbar)]^{-1}=\exp(-\tfrac{1}{2}\omega\beta\hbar)[1-\exp(-\omega\beta\hbar)]^{-1},$$

the usual partition function from such a mode. All the modes together contribute

$$\exp\left\{-\int\ln[2\sinh(\tfrac{1}{2}\omega\beta\hbar)]d^3\mathbf{K}(2\pi)^{-3}V\right\}, \quad (8)$$

the integral extending over all modes of wave number less than K_0. This factor in the partition function gives the usual Debye specific heat, varying as T^3 as long as our temperatures are, as we have assumed, small enough that $K_0c\hbar\beta\gg1$.

Multiplication of the factors (7) for all modes tells us that density fluctuations have a probability proportional to

$$\exp\left\{-\frac{mc^2}{V_A\hbar}\int[KB_0(\mathbf{K})]^2\tanh\left(\frac{\omega\beta\hbar}{2}\right)\frac{d^3\mathbf{K}}{(2\pi)^3\omega}\right\}, \quad (9)$$

where KB_0 is the Fourier transform of the density fluctuation [from (5)],

$$KB_0(\mathbf{K})=V_A\int(n(\mathbf{R})-n_0)e^{i\mathbf{K}\cdot\mathbf{R}}d^3\mathbf{R}. \quad (10)$$

Since, for large $\omega\beta\hbar$, $\tanh(\frac{1}{2}\omega\beta\hbar)$ is nearly 1, the density fluctuations of short wavelength are independent of the temperature. For these the factor (9) could just as well be combined with the temperature independent factor ρ to make a new effective ρ. This shows that the results do not depend on the exact choice of K_0. For long waves $\tanh(\frac{1}{2}\omega\beta\hbar)$ falls below 1 and wider fluctuations are permitted than would be expected from (3). A more accurate representation of (2) is then (3) with ρ containing general variations in density, these variations being weighed by multiplying by the factor (9), normalized. A factor $(m'/2\pi\beta\hbar^2)^{N/2}$ should be replaced by (8).

The purpose of this appendix is just to note that (3) does not automatically contain the phonon effects, but must be modified to include them. The study of such questions as the nature of higher-energy excitations and the character of the transition can presumably be made using (3) without modification. It is convenient that the sound wavelengths are so long that a nearly complete separation can be made of the local behavior and the behavior of the overlying compressional waves.

[14] In principle the final displacements could differ from the initial by an atomic spacing d (because atoms may be permuted). But the actual displacements permitted by (4) are very much smaller than d, so that the only important term from (4) is that for which the configuration is restored atom by atom.

[15] R. P. Feynman, Phys. Rev. 84, 108 (1951) Appendix C. The explicit answer is given in R. P. Feynman, Revs. Modern Phys. 20, 367 (1948) on page 386, by substituting $\gamma=0$, $q_i=q_0=B_0$, $\hbar=-iV_A\hbar^2/m\delta$, $\omega=i\omega\hbar$, $T=\beta$.

PHYSICAL REVIEW VOLUME 94, NUMBER 2 APRIL 15, 1954

Atomic Theory of the Two-Fluid Model of Liquid Helium

R. P. FEYNMAN
California Institute of Technology, Pasadena, California
(Received January 11, 1954)

It is argued that the wave function representing an excitation in liquid helium should be nearly of the form $\Sigma_i f(\mathbf{r}_i)\phi$, where ϕ is the ground-state wave function, $f(\mathbf{r})$ is some function of position, and the sum is taken over each atom i. In the variational principle this trial function minimizes the energy if $f(\mathbf{r})=\exp(i\mathbf{k}\cdot\mathbf{r})$, the energy value being $E(k)=\hbar^2k^2/2mS(k)$, where $S(k)$ is the structure factor of the liquid for neutron scattering. For small k, E rises linearly (phonons). For larger k, $S(k)$ has a maximum which makes a ring in the diffraction pattern and a minimum in the $E(k)$ *vs* k curve. Near the minimum, $E(k)$ behaves as $\Delta+\hbar^2(k-k_0)^2/2\mu$, which form Landau found agrees with the data on specific heat. The theoretical value of Δ is twice too high, however, indicating need of a better trial function.

Excitations near the minimum are shown to behave in all essential ways like the rotons postulated by Landau. The thermodynamic and hydrodynamic equations of the two-fluid model are discussed from this view. The view is not adequate to deal with the details of the λ transition and with problems of critical flow velocity.

In a dilute solution of He^3 atoms in He^4, the He^3 should move essentially as free particles but of higher effective mass. This mass is calculated, in an appendix, to be about six atomic mass units.

IN a previous paper,[1] II, a physical argument was given to interpret the fact that the excitations which constitute the normal fluid in the two-fluid theory of liquid helium were of two kinds. Those of lowest energy are longitudinal phonons. The main result of that paper was to give the physical reason for the fact that there can be no other excitations of low energy. It was shown that any others must have at least a minimum energy Δ. No quantitative argument was given to obtain this Δ nor to get an idea of the type of motion that such an excitation represents. In this paper we expect to determine Δ and the character of the excitations.

The physical arguments of II are carried a step further here to show that the wave function must be of a certain form. The form contains a function whose exact character is difficult to establish by intuitive arguments. However, the function can be determined, instead, from the variational principle as that function which minimizes the energy integral.

THE WAVE FUNCTION FOR EXCITED STATES

In II the exact character of the lowest excitation was not determined, but various possibilities were suggested. One is the rotation of a small ring of atoms. A second is the excitation of an atom in the local cage formed around it by its neighbors. Still a third is analogous to the motion of a single atom, with wave number k about $2\pi/a$, where a is the atomic spacing, the other atoms moving about to get out of the way in front and to close in behind. It is not clear that they are really distinct possibilities, for they might be merely different ways of describing roughly the same thing.

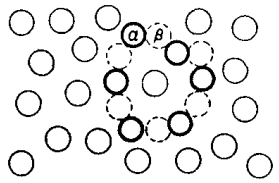


FIG. 1. Typical configuration of the atoms. If an excitation represents rotation of a ring of atoms such as the six in heavy outline the wave function must be plus if they are in the α positions and minus if they are moved to the intermediate β positions.

We shall now try to find the form of the wave function which we would expect under the assumption that one or another of these possibilities is correct. It will turn out that all of the alternatives suggest the same wave function, at least to within a function $f(\mathbf{r})$, of position $\mathbf{r}$, which is determined only vaguely.

First, suppose that the excitation is the rotation of a small ring of atoms. The number of atoms in the ring is determined, according to II, by the condition that it is the smallest ring that can be considered to be able to turn easily as an independent unit in view of the interatomic forces. For illustrative purposes we suppose this means that there are six atoms in the ring.

We can describe the wave function for this excitation by giving the amplitude associated with every configuration of the atoms. Suppose Fig. 1 represents a typical configuration, the six atoms of the ring in question (say ring A) being indicated by heavy outline. We discuss how the amplitude changes as we rotate this ring, leaving the other atoms out of account for a moment. Suppose the wave function is positive, say $+1$, if the atoms are in the position shown by the full circles in Fig. 1, which we arbitrarily call the α position. Suppose all the six atoms move around together, and let the ring turn about 60°. The atoms then appear again in α position, although which is which has been changed, so the wave function, by the Bose statistics, is still $+1$. On the other hand, for a 30° rotation, if the atoms are located as indicated in the figure by dotted circles (β position), the wave function will change to -1 for the first excited state. We need only discuss the real part of the wave function—the imaginary part, if any, can be dealt with in a similar way. (Actually since we deal with an eigenstate of the energy, the real part of

[1] R. P. Feynman, Phys. Rev. 91, 1291, 1301 (1953), hereafter called I, II, respectively.

the wave function is an eigenfunction also.) For orientations intermediate between α, β the function is correspondingly intermediate between $+1$ and -1, but to simplify the remarks we describe it for just the configurations α, β. The wave function for excitation of this ring we call ψ_A. It is $+1$ if the A ring is at α, and -1 if at β, and does not depend on how other rings of atoms are oriented. We can describe this wave function as follows. Consider a function of position $\mathbf{r}$ in space, $f_A(\mathbf{r})$ which is $+1/6$ if $\mathbf{r}$ is at one of the six positions of the centers of the atoms for the α position of ring A, is $-1/6$ if it is at a β position, and is zero if $\mathbf{r}$ is at any other place in the liquid far from the A ring. Then consider the quantity $\sum_i f_A(\mathbf{r}_i)$ where the sum is taken over all the atoms, i, in the liquid. For a configuration of the liquid for which there are atoms at the six α positions the quantity is $+1$, while if six atoms are at β position, it is -1. This suggests that we can write $\psi_A=\sum_i f_A(\mathbf{r}_i)$.

Actually this is incomplete because it does not correctly describe what happens if atoms in other parts of the liquid move. If ring A is in the α position, we wish the complete wave function to be $+1$ as far as this is concerned, but to drop to zero if two atoms overlap in other parts of the liquid, etc., just as for the ground state. That is, we expect (disregarding normalization)

$$\psi_A=\sum_i f_A(\mathbf{r}_i)\phi, \tag{1}$$

where ϕ is the ground-state wave function, a function of all the coordinates. This takes care of another matter also. What happens if some atoms are on α and some on β? This should be of very small amplitude because we do not wish the atoms to overlap on account of the repulsions. This is not correctly described by $\sum_i f_A(\mathbf{r}_i)$, but the ϕ factor does guarantee such a behavior. It is small for such overlaps. Of course, if the ring contained many atoms it could readjust just a little and the ϕ would not prevent, for example, all those near one side of the ring being α, and those on the opposite side of the ring being β. We are not guaranteed that (1) will describe well the amplitude for such a configuration. In fact, it wouldn't be expected that a function of just one variable could describe the motion of several atoms. However, by the arguments of II the ring is supposed to be small, in fact, so small that one part of the ring cannot move independently of the rest. The ring is so small that if one atom is at α, there cannot be a large amplitude for finding atoms at β because of the interatomic repulsions. This is represented in (1) by the factor ϕ which falls if two atoms approach (see II for a full description of the properties of ϕ).

Not knowing the exact size and shape of the ring we cannot say what the exact function $f_A(\mathbf{r})$ should be. But at least we conclude in this case the excited-state wave function is of the form

$$\psi=\sum_i f(\mathbf{r}_i)\phi, \tag{2}$$

where $f(\mathbf{r})$ is *some* function of position.

We might try to improve (1) by noting that, of course, the energy should be essentially the same if the excited ring were somewhere else in the liquid, say at B. The function

$$\psi_B=\sum_i f_B(\mathbf{r}_i)\phi \tag{3}$$

would describe this if $f_B(\mathbf{r})$ is $+1/6$ for $\mathbf{r}$ at some one of the six α positions of some other ring B, and $-1/6$ for intermediate β positions, and zero elsewhere. Or we could locate the ring at still another position, etc. Any one atom might be thought of as belonging to more than one ring. This produces a kind of interaction between adjacent rings. Because of this interaction, a better wave function than (1) might be some linear combination of these possibilities, say $c_A\psi_A+c_B\psi_B+\cdots$. But we can still conclude that the form of the wave function is given by (2), but now, with the function $f(\mathbf{r})=c_A f_A(\mathbf{r})+c_B f_B(\mathbf{r})+\cdots$, for any linear combination of functions of the form (2) is still of this form.

If the lowest excited state which we seek were something like the excitation of a single atom in a cage formed from its neighbors we would guess the wave function to be of the form (2) also. Because there would be a nodal plane across the cage, and we would take $f(\mathbf{r})$ to be positive if $\mathbf{r}$ is in the cage on one side of the plane, and negative if on the other, and to fall off to zero if $\mathbf{r}$ goes outside the cage. We do not care which atom is in the cage so the sum on i is taken over all atoms. Those which are outside the cage contribute nothing to the sum, because $f(\mathbf{r})$ is zero there. Further, there is no appreciable amplitude for there being more than one atom in the cage, because of the action of the factor ϕ which is very small if the atoms penetrate each other's mutual potential. The ϕ also takes care of the fact that the atoms in remote parts of the liquid behave independently of what the excited atom is doing, and act just as in the ground state. Further, linear combinations, representing the alternatives that the excited cage may be located at different places in the liquid, are still of the form (2).

The third possibility was only crudely described in II. It was noted that if the atoms were considered as roughly confined to cells, then a wave function representing the motion of an atom A could be $\exp(i\mathbf{k}\cdot\mathbf{r}_A)$, where $\mathbf{r}_A$ is the position of A, and it is assumed that as A moves about, the other atoms move around to make way for it so that the density is maintained roughly uniform. This would correspond in the liquid to a wave function

$$\exp(i\mathbf{k}\cdot\mathbf{r}_A)\phi, \tag{4}$$

where ϕ is the ground-state wave functions of all the atoms including A. The factor ϕ does the equivalent of keeping the atoms in cells so that the density is nearly uniform no matter where $\mathbf{r}_A$ is. For small $\mathbf{k}$ this is a possibility only if atom A is different from the others and does not obey the Bose statistics. If the symmetry is taken into account then we must replace this by

the symmetrical sum

$$\sum_i \exp(i\mathbf{k}\cdot\mathbf{r}_i)\phi. \tag{5}$$

If ϕ had no large scale density fluctuations this would be no wave function at all, because there would be just as many atoms in the region where $\exp(i\mathbf{k}\cdot\mathbf{r})$ is positive as where it is negative and the sum cancels out.[2] This is in concert with the idea that the wave function cannot depend on where atom A is on a large scale. For if A moves a long distance and the others readjust to keep the density uniform, on a large scale (the scale $1/k$ for small k), the result is just equivalent to the interchange of atoms and the wave function cannot change as a consequence of the Bose symmetry. On the other hand, if while the atom moves from one position to that of its neighbor the wave function changes sign and returns, then (5) may be allowed. That is, something like (5) with k of order $2\pi/a$ may be a possibility. This again is of the form (2), but with $f(\mathbf{r})=\exp(i\mathbf{k}\cdot\mathbf{r})$. The argument just given for this alternative is admittedly not as complete as for the others, mainly because the original idea of what the state is, was based on such a crude model of atoms in cells. Insofar as the idea can be carried over to the case of the true liquid perhaps we can say the form (5), or (2) will represent it.

Since all the examples have led to the same form, we might expect that a more general argument could be made for the validity of (2). This is, in fact, possible starting from the general argument given in II to show why the excited states, other than phonons, can be expected to have an excitation. It was pointed out there that the excited-state function ψ must be orthogonal to the ground state. For some configuration, say α, of the atoms it acquires its maximum positive value. Then it will be negative for some other, say β, which represents some stirring from the α configuration without change of large scale density (to avoid phonon states). But stirring reproduces a configuration nearly like α although with some atoms interchanged. Thus it is hard to get the configuration β to be very far (in configuration space) from α to keep the gradient of ψ small in going from α to β.

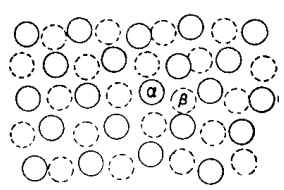


FIG. 2. In general the lowest excitation energy results if the configuration of atoms (solid circles) for which the wave function is most positive is as far as possible from that (dotted circles, β) for which it is most negative. All the β positions must be as far as possible from α positions, therefore.

[2] We shall see later that (5), for small $\mathbf{k}$, is actually a satisfactory wave function because ϕ does have the long wave density variations of the zero point motion of the sound field. We are trying to get excited states orthogonal to phonon states, and (5) for small $\mathbf{k}$ is not orthogonal. It is, in fact, just the wave function for such a phonon state. This is discussed later.

The lowest state would have the β configuration as far as possible from α. This means that in β as many atoms as possible are moved from sites (call them α positions) occupied by atoms in α. Hence β must be a configuration in which the atoms occupy sites (β positions) which are placed as well as possible between the α positions. (See Fig. 2.) In all these configurations, of course, the gross density must be kept uniform and the atoms should be kept from overlapping, to avoid high potential energy terms. If all atoms are on α positions ψ is maximum positive, and if all on β, maximum negative. The transition is made as smoothly as possible, and the kinetic energy thereby kept down, if for other configurations the amplitude is taken to be just the number of atoms on α positions minus the number on β positions. The number is just $\sum_i f(\mathbf{r}_i)$ where $f(\mathbf{r})$ is a function which is $+1$ if $\mathbf{r}$ is at an α position, and -1 if at a β position (and varies smoothly in between these limits as $\mathbf{r}$ moves about). It is of course a modulation to be taken on ϕ, because we wish to give small amplitude to configurations in which atoms overlap, etc., just as in the ground state. We are led, therefore, to (2). We can add the information that $f(\mathbf{r})$ must vary rapidly from plus to minus in distances of half an atomic spacing. That is, we expect that $f(\mathbf{r})$ will consist predominently of Fourier components of wave number $\mathbf{k}$ of absolute magnitude $k=2\pi/a$.

In the above argument it is not self-evident that in going from the configuration of all atoms at α positions to that of all at β, the amplitude must be just linear in the number on α, N_α minus the number of β, N_β. Perhaps some other smooth function of this number, like $\sin[\pi(N_\alpha-N_\beta)/2N]$ might be better. However, for the majority of possible configurations N_α and N_β are nearly equal; in fact, for almost all, $(N_\alpha-N_\beta)/N$ is of order $\pm N^{-\frac{1}{2}}$. For such a small range of the variable, the function, whatever it is, ought to behave nearly linearly. If the wave function (2) is wrong for a very few special configurations it will not be important as we shall determine the energies by the variational method, and the special configurations will contribute only a small amount to the integrals because of their small share of the volume in configuration space.

THE EXCITATION ENERGY

We have concluded that a function of the form (2) should be a good approximation to the wave function of the excited state.[2a] The function $f(\mathbf{r})$ is known only imperfectly, however. We shall determine this function $f(\mathbf{r})$ by using the variational principle. The Hamiltonian of the system is

$$H=-(\hbar^2/2m)\sum_i\nabla^2_i+V-E_0, \tag{6}$$

[2a] Wave functions of this form have been proposed before, for example by A. Bijl, Physica 7, 869 (1940). However, an argument establishing their validity for large $\mathbf{k}$ has been lacking, and it has not been clear that functions of other forms might not give much lower states.

where V is the potential energy of the system, and we measure energies above the ground-state energy E_0, so E_0 is subtracted in (6). Therefore the ground-state wave function satisfies

$$H\phi=0. \tag{7}$$

If we write

$$\psi=F\phi, \tag{8}$$

where F is a function of all the coordinates, then we can verify, using (7), that

$$H\psi=H(F\phi)=-(\hbar^2/2m)\textstyle\sum_i(\phi\nabla_i^2F +2\nabla_i\phi\cdot\nabla_iF)=\phi^{-1}(-\hbar^2/2m)\sum_i\nabla_i\cdot(\rho_N\nabla_iF), \tag{9}$$

where $\rho_N(\mathbf{r}^N)=\phi^2$ is the density function for the ground state, that is, the probability of finding the configuration $\mathbf{r}^N$ (we use $\mathbf{r}^N$ to denote the set of coordinates $\mathbf{r}_i$ of all the atoms, and $\int\cdots d^N\mathbf{r}$ to represent the integral over all of them).

The energy values come from minimizing the integral, (note ϕ is real)

$$\mathcal{E}=\int\psi^*H\psi d^N\mathbf{r} =(\hbar^2/2m)\textstyle\sum_i\int(\nabla_iF^*)\cdot(\nabla_iF)\rho_Nd^N\mathbf{r}, \tag{10}$$

subject to the condition that the normalization integral,

$$\mathcal{I}=\int\psi^*\psi d^N\mathbf{r}=\int F^*F\rho_Nd^N\mathbf{r}, \tag{11}$$

is fixed. The energy is then $E=\mathcal{E}/\mathcal{I}$.

In these expressions we must substitute

$$F=\textstyle\sum_if(\mathbf{r}_i). \tag{12}$$

Consider the normalization integral first. It is

$$\mathcal{I}=\textstyle\sum_i\sum_j\int f^*(\mathbf{r}_j)f(\mathbf{r}_i)\rho_Nd^N\mathbf{r}.$$

For a fixed i and j we can integrate first over all of the other atomic coordinates. This integral on ρ_N gives the probability for finding the ith atom at $\mathbf{r}_i$ and the jth at $\mathbf{r}_j$; therefore

$$\mathcal{I}=\int f^*(\mathbf{r}_1)f(\mathbf{r}_2)\rho_2(\mathbf{r}_1,\mathbf{r}_2)d^3\mathbf{r}_1d^3\mathbf{r}_2, \tag{13}$$

where ρ_2 is the probability of finding an atom at $\mathbf{r}_1$ per cm^3, and at $\mathbf{r}_2$ per cm^3. These density functions can be defined in general by

$$\rho_k(\mathbf{r}_1',\mathbf{r}_2'\cdots\mathbf{r}_k')=\textstyle\sum_i\sum_j\cdots\sum_n\int\delta(\mathbf{r}_i-\mathbf{r}_1') \times\delta(\mathbf{r}_j-\mathbf{r}_2')\cdots\delta(\mathbf{r}_n-\mathbf{r}_k')\rho_N(\mathbf{r}^N)d^N\mathbf{r}. \tag{14}$$

For example, $\rho_1(\mathbf{r})$ is simply the chance of finding an atom at $\mathbf{r}_1'$, for the liquid in the ground state. This is independent of $\mathbf{r}$ and is the number density ρ_0 in the ground state. In the same way $\rho_2(\mathbf{r}_1,\mathbf{r}_2)$ can be written as $\rho_0p(\mathbf{r}_1-\mathbf{r}_2)$ where p is the probability of finding an atom at $\mathbf{r}_2$ per unit volume if one is known to be at $\mathbf{r}_1$. Except near the liquid surface it is a function of only the distance from $\mathbf{r}_1$ to $\mathbf{r}_2$, so (13) is

$$\mathcal{I}=\rho_0\int f^*(\mathbf{r}_1)f(\mathbf{r}_2)p(\mathbf{r}_1-\mathbf{r}_2)d^3\mathbf{r}_1d^3\mathbf{r}_2. \tag{15}$$

The energy integral (10), with the substitution (12) becomes

$$\mathcal{E}=(\hbar^2/2m)\textstyle\sum_i\int\nabla_if^*(\mathbf{r}_i)\cdot\nabla_if(\mathbf{r}_i)\rho_Nd^N\mathbf{r}.$$

The integral of ρ_N over all atomic coordinates except $\mathbf{r}_i$ gives a result involving only $\rho_1(\mathbf{r}_i)=\rho_0$. Therefore we have simply

$$\mathcal{E}=\rho_0(\hbar^2/2m)\int\nabla f^*(\mathbf{r})\cdot\nabla f(\mathbf{r})d^3\mathbf{r}. \tag{16}$$

The best choice of f is that which minimizes the ratio of (16) to (15). The variation with respect to f^* gives the equation

$$E\int p(\mathbf{r}_1-\mathbf{r}_2)f(\mathbf{r}_2)d^3\mathbf{r}_2=-(\hbar^2/2m)\nabla^2f(\mathbf{r}_1),$$

where the energy E is $\mathcal{E}/\mathcal{I}$. This has the solution

$$f(\mathbf{r})=\exp i(\mathbf{k}\cdot\mathbf{r}), \tag{17}$$

with the energy value

$$E(\mathbf{k})=\hbar^2k^2/2mS(\mathbf{k}), \tag{18}$$

where $S(\mathbf{k})$ is the Fourier transform of the correlation function,

$$S(\mathbf{k})=\int p(\mathbf{r})\exp(i\mathbf{k}\cdot\mathbf{r})d^3\mathbf{r}. \tag{19}$$

It is a function only of k, the magnitude of $\mathbf{k}$.

It is readily verified that the solution is orthogonal to the ground state if we exclude $\mathbf{k}=0$. In fact, the solutions for different values of $\mathbf{k}$ are orthogonal to each other. This is because they all belong to different eigenvalues, $\hbar\mathbf{k}$, of the total momentum operator

$$\mathbf{P}=(\hbar/i)\textstyle\sum_i\nabla_i,$$

as is directly verified from (2) with (17), taking $\mathbf{P}\phi=0$ since the ground state has zero total momentum.[3] Since

[3] The argument is not rigorous because the momentum of the entire liquid can be changed without appreciable energy change by moving the center of gravity. This multiplies the wave function by a factor like $\exp(-i\hbar\mathbf{k}N^{-1}\cdot\Sigma_i\mathbf{r}_i)$. This function is so different from (2), however, that the orthogonality is probably not destroyed.

this operator commutes with the Hamiltonian, we have in (18) an upper limit to the energy for each value of k. (In fact, we could have obtained (17) from (2) by this argument.) Since we expect that (2) is a good wave function for functions f which vary from plus to minus in a distance of order $a/2$, we expect that (18) is not only an upper limit, but also a good estimate of the energy in a range of k's in the neighborhood of $k=2\pi/a$. In fact, our arguments suggest that $E(\mathbf{k})$ should have a minimum as a function of k in that region. These expectations are verified in the next section.

DISCUSSION OF THE ENERGY SPECTRUM

To find the consequences of (18) we shall have to discuss the behavior of $S(k)$ defined in (19).

The function $p(\mathbf{r}_1-\mathbf{r}_2)$ gives the probability per unit volume that a particle is at $\mathbf{r}_2$ if one is known to be at $\mathbf{r}_1$. If $\mathbf{r}_2$ is close to $\mathbf{r}_1$ it is zero, for the atoms cannot overlap. On the other hand, if $\mathbf{r}_2$ coincides with $\mathbf{r}_1$ there is an atom there, so p contains a delta function $\delta(\mathbf{r}_1-\mathbf{r}_2)$. For large r_2 it approaches ρ_0. Since the structure of the liquid ought to be more or less like that in a classical fluid, as r increases from zero, $p(r)$ probably rises to a maximum at the nearest neighbor spacing, falls, then rises again to a lower and wider maximum for next nearest, and with rapidly decreasing smaller oscillations approaches unity.[4] The integral $\int(p(\mathbf{r})-1)d^3\mathbf{r}$ vanishes since the integral of $\rho_2(\mathbf{r}_1,\mathbf{r}_2)$ with respect to $\mathbf{r}_2$ is exactly ρ_1 times the number of atoms $N=\rho_0 V$.

The Fourier transform function $S(k)$ is just the liquid structure factor which determines the scattering of neutrons (or x-rays, after multiplication by the atomic structure factor) by the liquid at absolute zero. It is therefore a quantity which can be directly determined experimentally. For large k it approaches 1 because of the delta function in $p(\mathbf{r})$. It has a delta function at $\mathbf{k}=0$, but this value of k is not of interest to us in (18), because the wave function ψ must be orthogonal to the ground state. The behavior at small k depends on the variations of $p(r)$ over long distances, that is, on long wavelength density fluctuations. These are the zero point fluctuations of the sound field in the ground state, since for wavelengths longer than the atomic spacing the approximation of a continuous sound field is good. This may be analyzed as follows. The operator representing the density at a point $\mathbf{r}$ is

$$\rho(\mathbf{r})=\sum_i\delta(\mathbf{r}-\mathbf{r}_i). \tag{20}$$

Its Fourier transform is

$$q_\mathbf{k}=\int\rho(\mathbf{r})\exp(i\mathbf{k}\cdot\mathbf{r})d^3\mathbf{r}=\sum_i\exp(i\mathbf{k}\cdot\mathbf{r}_i). \tag{21}$$

Evidently, the $S(\mathbf{k})$ is the expected value of $|q_\mathbf{k}|^2$ in the ground state. For long wave sound $q_\mathbf{k}$ is just the coordinate of the normal mode, so its mean square can be easily determined, for example, by noting that the mean potential energy is half of the ground-state energy $\frac{1}{2}\hbar\omega$. In this way one finds $S(\mathbf{k})=\hbar k/2mc$ for small k, where c is the velocity of sound.

The behavior of $S(\mathbf{k})$ for intermediate k is familiar to us from the x-ray studies of classical liquids. The density distribution in the ground state is roughly similar to such a liquid. There is some local structure produced by the tendency of the atoms to stay apart. This quasi-crystalline local order makes a maximum in the $S(k)$ curve for k near $2\pi/a$. There may be smaller subsidiary maxima for near multiples of this k. For helium, because of the large zero point motion, these maxima may be broader and less marked than in other liquids. The main maximum is responsible for the main ring in the x-ray diffraction pattern. It is shown clearly in the preliminary neutron diffraction data reported by Henshaw and Hurst.[5]

To summarize: with rising k, $S(k)$ starts linearly as $\hbar k/2mc$, rises then to a maximum near $k=2\pi/a$, and falls again to approach, with possible minor oscillations, the limit unity. Consequently the quantity $E(k)=\hbar^2k^2/2mS(k)$ should start linearly as $\hbar kc$, but should then show a dip with a minimum at $k=k_0$ say, near $2\pi/a$, finally rising, eventually as $\hbar k^2/2m$. These relations are shown in Fig. 3.

We have argued that (2) should be a good approximation to the wave function for functions that contain wave numbers in the vicinity of $2\pi/a$. Therefore we can expect the energy values (18) to be good in the neighborhood of this wave number. It is gratifying to see that there is a minimum in this region. The minimum value we shall call Δ. Ordinarily the variational method only permits one to interpret the minimum value of E as one varies a parameter such as k. On the other hand, in our case each value of k has significance since these values correspond to different eigenvalues of the momentum operator, as has been remarked. Therefore we can believe the behavior of the curve through a range of k near k_0, where it behaves parabolically, so we can write $E(k)$ in Landau's form $\Delta+\hbar^2(k-k_0)^2/2\mu$ where μ is a constant determining the curvature.

It is at first disconcerting that values of the energy lower than Δ can be obtained by going to very small values of k. But the energy here varies as $\hbar kc$, just that expected for phonon excitation. In fact, a moment's reflection shows that, for small k, the wave function (2) is just that which represents phonon excitation. Excitation of a given phonon means that the harmonic oscillator representing the corresponding normal mode is in the first excited state. The wave function is therefore $q_\mathbf{k}\phi$, if $q_\mathbf{k}$ is the normal coordinate of the mode excited. This coordinate is the Fourier transform of the density, so (21) shows that (2) with (17) represents a phonon for small k. Since the wave function is correct the energy must be exact, and is therefore $\hbar kc$.

[4] J. Reekie and T. S. Hutchison, Phys. Rev. 92, 827 (1953), have determined $p(r)$ by x-ray scattering.

[5] D. G. Henshaw and D.G. Hurst, Phys. Rev. 91, 1222 (1953).

Although we have made an argument only to show that (2) should be valid for high k, we see now that it is also valid for small k, that is for $f(\mathbf{r})$ which vary slowly. Since the energy curve is valid for the smaller k and for a range about $2\pi/a$, we can accept it as reasonable for all k from zero up to and slightly beyond the minimum.

On the other hand, for still larger k, another state of lower energy exists with the same total momentum. It is the state of double excitation, one of $\mathbf{k}_1$, the other of $\mathbf{k}_2$, such that $\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}$ and still $E(\mathbf{k}_1)+E(\mathbf{k}_2)<E(\mathbf{k})$. This becomes possible for k so high that the slope dE/dk of the energy curve exceeds $\hbar c$, the initial slope. The curve for very large k, therefore, does not have the same validity as that for lower k, but we need not enter into this matter, because at temperatures of a few degrees such high-energy states would not be appreciably excited. Such questions may be of importance in discussing nonequilibrium phenomena. One process by which the number of excitations can change is for an excitation to pick up enough momentum that it can divide spontaneously into two.

It is easy to misinterpret the meaning of the wave function

$$\psi=\sum_i \exp(i\mathbf{k}\cdot\mathbf{r}_i)\phi, \tag{22}$$

so a few remarks might be appropriate here. It looks at first, on inspection of the first factor, that this represents the excitation of a single particle. This is correct at very high $k(ka\gg 2\pi)$ and it is also correct for the ideal gas case for which the atoms do not interact (ϕ is constant then). But our arguments for intermediate k show that this is not the case. Because of the correlations in position implied by the factor ϕ, the motion of one atom implies the motion of others. Thus the factor in front of ϕ selects from that function certain correlated motions, in spite of the fact that each term in the factor depends on just one variable.

We can get a better idea of how this works by taking the extreme case of very low k. Here (22) represents a sound wave but at first sight there is no sign of the density variations that such a wave usually brings to mind. Let us take the real part and consider

$$\sum_i \cos(\mathbf{k}\cdot\mathbf{r}_i)\phi \tag{23}$$

for small k. Now, for most configurations, allowed by ϕ, the atoms are fairly uniformly distributed, so that there are just as many in the regions where the cosine is positive, as where it is negative. Therefore the sum over all the atoms of $\cos(\mathbf{k}\cdot\mathbf{r}_i)$ is zero. The wave function is zero for nearly all configurations. It is only for the rare configurations in which the number in positive regions exceeds that in regions where the cosine is negative that the wave function does not vanish. In this way (23) selects configurations for which the mean density varies as $\cos(\mathbf{k}\cdot\mathbf{r})$. Since such density fluctuations are, according to the behavior of ϕ, most likely produced by small cooperative motions of large numbers of atoms, the state described is very far from the one particle state it would be if the cosine factor appeared alone, not multiplied by ϕ.

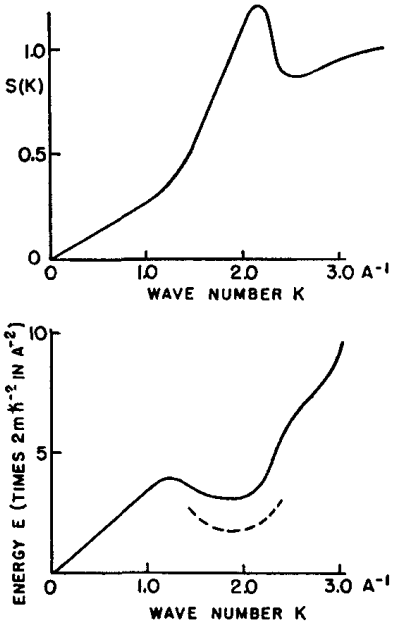


FIG. 3. The upper curve gives the liquid structure factor determined from neutron diffraction (reference 5) and extrapolated to zero k. The lower curve gives the energy spectrum of excitations as a function of wave number (momentum$\cdot\hbar^{-1}$) which results from the formula $E=\hbar^2k^2/2mS(k)$ derived in the text. The initial linear portion represents excitation of phonons while excitations near the minimum of the curve, where it behaves as $\Delta+\hbar^2(k-k_0)^2/2\mu$, correspond to Landau's rotons. However, data on the specific heat indicate that the theoretical curve should lie lower, closer to the dashed curve.

In the region of the energy minimum at k_0 the wave function represents a situation intermediate between the cooperative motion of phonons, and the excitation of a single particle. Several atoms move together because of the correlations implied by ϕ. It is hard to make a clear picture out of this vague idea. There is nothing to indicate that the state carries an intrinsic angular momentum. One must be careful because the state is degenerate, as all directions of k with the same magnitude k_0 give the same energy Δ. Perhaps, if more complicated wave functions were tried, some special linear combination representing a kind of microscopic vortex ring or one with intrinsic angular momentum has in fact a lower energy. States of low k will be called phonons, and states of momentum near k_0 will be called rotons in this paper, in accordance with the terminology of Landau,[6] although we do not necessarily mean to imply that rotons carry intrinsic angular momentum or represent vortex motion.

MULTIPLE EXCITATION

We have obtained the energy spectrum $E(k)$ of what we may call single excitations. They have the form of

[6] L. Landau, J. Phys. U.S.S.R. 5, 71 (1941); 8, 1 (1941). See also R. B. Dingle, Supplement to Phil. Mag. 1, 112 (1952).

plane waves through the liquid. By taking linear combinations we can make wave packets that are more or less confined to a local region. Unless the region is very small there would be only a negligible energy addition required to do this. The remainder of the liquid is quiet as in the ground state. It is conceivable that another packet could be located somewhere far away and the energy would be close to $E(\mathbf{k}_1)+E(\mathbf{k}_2)$ if $\mathbf{k}_1$ and $\mathbf{k}_2$ are the momenta of the majority of the waves in each packet. Thus we should expect states with several excitations, the energy being the sum of the energies of each packet separately. This neglects a kind of interaction energy between them. It will be valid if the density of excitations is very small, but one cannot expect to apply it to situations in which the number of excitations is any appreciable fraction of the number of atoms in the liquid.

Mathematically, if the function

$$F=\sum_i \exp i\mathbf{k}_1\cdot\mathbf{r}_i$$

gives $E(k_1)$ for the energy, we might expect the wave function for two excitations to correspond to $\psi=F\phi$, with

$$F=[\sum_i \exp(i\mathbf{k}_1\cdot\mathbf{r}_i)][\sum_j \exp(i\mathbf{k}_2\cdot\mathbf{r}_j)]. \quad (24)$$

It is readily verified, by substitution into the variation integral, that the energy is $E(\mathbf{k}_1)+E(\mathbf{k}_2)$ within correction terms of order $1/V$, where V is the volume of the entire fluid. This is just what one would expect if the excitations behaved like interacting particles, for the relative probability of their being within their range of interaction varies inversely as the volume. The expression (24) is unaltered on reversing the order of the factors, so the state in which the first excitation has momentum $\mathbf{k}_1$ and the second has $\mathbf{k}_2$ is the same as that in which the momenta are reversed. Thus the excitations obey Bose statistics. The expression (24) is not orthogonal to (17) with $\mathbf{k}=\mathbf{k}_1+\mathbf{k}_2$, so there undoubtedly are matrix elements between states of different numbers of excitations, and collisions must be possible which change this number. In summary, the excitations behave much like interacting Bose particles which may be created and destroyed, and whose energy as a function of momentum is given by $E(k)=\hbar^2k^2/2mS(k)$.

THERMODYNAMIC PROPERTIES OF HELIUM II

From this we may determine the thermodynamic behavior of liquid helium at low temperature. At sufficiently low temperatures the number of excitations will be small, so the interactions between them can be neglected. The approximation of independence leads in the usual way to the formula for the Gibbs free energy (taking the ground-state energy as zero),

$$F=kTV\int \ln[1-\exp(-\beta E[\mathbf{k}])]d^3\mathbf{k}(2\pi)^{-3}, \quad (25)$$

with $\beta=1/kT$. The number of excitations of momentum $\mathbf{k}$ is

$$n_\mathbf{k}=[\exp\beta E(\mathbf{k})-1]^{-1}. \quad (26)$$

We need not enter into further details as this has been thoroughly analyzed by Landau,[6] who first proposed the form of energy spectrum we have deduced here. At low temperatures only the lowest energy excitations can become excited. That is, only the phonons are excited and the specific heat varies as T^3. At higher temperatures some of the states near the minimum of the curve, at k_0 become excited. The specific heat then rises rapidly, controlled predominantly by the $\exp(-\beta\Delta)$ factor, governing the number of rotons excited. For temperatures of a few degrees few rotons are excited and only the phonon part, and the part of the curve near the minimum, are important. Landau[7] has shown that one obtains good agreement with the specific heat (and with the measured values of the velocity of second sound) if one chooses the parameters $\Delta=9.6°\text{K}$, $k_0=1.95\ \text{A}^{-1}$, and $\mu=0.77$. This means the energy curve near the minimum behaves as $2mE/\hbar^2=1.6+1.3(k-1.95)^2$, with k in reciprocal angstroms (one A^{-2} corresponds to a temperature of 6°K). In the phonon region the curve is

$$2mE/\hbar^2=2.6k,$$

in the same units, if the speed of sound is 240 meters/sec. Henshaw and Hurst[5] have published some preliminary data on the neutron scattering by liquid helium at 4.2°K. From it $S(k)$ may be directly determined (see Fig. 3). The curve for $E(k)$ calculated in this way behaves as

$$2mE/\hbar^2=3.0+1.0(k-2.0)^2$$

near the minimum (and is consistent with $2.6k$ for small k). This corresponds to a value of Δ of 18° which is impossibly large. Such a discrepancy may be due to the inaccuracy of the trial function (2), the true energy being lower than that calculated with this trial function. Such a large discrepancy in energy is discouraging, because the physical arguments did seem to indicate that (2) should be a reasonably good first approximation.

The expression (25) should not hold at high temperature because it neglects the interactions among the large number of excitations which (26) demands at such a temperature. Without an estimate of these interactions it is hard to judge the region in which deviations are to be expected. We shall make a very rough preliminary argument here.

To the approximation that the energy in a mode is proportional to an integer n, this mode behaves like a harmonic oscillator. The coordinate of this oscillator $q_\mathbf{k}$ has a mean square value $2n+1$ times its value in the ground state. For what size $q_\mathbf{k}$ is the harmonic oscillator approximation poor? If we knew this we could put a limit on the ranges of $q_\mathbf{k}$ and hence of $n_\mathbf{k}$, for which (25) might be expected to be valid. In our case the various $q_\mathbf{k}$ from (21) are not independent, because they can all be defined in terms of the same $3N$ variables $\mathbf{r}_i$.

[7] L. Landau, J. Phys. U.S.S.R. 11, 91 (1947); Phys. Rev. 75, 884 (1949).

Thus, for example, if q_k is known for $3N$ values of $\mathbf{k}$, it is known in principle for all others. It is very difficult to see what this interdependence means to (25).

But one can notice that one restriction is

$$\int |q_k|^2 d^3\mathbf{k}(2\pi)^{-3} = N, \tag{27}$$

where the integral is taken over all $\mathbf{k}$. Hence we may guess that we should restrict (25) by the condition that the excess $|q_k|^2$ over that for the ground state, summed on all states, cannot exceed N. This excess for a given mode is $2n_k$ times the ground-state mean value of $|q_k|^2$. This latter is $S(k)=\hbar^2k^2/2mE(k)$, so we obtain the restriction

$$\hbar^2 m^{-1} \int (k^2/E(k)) n_k d^3\mathbf{k}(2\pi)^{-3} = 1. \tag{28}$$

The thermodynamics which would result (by adding a chemical potential μ to $E(k)$ in (25), (26) to allow for (28)) will show a second-order transition.[8] But without a deeper analysis we are in no position to take Eq. (28) literally. We can only use it as a rough criterion for validity of (25). If the integral on the left is much less than 1, then (25) should hold. At 2°K the integral amounts to roughly 0.2, so perhaps we are entitled to trust (25) even to within a few tenths of a degree of the transition temperature.

MOTION OF THE FLUID AS A WHOLE

The existence of such excitations moving as nearly free particles in a background fluid is the central concept of the two-fluid model of Tisza[9] and Landau. The consequences of these ideas for excitations with a spectrum such as (18) have been carefully analyzed in a general manner by Landau and Dingle.[6] There is nothing to add that is new in this direction. However, we shall review briefly how the equations of this model arise, emphasizing the behavior of the wave function.

Beside the states which represent local internal excitation of part of the fluid, there are, of course, states in which the entire body of fluid moves. In general, in these cases the boundaries of the fluid move also. For example, at absolute zero, the entire fluid may move as a body with velocity $\mathbf{v}$. This center-of-gravity motion is described by

$$\psi = \exp(im\mathbf{v}\cdot\textstyle\sum_i \mathbf{r}_i)\phi,$$

if we assume ϕ corresponds to the ground state at rest in the laboratory system.

Suppose we wish to represent a situation in which the velocity $\mathbf{v}(\mathbf{r})$ varies from point to point, but only very gradually on an atomic scale. We might try something like this. The atoms in a region about some point P have their center of gravity moving at velocity $\mathbf{v}_P$ corresponding to this point. They must contribute a phase $m\mathbf{v}_P\cdot\sum_i \mathbf{r}_i$ where the sum is taken only over those near P. Corresponding contributions would come from sums near other points so the total factor ought to be $\exp[im\sum_i \mathbf{v}(\mathbf{r}_i)\cdot\mathbf{r}_i]$. The wave function is, therefore, of the form

$$\psi = \exp[i\textstyle\sum_i s(\mathbf{r}_i)]\phi, \tag{29}$$

where $s(\mathbf{r})$ is some function of position. We have suggested that it is $m\mathbf{v}(\mathbf{r})\cdot\mathbf{r}$. However, as is usual when one has waves whose wavelength varies from point to point, the wave number is not the phase divided by r, but more accurately it is the gradient of the phase. Therefore (29) represents the fluid in motion, the velocity at any point being given by

$$\mathbf{v}(\mathbf{r}) = m^{-1}\nabla s. \tag{30}$$

As a consequence of (31), $\nabla\times\mathbf{v}=0$. Velocity fields for which this is not true cannot be represented in such a simple manner, and represent, as we have seen in II, states involving large numbers of excitations. The problem they present is being studied. For regions which are not simply connected, such as a torus, s need not be single-valued. For example, in the torus we could take $s=\phi$, the cylindrical angle. This would represent a permanent circulation[10] even though $\nabla\times\mathbf{v}=0$ locally.

Substitution of (29) into the variational principle to obtain a steady-state solution leads [see (10), (11) with $F=\exp(i\sum_i s(r_i))$] to the energy expression

$$\frac{\rho_0}{2m}\int \nabla s(\mathbf{r})\cdot\nabla s(\mathbf{r}) d^3\mathbf{r}, \tag{31}$$

which is the kinetic energy $\rho_0 m\mathbf{v}^2/2$ per unit volume. It is minimum for variations in s if $\nabla\cdot(\nabla s)=0$, that is, $\nabla\cdot\mathbf{v}=0$. The flow must be incompressible. We have not allowed, in (29), for variations in density. For a singly connected region this has but one solution $\mathbf{v}=0$, unless the boundaries move. In a multiply connected region, like a ring, circulation of angular momentum in multiples of $N\hbar$ is possible. There are so few of these special states that the statistical mechanics is not affected. The variables, $\mathbf{v}(\mathbf{r})$, representing such motions can be specified as external known variables like pressure and volume.

EXCITATIONS IN A MOVING FLUID

Next we study the motion and energy of an excitation in a moving fluid. The wave function is

$$\psi = \textstyle\sum_i f(\mathbf{r}_i)\exp[i\sum_j s(\mathbf{r}_j)]\phi.$$

[8] This conclusion is modified if the interaction of the rotons and the hydrodynamic modes is taken into account.

[9] L. Tisza, Phys. Rev. 72, 838 (1947).

[10] It was suggested in II, reference 9, that to observe this experimentally one might have to avoid letting the liquid have a free surface. But R. Peierls has pointed out (private discussion) that although the atoms evaporating from the moving liquid to the gas carry angular momentum out, only those of the gas which are moving along with the liquid can condense, bringing back angular momentum—so in equilibrium there would be no damping of the motion from this effect. There are other effects, however, such as those which cause resistance to capillary flow at high velocities, which might be expected to damp the motion.

When this is put into the variational principle, one finds directly

$$\mathcal{E}/\mathcal{g}=E(k)+m\int \mathbf{j}(\mathbf{r})\cdot\mathbf{v}_s(\mathbf{r})d^3\mathbf{r} +\tfrac{1}{2}m\int \rho(\mathbf{r})\mathbf{v}_s(\mathbf{r})\cdot\mathbf{v}_s(\mathbf{r})d^3\mathbf{r}, \quad (32)$$

where $E(k)$ is given in (18), $\mathbf{v}_s(\mathbf{r})=m^{-1}\nabla s$, and we have defined

$$\rho(\mathbf{a})=\int \rho_3(\mathbf{a},\mathbf{r}_1,\mathbf{r}_2)f^*(\mathbf{r}_1)f(\mathbf{r}_2)d^3\mathbf{r}_1d^3\mathbf{r}_2/I, \quad (33)$$

$$\mathbf{j}(\mathbf{a})=\hbar\int \rho_2(\mathbf{a},\mathbf{r})[f^*(\mathbf{r})\nabla f(\mathbf{a}) -f(\mathbf{r})\nabla f^*(\mathbf{a})]d^3\mathbf{r}/Imi, \quad (34)$$

with

$$I=\int \rho_2(\mathbf{r}_1,\mathbf{r}_2)f^*(\mathbf{r}_1)f(\mathbf{r}_2)d^3\mathbf{r}_1d^3\mathbf{r}_2. \quad (35)$$

These ρ, $\mathbf{j}$, are the expected values of density and current density that belong to the state representing a single excitation.

We shall analyze this for the case of a single excitation of nearly definite momentum in the form of a large wave packet, large compared with the central wavelength. That is, we take $f(\mathbf{r})=\exp(i\mathbf{k}\cdot\mathbf{r})g(\mathbf{r})$ where $g(\mathbf{r})$ is a smooth amplitude function, such as a Gaussian, with width very large compared to $1/k$, but small compared to the size of the vessel. Such a packet will drift and spread slowly in a way completely determined by $E(k)$ and the principle of superposition. We wish to determine the additional effects of the possibility of general liquid flow.

The current density associated with this excitation is found from (34). We make the approximation that $\nabla f=i\mathbf{k}f$ since g varies so slowly, obtaining,

$$\mathbf{j}(\mathbf{a})=\hbar\mathbf{k}m^{-1}\int \rho_2(\mathbf{a},\mathbf{r})\{g^*(\mathbf{r})g(\mathbf{a})\exp[-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{a})] +g(\mathbf{r})g^*(\mathbf{a})\exp[+i\mathbf{k}\cdot(\mathbf{r}-\mathbf{a})]\}d^3\mathbf{r}/I.$$

Now, because of the variation of the exponentials, contributions to this integral come only from $\mathbf{r}$ within a limited distance from $\mathbf{a}$. Within such a distance $g(\mathbf{r})$ is nearly the same as $g(\mathbf{a})$, so all the g factors can be evaluated at $\mathbf{a}$ and taken outside the integral. The integral on $\mathbf{r}$ is then easy by (19) and one finds

$$\mathbf{j}(\mathbf{a})=\hbar\mathbf{k}\rho_0S(k)|g(\mathbf{a})|^2/mI.$$

The normalization integral may be done in a similar manner. It is

$$I=\int\int g^*(\mathbf{r}_1)g(\mathbf{r}_2)\rho_2(\mathbf{r}_1,\mathbf{r}_2) \times\exp[-i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)]d^3\mathbf{r}_1d^3\mathbf{r}_2.$$

Since $\mathbf{r}_2$ must be near $\mathbf{r}_1$ for a large contribution, we may replace $g(\mathbf{r}_2)$ by $g(\mathbf{r}_1)$, integrate $\mathbf{r}_2$ directly and obtain

$$I=\rho_0S(k)\int |g(\mathbf{r})|^2d^3\mathbf{r}.$$

If we assume $g(\mathbf{r})$ is normalized, $\int|g(\mathbf{r})|^2d^3\mathbf{r}=1$, so that $|g(\mathbf{a})|^2=d(\mathbf{a})$ is the density in the packet at $\mathbf{a}$, or roughly the probability of finding the excitation at $\mathbf{a}$, the current is

$$\mathbf{j}(\mathbf{a})=(\hbar\mathbf{k}/m)d(\mathbf{a}), \quad (36)$$

that is, a total current $\hbar\mathbf{k}/m$ distributed at density $d(\mathbf{a})$.

The particle density at $\mathbf{a}$ is

$$\rho(\mathbf{a})=\int \rho_3(\mathbf{a},\mathbf{r}_1,\mathbf{r}_2)g^*(\mathbf{r}_1)g(\mathbf{r}_2) \times\exp[-i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)]d^3\mathbf{r}_1d^3\mathbf{r}_2/I. \quad (37)$$

The points $\mathbf{r}_1$ and $\mathbf{r}_2$ must be close. If point $\mathbf{a}$ is not close to these points we can use the asymptotic form,

$$\rho_3(\mathbf{a},\mathbf{r}_1,\mathbf{r}_2)=\rho_0\rho_2(\mathbf{r}_1,\mathbf{r}_2), \quad (38)$$

to show directly that $\rho(\mathbf{a})$ is the density ρ_0 of the fluid, far from the packet. It is nearly so, even in the region of the packet, for since its dimensions are large, $\mathbf{a}$ is nearly always far from $\mathbf{r}_1,\mathbf{r}_2$, and further, the integral over all $\mathbf{a}$ of $\rho_3(\mathbf{a},\mathbf{r}_1\mathbf{r}_2)-\rho_0\rho_2(\mathbf{r}_1,\mathbf{r}_2)$ is exactly zero.

It is true that the distance of influence in ρ may not be very small, because of the correlations in the sound field. That is, the excitation produces a small strain in the fluid which makes a field of stress in the vicinity. Such fields provide a mechanism of interaction between excitations (as well as a correction to the energy of one). In a more detailed analysis such effects should be taken into account. Here we proceed to a first approximation and neglect them. To the approximation of neglecting compressibility, then, we find $\rho(\mathbf{a})=\rho_0$; the presence of an excitation does not change the fluid density.

Thus we picture an excitation in the form of a drifting wave packet as carrying a total current $\hbar\mathbf{k}/m$, and drifting (if $\mathbf{v}=0$) at the group velocity $\mathbf{v}_g=\partial E/\partial\mathbf{k}$, but as not appreciably altering the density.

This clearly violates the conservation of matter. For a moment we overlook this difficulty. It is discussed in the section following the next.

If this packet is in a general velocity field $\mathbf{v}_s(\mathbf{r})$ we may determine its energy from (32). In integral $\int\mathbf{j}(\mathbf{r})\cdot\mathbf{v}_s(\mathbf{r})d^3\mathbf{r}$ we shall assume that $\mathbf{v}_s$ does not vary appreciably over a region as small as the packet, and may be taken outside the integral sign. The integral of $\mathbf{j}$ is then $\hbar\mathbf{k}/m=\mathbf{p}/m$, giving the following results:

The energy of an excitation in a moving fluid is

$$E=E(p)+\mathbf{p}\cdot\mathbf{v}_s, \quad (39)$$

where $\mathbf{v}_s$ is the velocity of the fluid where the excitation (considered as a packet) is located. The total momentum

associated with the packet is $\mathbf{p}$, and it contributes a current $\mathbf{p}/m$ to the total in the fluid. The energy contributed by the moving fluid has density $\rho_0 m\mathbf{v}_s^2/2$, and it contributes to the current density $\rho_0\mathbf{v}_s$.

The group velocity of the excitation is $\partial E/\partial\mathbf{p}$ so that $\mathbf{v}_g=\mathbf{v}_{g0}+\mathbf{v}_s$ where $\mathbf{v}_{g0}$ is the group velocity in liquid at rest, $\partial E(p)/\partial\mathbf{p}$. Thus the excitation just drifts along with the background fluid motion, of velocity $\mathbf{v}_s$. Equation (39) can be obtained much more simply by a Galilean transformation of coordinates. Indeed, it is in this way that it was obtained by Landau and Dingle.[6]

RELATION TO THE TWO-FLUID MODEL

The two paragraphs at the end of the previous section contain the main relations by which the hydrodynamics and thermodynamics of the two-fluid model is derived. We review here a few of the steps, very briefly, in order to make clear the relation of the excitations to what is called the normal fluid. For further details see reference 6.

Thermodynamic equilibrium results, for a system with Bose statistics, if the excitations are distributed so that the number of those with energy E is

$$n=[\exp(\beta E)-1]^{-1}.$$

This may be obtained, for example, by maximizing the entropy, keeping the total energy constant. Another distribution which is also in equilibrium can be got by maximizing the entropy, keeping both total energy and total momentum constant. It is

$$n=\{\exp[\beta(E-\mathbf{p}\cdot\mathbf{u})]-1\}^{-1}, \tag{40}$$

where $\mathbf{u}$ is a constant.

In our liquid, the density of excitations per unit volume at a point $\mathbf{r}$ in this case would be, substituting (39) into (40),

$$n=\{\exp[\beta(E(\mathbf{p})+\mathbf{p}\cdot\mathbf{v}_s(\mathbf{r})-\mathbf{p}\cdot\mathbf{u})]-1\}^{-1}. \tag{41}$$

In order to interpret $\mathbf{u}$, we study the total current density. Since each excitation contributes a current $\mathbf{p}/m$ the total current density contributed by the excitations is

$$m^{-1}\int\mathbf{p}\{\exp[\beta(E(\mathbf{p})+\mathbf{p}\cdot(\mathbf{v}_s-\mathbf{u}))]-1\}d^3\mathbf{p}(2\pi)^{-3}.$$

At this stage we shall only consider the case of low macroscopic velocities. Expanding to the first order in in $\mathbf{v}_s-\mathbf{u}$ this may be written in the usual way as

$$\rho_n(\mathbf{u}-\mathbf{v}_s),$$

where ρ_n is defined as

$$\rho_n=\frac{\beta}{3m}\int p^2\{\exp[\beta E(p)]-1\}^{-2}d^3\mathbf{p}(2\pi)^{-3}. \tag{42}$$

To this we must add the current of the background $\rho_0\mathbf{v}_s$, so that the total macroscopic current density can be written

$$\mathbf{j}=\rho_n\mathbf{u}+\rho_s\mathbf{v}_s, \tag{43}$$

if we put $\rho_s=\rho_0-\rho_n$.

In view of these separations we can say, artificially, that the liquid behaves as though there were two parts, superfluid at density ρ_s moving at velocity $\mathbf{v}_s$, and normal at density ρ_n and velocity $\mathbf{u}$ (which we write hereafter as $\mathbf{v}_n$). The current is the sum of these two partial currents. In a similar manner the change of the internal energy, at constant entropy, produced by the velocities can be shown to second order to be the sum of the kinetic energies $\frac{1}{2}\rho_s\mathbf{v}_s^2+\frac{1}{2}\rho_n\mathbf{v}_n^2$.

In a vessel with fixed walls in thermal equilibrium, if the liquid background is flowing, its velocity $\mathbf{v}_s$ must have no component normal to the wall. Further, the total current normal to the wall must vanish, so that the normal component of $\mathbf{v}_n$ must also vanish at the walls. But in equilibrium $\mathbf{v}_n$ is constant everywhere and must therefore vanish everywhere. We say the normal fluid is stationary in equilibrium with fixed walls, even though the superfluid moves with the velocity $\mathbf{v}_s$. Incidentally, the superfluid velocity is irrotational, $\nabla\times\mathbf{v}_s=0$.

If the walls move together at constant velocity, then equilibrium results if $\mathbf{v}_n$ is this velocity; the normal fluid moves at the same velocity as the walls.

We extend Eq. (40) to situations slightly out of thermal equilibrium by assuming $\mathbf{v}_n$ is not constant but varies from place to place. The failure of equilibrium will bring in various irreversible processes associated with the normal fluid, such as viscosity. If we leave these out of account, the remainder of the hydrodynamic equations which result can be derived in exactly the manner already given by Dingle.[5] We need enter no further in this direction, as nothing new is gained. The resultant hydrodynamical equations can most easily be interpreted from the model of helium as consisting of two interpenetrating fluids.

Nevertheless, it is difficult to understand these partial fluids from a detailed kinematic point of view. Kinematically we have a general, or background fluid in which excitations move. The velocity of the superfluid, $\mathbf{v}_s$, is the general velocity of this background, but the density of superfluid is not ρ_0. The velocity of the normal fluid, $\mathbf{v}_n$, appears as a parameter in the distribution function. It can be shown to be the average group velocity of the excitations. But the difficulties arise if one tries to interpret the formula (42) for ρ_n from a direct kinematical point of view. It is not the average value of any quantity that can reasonably be ascribed to an individual excitation. It appears to have meaning only for the entire group of excitations in, or near, thermal equilibrium.

The division into a normal fluid and superfluid, although yielding a simple model for understanding

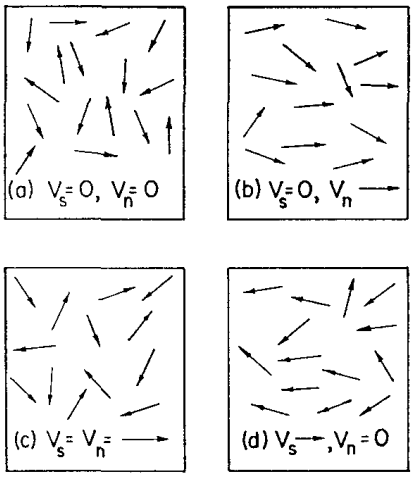


FIG. 4. At intermediate temperatures the main excitations are rotons which carry an intrinsic momentum, indicated by the arrows. If they drift relative to the background fluid, they tend to polarize upstream. This is illustrated for various values of v_s the background fluid velocity, and v_n the absolute drift velocity of the rotons. The total current is $\rho_0 v_s$ plus the polarization current of the rotons. Although mathematically correct, the separation of this current into the two parts, $\rho_s v_s$ and $\rho_n v_n$, characteristic of the two fluid model seems somewhat artificial from the microscopic viewpoint.

the final equations, appears artificial from a microscopic point of view. This opinion is shared by Landau and by Dingle.[5]

It is interesting to look at what is happening on a microscopic scale for various conditions of the velocities $\mathbf{v}_s$, $\mathbf{v}_n$. Consider a temperature not too low so that the predominant excitations are rotons. If the fluid is at rest a roton created with exactly the minimum energy Δ has no group velocity, but it has a momentum, of magnitude p_0 pointing in some direction. We will call it the direction of polarization and represent our roton by an arrow in this direction in Fig. 4. Not all rotons have exactly this energy Δ, but may differ by order kT from it, and have therefore nonzero group velocity. (Incidentally, the group velocity is parallel to or opposite to the polarization.) The rotons therefore may move about in a random manner like the molecules of a gas. Like gas molecules they can also have an average drift velocity relative to the fluid. Now we will assume [as required by (39), (41)] that if the rotons are drifting relative to the fluid in a certain direction, they tend (as a result of collisions among themselves) in equilibrium to polarize themselves in the direction in which they drift (that is, opposite to the velocity of the liquid moving past them). The general drift velocity of the rotons in space we call the normal fluid velocity, v_n. The motion of the fluid as a whole we call the superfluid velocity, v_s. Let us consider some examples.

First, with the fluid at zero velocity ($v_s=0$), and no drift of the rotons ($v_n=0$), they remain unpolarized (Fig. 4 (a)). If they are drifting to the right ($v_n>0$) they will tend to polarize in this direction, lining up to oppose the liquid passing them (Fig. 4 (b)). Now if the liquid is in motion ($v_s=v$), and all the rotons drift with the liquid in the same direction ($v_n=v$), there is no relative motion and no tendency to polarize (Fig. 4 (c)). This situation is not in equilibrium with stationary walls. Collisions of the rotons with the walls will stop their drifting motion and they will remain at rest relative to the walls. However, they will become polarized opposite to the fluid passing them (Fig. 4 (d)).

This interpretation of the velocities of the partial fluids of the two-fluid model is fairly simple and direct. It is otherwise with the current. The natural way to discuss the current (or momentum) from the microscopic view is to split it into two parts. First the current produced by the flow of the moving fluid, $\rho_0 v_s$, and second the current produced by the polarization. It is then easy to see what the current is in each case. In the first two cases, 4a, 4b, we have no general fluid motion so that the current is all due to polarization, zero in case *a*, to the right in case *b*. If the fluid moves, but the rotons remain unpolarized, as in 4c, the current is $\rho_0 v_s$, purely due to liquid motion. When the rotons polarize, as in 4d, to oppose this background motion the total current is reduced. But this natural separation is not the same as that utilized in the two fluid view. It is difficult to identify, for example, what is called the current of the normal fluid. It is not current carried by the rotons as they drift from one place to another (which they do with the velocity v_n) because the roton as such carries no mass but is only a disturbance in the liquid. Certainly the value of ρ_n would be hard to obtain this way because the rotons contribute to the current mainly by their polarization, and not by their drift motion. (Of course in a case such as 4b the polarization is in the direction of the drift so we could say the drift acts *as if* it carries current, because it induces polarization. ρ_n is then the ratio of the polarization current to the drift velocity which produces it.)

On the other hand, it is evident that the entropy flow is produced entirely by the drift motion (and not the polarization) of the rotons. Hence it is easy to see why all the entropy flows with the velocity v_n.

THE CONSERVATION OF CURRENT

In the last section we considered a packet of solutions (2) and found that we could picture an excitation in the form of a packet carrying a total current $\hbar\mathbf{k}/m$, and drifting at a group velocity $\partial E/\partial \mathbf{p}$, but as not appreciably altering the density. But such a picture is inconsistent with the conservation of matter. To take an extreme example, for a roton of the minimum energy Δ the group velocity $\partial E/\partial \mathbf{k}$ is zero, but the current $\hbar k_0/m$ is large. If such a current is distributed over a finite region in such a way that the direction is everywhere the same, we evidently cannot conserve material.

On the other hand, it is well known that one can

demonstrate the conservation of matter,

$$\partial\rho(\mathbf{a})/\partial t=\nabla\cdot\mathbf{j}(\mathbf{a}), \tag{44}$$

from the Schrödinger equation. The reason that our wave function does not satisfy (44) is that it is not an exact solution of the wave equation. This shows an inaccuracy in our approximate wave function (17).

One way that suggests itself to resolve it, in the case of rotons with $k=k_0$, is to propose a superposition of two waves with opposite momenta $\mathbf{k}$ and $-\mathbf{k}$, like $g(\mathbf{r})\cos(\mathbf{k}\cdot\mathbf{r})$. In this case the current density is zero, and everything is all right. Furthermore, if the same small momentum $\mathbf{l}$ is given to each, so the momenta become $\mathbf{k}+\mathbf{l}$ and $-\mathbf{k}+\mathbf{l}$ with $k=k_0$, the drift velocity $\partial E/\partial\mathbf{k}$ is the same for each partial wave, so that the packet stays together.

On the other hand, with stronger collisions with walls and phonons perhaps the two momentum components would become separated. Further $g(\mathbf{r})\sin(\mathbf{k}\cdot\mathbf{r})$ is just as good a solution, and it must have almost exactly the same energy even if interactions are taken into account, because the exact position of the nodes in a large packet cannot be important. Therefore, a linear combination must again be a possibility and we are led back to the exponential, and to the difficulty of current conservation. This lack of conservation is a symptom that all is not too well with our wave function. It is true that in the cosine case the symptom is hidden, but the conclusion should stand that the wave function could be improved.

The problem can be resolved by considering more complicated functions representing interaction of the excitation with the flow of fluid in its surroundings. One way the current could be conserved would be to have a general return flow of fluid in the region outside the packet. We therefore try the solution

$$\psi=\sum_i g(\mathbf{r}_i)\exp(i\mathbf{k}\cdot\mathbf{r}_i)\exp[i\textstyle\sum_j s(\mathbf{r}_j)]\phi, \tag{45}$$

with the hope of finding an s which produces a velocity distribution $\mathbf{v}=m^{-1}\nabla s$ which shows such a reverse flow. Let us first consider such a packet in otherwise stationary liquid. Then as a boundary condition s should go to zero as we go far from the packet. Substitution into the variation integral gives (32). For the current and density we use our approximations, that $\mathbf{j}(\mathbf{a})$ is given by (36), and $\rho(\mathbf{a})=\rho_0$. There results

$$\mathcal{E}/g=E(k)+\int\mathbf{j}(\mathbf{r})\cdot\nabla s(\mathbf{r})d^3\mathbf{r}+\frac{\rho_0}{2m}\int\nabla s\cdot\nabla s d^3\mathbf{r}, \tag{46}$$

where we have put, for the packet energy, $E(k)$, which is nearly correct. Variation of s to find a minimum gives the equation

$$\nabla\cdot(\mathbf{j}+\rho_0\nabla s)=0. \tag{47}$$

This equation determines s if we impose the boundary conditions $s\to 0$ far from the packet. Call this solution s_0 and the velocity distribution $\mathbf{v}_0=\nabla s_0/m$. It is like the field produced by the charge density $\nabla\cdot\mathbf{j}$, that is, at large distances the field of a dipole. It represents the back flow expected. Furthermore the total current operator has for our function the value

$$\mathbf{J}_0=\mathbf{j}+\rho_0\nabla s_0 \tag{48}$$

so that (47) says that now the total current is conserved.

There is a small shift in energy. Substitution of (47) into (46) gives the extra energy (reduction)

$$-\frac{\rho_0 m}{2}\int\mathbf{v}(\mathbf{r})\cdot\mathbf{v}(\mathbf{r})d^3\mathbf{r}.$$

If the order of the dimensions of the packet are L, the current $\hbar\mathbf{k}/m$ is distributed over a volume L^3, so the velocities are of order $\hbar k/mL^3$ and the kinetic energy $(\rho_0\hbar^2k^2/mL^6)L^3$ varies as $1/L^3$. But to confine the packet to such a dimension wave numbers of order $1/L$ in $g(r)$ must be used, so we find from (18) (for the case $k=k_0$) excess energies of order $1/L^2$ needed to confine the packet. Thus, for large packets, spreading the packet over even larger dimensions will decrease the energy, in spite of the energy of the currents we have just calculated. For extremely small packets our analysis does not hold because of the approximations made.

Here we have just gone far enough to save the theorem of conservation of current. We have only dealt with the background current in a semiclassical way. More complex states consisting of superpositions of expressions like (45) should be considered if a correct calculation of the quantum-mechanical "self-energy" of a roton due to coupling with the general velocity field is to be carried out. Since the "self-energy" is negative, the corrected value of Δ will be nearer the experimental result. This problem is being studied.

A more correct picture of a packet excitation, then, is that of a kind of region of polarization (that is, $\mathbf{j}$) which induces a distribution of velocity field around it, $\rho_0\nabla s_0$. The field is analogous to that produced by electrical polarization, the electric field $\mathbf{E}$ corresponding to the velocity field $\nabla s/m$, and the electric displacement vector $\mathbf{D}$ being analogous to the total current density, since its divergence vanishes. We can use this analogy to determine the behavior of the system directly, as it is easily verified that all the equations correspond. There is, however, one important difference. The signs of interaction are reversed. Thus (39) shows that rotons tend to line up opposed to the external field $\mathbf{v}$, while electric dipoles line up with the field. The analogy must therefore be completed with the remark that the rotons are dipoles of a gravitational type; that is, like poles attract, unlike repel.

The energy of a dipole in an external field is still the moment times the external field, even though the dipole itself creates some field of its own. None of the conclusions of the previous section are changed, therefore.

ROTON INTERACTION VIA THE VELOCITY FIELD

The fact that one roton creates a velocity field in which another may interact produces a kind of interaction between rotons. This is possibly one of the major sources of interaction, especially for not too high roton density. It is interesting to try to see what effect it has. Suppose we consider rotons as small packets all separated from one another and acting as dipoles of strength $\mathbf{p}/m$. Let us suppose there is an average polarization $\mathbf{P}$ per unit volume, and an average background velocity $\mathbf{v}_s$. The mean current density is then

$$\mathbf{J}=\rho_0\mathbf{v}_s+\mathbf{P}. \tag{49}$$

The actual field at any point is not $\mathbf{v}_s$ because of the local variations produced by the individual dipoles. Call the $\mathbf{w}$ velocity that an average dipole feels both from the average effect $\mathbf{v}_s$ and from its neighbors. The latter contribution is proportional to the polarization. In fact, as Lorentz showed for dipoles in random positions, it is $\frac{1}{3}\mathbf{P}/m$, hence

$$\mathbf{w}=\mathbf{v}_s+\alpha\mathbf{P}\rho_0^{-1}, \tag{50}$$

where[11] $\alpha=\frac{1}{3}$. The case $\alpha=0$ is the case previously studied which neglects direct effects between the dipoles.

The energy of a roton in this field is $E(\mathbf{p})+\mathbf{p}\cdot\mathbf{w}$. The statistical mechanics will then be governed by the function,

$$f=kT\int \ln\{1-\exp[-\beta(E(\mathbf{p})+\mathbf{p}\cdot\mathbf{w}-\mathbf{p}\cdot\mathbf{u})]\}d^3\mathbf{p}/(2\pi)^3. \tag{51}$$

Here $\mathbf{u}$ is zero for equilibrium with fixed walls, and is the normal fluid velocity, $\mathbf{u}=\mathbf{v}_n$. The average polarization then is given by

$$\mathbf{P}=m^{-1}\partial f/\partial\mathbf{w}. \tag{52}$$

The internal energy of the system is

$$U=\tfrac{1}{2}m\rho_0\mathbf{v}_s^2+m\mathbf{P}\cdot\mathbf{w}-(\alpha m/2\rho_0)\mathbf{P}^2+\langle E(\mathbf{p})\rangle, \tag{53}$$

where the average value of $E(\mathbf{p})$ is

$$\langle E(\mathbf{p})\rangle=\int E(p)\{\exp[\beta(E(p)+\mathbf{p}\cdot(\mathbf{w}-\mathbf{u}))]-1\}^{-1}d^3\mathbf{p}/(2\pi)^3$$
$$=f+TS-(\mathbf{w}-\mathbf{u})\cdot m\mathbf{P}.$$

The entropy is $S=-\partial f/\partial T$.

Expanding up to second order in the velocities one finds that the current can still be written as $\rho_s\mathbf{v}_s+\rho_n\mathbf{u}$, and the excess internal energy (at constant entropy and total current) as $\frac{1}{2}\rho_s\mathbf{v}_s^2+\frac{1}{2}\rho_n\mathbf{u}^2$, provided that one writes $\rho_s=\rho_0-\rho_n$ and

$$\rho_n=\rho_n^0(1+\alpha\rho_n^0/\rho_0)^{-1}, \tag{54}$$

where ρ_n^0 is the old expression (42) valid for the case $\alpha=0$. Therefore the expression for the velocity of second sound

$$c_2=\left(\frac{\rho_n}{\rho_s}\frac{TS^2}{c_v}\right)^{\frac{1}{2}},$$

is unaltered when expressed in terms of ρ_n, etc. Only the theoretical formula for ρ_n is slightly modified. However, the modification is appreciable only when ρ_n^0/ρ_0 is not small, that is, near the transition. At the transition where c_2 goes to zero, ρ_n must equal ρ_0 so that ρ_n^0 given in (42) must equal $1/(1-\alpha)$ or 1.5. Actually ρ_n^0 varies very rapidly in this region so this makes no appreciable change in evaluating Δ and p_0 from the data. Furthermore in this region there may be other interactions which should alter our statistical mechanical analysis anyway. (Actually we cannot even be sure that rotons act as small individual dipoles until we have improved the wave function to include the interaction with the velocity field, as a quantum field.)

With interacting dipoles we would expect the analogue of a transition corresponding to the Curie point for electric dipoles. The analog of the condition for the Curie point comes out to be exactly the criterion that the expression (54) for ρ_n (with ρ_n^0 substituted from (42)) becomes equal to ρ_0. There are a few surprises here, though. Firstly, ordinarily the Curie transition occurs as we lower the temperature, but here it appears on raising the temperature. That is because the Curie point depends markedly on the density. Dipoles polarize if they are cold and dense. In our case at low temperatures they are cold, but not dense enough. As the temperature rises, the density does also, very rapidly, until a point is reached where spontaneous polarization appears even though the temperature has been raised. Another surprise is the fact that there is a transition even if the local field effect is neglected ($\alpha=0$). This difference is a result of the change of sign of the forces. Our dipoles polarize most easily if arranged in a flat region, while for electrical dipoles a needle-like region is preferred. With all the dipoles polarized parallel in the sheet the outside field is zero if the internal field opposed the polarization. But this opposition of polarization and field is just the stable condition for the rotons. In fact, the mutual local field α tends to *de*polarize them and *raises* the transition temperature. (For Onsager's value of α, (54) shows no transition for any temperature.)

One might be tempted to speculate that we could carry the statistical mechanical analysis right up to the transition point and beyond, by simply assuming that the only interaction of importance among rotons is the coupling with the general velocity field. One would just hope that other interactions or limitations to the number of degrees of freedom are not as important as

[11] Onsager has shown that if one deals with permanent dipoles, mutually impenetrable and roughly spherical, this value of α is in error. If his analysis applies to our case, the value $\alpha=\rho_0(\rho_0+2\rho_s)^{-1}$ results [L. Onsager, J. Am. Chem. Soc. 58, 1486 (1936)].

one would otherwise guess. Aside from the amusing twist that helium I would then be the polarized, organized state, serious difficulties arise. One can analyze these things from the statistical formulas, if the velocities are not considered small and are not expanded. One can, without loss of generality, take states of total current zero, and for simplicity take $\alpha=0$. What happens is this. For any temperature below the transition there are two equilibrium states possible, one unpolarized with $v_s=0$, and the other polarized at finite v_s. Since the latter has higher free energy it is an unstable equilibrium, but the $v_s=0$ is stable (actually only metastable[12]). As we approach the transition point the polarization of the unstable state approaches zero. Above the transition point (more correctly, the point when $\rho_n=\rho_0$) only the $v_s=0$ state is in equilibrium and that is unstable. The instability arises this way. There is a high density of rotons. If a little polarization develops, their energy is reduced. This increases the number of rotons in equilibrium at a fixed temperature as well as the polarization, so that if there is no limit to the number of rotons there is no stable state.

On the other hand, if a limitation of roton number such as (28) is imposed, stable polarized states exist at the higher temperature. But the transition to that state occurs as a first-order transition, and there is another transition at still higher temperatures when the polarization disappears again.

It is therefore evident that we do not correctly describe the region very close to the transition by the usual energy expression (53) (with or without $\alpha=0$). The interactions between rotons is playing a more complicated role than (53) can describe.

In a previous paper an expression was given for the partition function which was presumably reasonably satisfactory right across the transition. However, the analysis was too difficult to carry out. Now that a more detailed picture of the behavior below the transition is available it may be easier to see how that expression can be treated. We still lack a clear picture of what happens in the few tenths of a degree on either side of the λ point.

INTERACTIONS BETWEEN EXCITATIONS

Interactions between the excitations will lead to various irreversible processes, such as viscosity, attenuation of second sound, etc. These questions have been studied by Landau and Khalatnikov.[13] The most important factor in the various mean free paths which are involved is the change in density of the phonons and rotons with temperature. The absolute cross sections depend on the details of the interactions. Interactions between phonons can be thought of as arising from a nonlinear equation of state. An interaction between a phonon and a roton would result if rotons have a different energy for different pressures. According to the theory presented here their energy is $\hbar^2k^2/2mS(k)$. If the liquid is compressed k^2 increases. On the other hand $S(k)$ probably increases even more rapidly from the increase in local order produced by squeezing the nearly impenetrable atoms into a smaller space. Therefore we expect Δ to decrease with pressure. This provides a mechanism for roton-phonon interaction. It also has other effects. The presence of a roton would cause a small increase in density in its neighborhood with the effect falling off inversely as the distance from the roton. This provides a mechanism of long-range interaction between rotons in addition to that due to coupling with the general velocity field. The roton-roton interaction at short distances is a more difficult problem, which probably cannot be adequately solved until a more accurate wave function is available for the roton state.

If the roton energy Δ decreases when the liquid density increases then we would expect that in equilibrium the liquid would shrink if the number of rotons is increased. The volume decrease as the λ point is approached is probably a consequence of this effect. The fall of the λ temperature with rising pressure is thermodynamically related. It may also be seen directly from (42) (supposing the λ point to be $\rho_n=\rho_0$) considering that Δ decreases as the density rises.

SUPERCONDUCTIVITY

It has been suggested that superconductivity is analogous to superfluidity. What can we learn of the former from our study of the latter? It is interesting that if the He atoms were charged (and their net charge canceled by a uniform fixed background charge of opposite sign) the liquid would imitate many of the features of a superconductor. In a magnetic field at absolute zero the London[14] equation, $\mathbf{j}=e\mathbf{A}/mc$, would hold. This is because stirring of atoms is equivalent to interchange so that in the lowest state the wave function cannot vary if atoms are stirred, and the part of the current depending on the wave function gradient vanishes. Other states would take a finite energy to create, there would be states of permanent circulation in multiply connected rings, there would be a second-order transition, etc. How this close analogy is to be interpreted is not clear. The things which, it has been argued here, apply physically to helium cannot be justifiably taken over to the case of Fermi particles, or such particles in interaction with lattice waves, without a complete investigation of their validity in the new environment. For example, at present there is not, in the author's opinion, justification for assuming

[12] That is, the conventional free-energy expressions for arbitrary v_s are, strictly speaking, not self-consistent. For any temperature there is always some value of v_s for which the free energy is less than its value for $v_s=0$.

[13] L. Landau and I. Khalatnikov, J. Exptl. Theoret. Phys. (U.S.S.R.) **19**, 637, 709 (1949).

[14] F. London, *Superfluids* (John Wiley and Sons, Inc., New York, 1950), Vol. I.

that the form (2) is a reasonable wave function for an assembly of Fermi particles, with ϕ the ground-state function for such particles. In fact, there are definite arguments against it. Possibly the close analogy should only be used to tell us what the problem of superconductivity is. It is, from this point of view, the problem of showing that, in the metal, aside from phonons there are no (or only very few) states of very low energy just above the ground state.

DISCUSSION

We still have left unsolved at least three basic questions. One is to find a clear description of the neighborhood of the transition. A second is to obtain a more perfect roton wave function. The third is to describe states for which the superfluid velocity is not vortex-free. So far we have $\nabla\times\mathbf{v}_s=0$. At high velocities when more energy is available, more complicated motions might be excited. The evidence of high resistance to capillary flow at the higher velocities indicates this. A new element must presumably be added to our picture. We hope to publish some views on this third problem at a later time.

We have limited ourselves to a qualitative analysis of the more curious features of the behavior of liquid helium. The problem of obtaining $S(k)$ or the correlation function for the ground state quantitatively from first principles is beyond the scope of this work.

It has been argued[1] that He^3 atoms in low concentration, in He^4 would act as a gas of free particles, but with an effective mass m'' higher than that of one atom. This mass m'' is calculated in the appendix, where it is found to be about six atomic mass units.

The author has profited from conversations with R. F. Christy and with Michael Cohen.

APPENDIX

According to II an impurity atom of He^3 (at infinite dilution) should behave as an essentially free particle except that its effective mass m'' should exceed the true mass of He^3 due to the inertia of the He^4 atoms which must make way for it as it moves. We shall calculate this excess mass here. First we suppose the impurity atom had the same mass m as the other He^4 atoms (i.e., we neglect the difference in mass of He^3 and He^4).

The wave function for such an atom (coordinates $\mathbf{r}_A$) moving with momentum $\hbar\mathbf{k}$ might be conjectured to be $\exp(i\mathbf{k}\cdot\mathbf{r}_A)\phi$, where ϕ is the ground state of the system (which is the same as if all the atoms were identical). This, as a trial function, gives the variational energy as $\hbar^2k^2/2m$ and shows no mass correction. It does not represent with sufficient accuracy the other atoms moving back when atom A moves forward. This suggests the trial function

$$\psi=\exp(i\mathbf{k}\cdot\mathbf{r}_A)\exp[i\textstyle\sum_i s(\mathbf{r}_i-\mathbf{r}_A)]\phi, \qquad (1\text{-a})$$

where the velocity field $\nabla s(\mathbf{r}_i-\mathbf{r}_A)$ represents a backflow which depends only on the distance from the impurity. To omit the term $i=A$ in the sum we take $s(0)=0$. Substitution into the variational principle gives

$$\mathcal{E}=(\hbar^2/2m)\int[k^2-2\mathbf{k}\cdot\textstyle\sum_i\nabla s(\mathbf{r}_i-\mathbf{r}_A)+\sum_i\sum_j\nabla s(\mathbf{r}_i-\mathbf{r}_A)\cdot\nabla s(\mathbf{r}_j-\mathbf{r}_A)+\sum_i\nabla s(\mathbf{r}_i-\mathbf{r}_A)\cdot\nabla s(\mathbf{r}_i-\mathbf{r}_A)]\rho_N d^N\mathbf{r}, \qquad (2\text{-a})$$

with $\mathcal{I}=1$, as ψ is normalized. We write $\rho_2(\mathbf{r}_i,\mathbf{r}_A)=\rho_0 p(\mathbf{r}_i-\mathbf{r}_A)$, and $\rho_3(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_A)=\rho_0 p_3(\mathbf{r}_i-\mathbf{r}_A,\mathbf{r}_j-\mathbf{r}_A)$, and measure all distances from the point $\mathbf{r}_A$, so that (2-a) becomes

$$2m\hbar^{-2}\mathcal{E}=k^2-2\mathbf{k}\cdot\int\nabla s(\mathbf{r})p(\mathbf{r})d^3\mathbf{r}+\int\int\nabla s(\mathbf{r})\cdot\nabla s(\mathbf{r}')p_3(\mathbf{r},\mathbf{r}')d^3\mathbf{r}d^3\mathbf{r}'+\int\nabla s(\mathbf{r})\cdot\nabla s(\mathbf{r})p(\mathbf{r})d^3\mathbf{r}. \qquad (3\text{-a})$$

Then s is to be chosen to minimize this expression.

The function $p_3(\mathbf{r},\mathbf{r}')$ is the probability of finding one atom at $\mathbf{r}$ and another at $\mathbf{r}'$ if there is an atom at the origin ($p(\mathbf{r})$ is just the probability of finding one at $\mathbf{r}$ if one is at the origin). We do not know what this function p_3 is, but in this problem we can approximate it by $p(\mathbf{r})p(\mathbf{r}')$ (except at the origin). A given atom is surrounded by many (eight or ten?) nearest neighbors, and the distribution along one radius r and another r' must be nearly independent except for the relatively small solid angle where the atoms at $\mathbf{r}$ and $\mathbf{r}'$ are close together. Even here, if the functions $s(\mathbf{r})$, $s(\mathbf{r}')$ are smooth enough the average of p_3 over such angles may still give nearly the same result as $p(r)p(r')$. With this substitution our problem is that of minimizing

$$2m\hbar^{-2}\mathcal{E}=\left[k-\int\nabla s(\mathbf{r})p(\mathbf{r})d^3\mathbf{r}\right]^2+\int\nabla s(\mathbf{r})\cdot\nabla s(\mathbf{r})p(\mathbf{r})d^3\mathbf{r}. \qquad (4\text{-a})$$

The error made by this approximation is

$$\Delta\mathcal{E}=(\hbar^2/2m)\int\int\nabla s(\mathbf{r})\cdot\nabla s(\mathbf{r}')[p_3(\mathbf{r},\mathbf{r}')-p(\mathbf{r})p(\mathbf{r}')]d^3\mathbf{r}d^3\mathbf{r}'. \qquad (5\text{-a})$$

The $s(\mathbf{r})$ which minimizes (4-a) behaves as a dipole field (as $\mathbf{k}\cdot\mathbf{r}/r^3$) at large distances and this produces some convergence difficulties in the first integral in (4-a). They may be easily straightened out as follows. The minimum energy is only very slightly altered if the function $s(\mathbf{r})$ is altered at very large distances in such

a way that it falls off eventually more rapidly than $1/r^2$. For such a function the integral may be done by parts, the integrated part vanishing, so we may write

$$2mh^{-2}\mathcal{E}=\left(k+\int s(\mathbf{r})\nabla p(\mathbf{r})d^3\mathbf{r}\right)^2 + \int \nabla s(\mathbf{r})\cdot\nabla s(\mathbf{r})p(\mathbf{r})d^3\mathbf{r}. \quad (6\text{-a})$$

But in this form all the integrals have a definite limit even if $s(\mathbf{r})$ has no convergence factor (and therefore varies as $1/r^2$). We may therefore use (6-a) and avoid ambiguities from conditionally convergent integrals.

The variational principle shows that s must be a solution of

$$\nabla\cdot[p(\mathbf{r})((1-\beta)\mathbf{k}-\nabla s(\mathbf{r}))]=0, \quad (7\text{-a})$$

where we have set

$$-\int s(\mathbf{r})\nabla p(\mathbf{r})d^3\mathbf{r}=\beta\mathbf{k}. \quad (8\text{-a})$$

Multiplication of (7-a) by $s(\mathbf{r})$ and integration, using (8-a), tells us further that

$$\int \nabla s(\mathbf{r})\cdot\nabla s(\mathbf{r})p(\mathbf{r})d^3\mathbf{r}=\beta(1-\beta)k^2,$$

so (6-a) says

$$\mathcal{E}=\frac{h^2k^2}{2m}[(1-\beta)^2+\beta(1-\beta)]=\frac{h^2k^2}{2m}(1-\beta),$$

and the effective mass is $m/(1-\beta)$, an increase over m of

$$\Delta m=\beta m/(1-\beta). \quad (9\text{-a})$$

If the direction of $\mathbf{k}$ is taken as the z axis, the solution of (7-a) may be written in the form,

$$s(\mathbf{r})=(1-\beta)k(z-zv(r)/r), \quad (10\text{-a})$$

where $v(r)$ is a function of radius $r=(\mathbf{r}\cdot\mathbf{r})^{\frac{1}{2}}$ only, satisfying

$$\frac{d}{dr}\left(r^2p(r)\frac{dv}{dr}\right)=2pv, \quad (11\text{-a})$$

and such that v approaches r at large distances. This is easily solved numerically. We used the values of $p(r)$ determined by Reekie and Hutchison.[4] Starting for small r where $p(r)=0$, so that $dv/dr=0$, we chose v at some convenient value and integrated out to radii so large that $p(r)$ was effectively its asymptotic value ρ_0. Asymptotically v has the form $c(r+B/r^2)$ where B, c are constants. Since the equation is homogeneous we may divide the entire solution by c to obtain one with the correct asymptotic form. By substitution of (10-a) into (8-a) one can show, using (11-a), that the expression for (9-a) can be written

$$\Delta m/m=4\pi\rho_0B-1, \quad (12\text{-a})$$

where we have used the fact that

$$\int_{0+}^{\infty}(\rho_0-p(r))4\pi r^2dr=1 \quad (13\text{-a})$$

[the origin, where $p(r)$ has a δ function being excluded in this integral]. Actually it is difficult to obtain accuracy with this method because the asymptotic form of v is sensitive to the values of $p(r)$ used. Those of Reekie and Hutchinson[4] extended only up to $r=6$A and had to be taken from a graph, so that (13-a) was not accurately satisfied without some small arbitrary readjustments of the values.

On the other hand, the solution showed that $s(\mathbf{r})$ was nearly proportional to z/r^3. Since (6-a) is a variational principle we can therefore obtain a good value of the energy much more simply. We substitute the trial function,

$$s(\mathbf{r})=Az/r^3, \quad (14\text{-a})$$

directly into (6-a) and determine the parameter A to minimize $\mathcal{E}$. This gives

$$\Delta m/m=\tfrac{1}{2}\left(\frac{4\pi}{3}\rho_0\right)^2\left[\int_{0+}^{\infty}r^{-6}p(r)4\pi r^2dr\right]^{-1}, \quad (15\text{-a})$$

or, with the data of reference 4, $\Delta m=0.70m$, or 2.8 atomic mass units. (The numerical solution of the differential equation gave the same result within its accuracy of about 10 percent.)

It is difficult to evaluate the small correction arising from the term $\Delta\mathcal{E}$ of (5-a), for p_3 is unknown. If the atoms locally are nearly on a lattice, say face-centered, or body-centered, of cubic symmetry, $\Delta\mathcal{E}$ vanishes with the trial function (14-a).

If the mass of the impurity atom is not four atomic units it is readily shown that Δm is unchanged, provided that the distribution $p(r)$ of atoms around the impurity is assumed to be unchanged. This is also expected physically for the extra mass is due to the motion of the He^4 atoms in the environment of the impurity. Therefore for a He^3 atom the effective mass should be 5.8 atomic mass units. The higher zero-point motion of a lighter atom changes $p(r)$ by pushing the neighbors farther away, thereby raising Δm a little (but this effect must be a fraction of a mass unit because the effective mass is almost certainly less than the 6.8 which it would be if the He^3 had the larger mass of 4 units).

This result is not in good agreement with the determinations which have been made from experiment, as summarized by Daunt.[15] These give values nearer 8 or 9 atomic mass units.

[15] J. G. Daunt, Advance in Physics (Phil. Mag. Supplement) 6, 209 (1952), particularly p. 258.

CHAPTER II

APPLICATION OF QUANTUM MECHANICS TO LIQUID HELIUM

BY

R. P. FEYNMAN
CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA

1. Introduction

Liquid helium exhibits quantum mechanical properties on a large scale in a manner somewhat differently than do other substances. No other substance remains liquid to a temperature low enough to exhibit the effects. These effects have long been a puzzle. It is supposed that they can all be ultimately understood in terms of the properties of Schrödinger's equation. We cannot expect a rigorous exposition of how these properties arise. That could only come from complete solutions of the Schrödinger equation for the 10^{23} atoms in a sample of liquid. For helium, as for any other substance today we must be satisfied with some approximate understanding of how, in principle, that equation could lead to solutions which indicate behavior similar to that observed.

Since the discovery of liquid helium considerable progress has been made in understanding its behavior from first principles. Some of the properties are more easily understood than others. The most difficult of these concern the resistance to flow above critical velocity. If we permit some conjectures of Onsager [1], however, perhaps a start has been made in understanding even these. The aim of this article is to describe those physical ideas which have been suggested to explain the behavior of helium which can most easily be related to properties of the Schrödinger equation.

We shall omit references to the phenomena involved in the Rollin

film. It appears that the film can be understood as being maintained by van der Waals attraction to the wall. The flow properties of the film are interpreted as a special case of flow properties of helium in leaks in general.

The article falls naturally into two main sections. First, there are phenomena in which the superfluid velocity is irrotational. Here we can give a fairly complete picture. The second part concerns the case in which vorticity of the superfluid exists. Our position here is less satisfactory and more uncertain. It is described here in considerable detail because of the interesting problems it presents.

2. Summary of the Theoretical Viewpoint

The first striking way that helium differs from other substances is that it is liquid even down to absolute zero. Classically at absolute zero all motion stops, but quantum mechanically this is not so. In fact the most mobile substance known is one at absolute zero, where on the older concepts we should expect hard crystals. Helium stays liquid, as London [2] has shown, because the inter-atomic forces are very weak and the quantum zero point motion is large enough, since the atomic mass is small, to keep it fluid even at absolute zero. In the other inert gases the mass is so much higher that the zero-point motion is insufficient to oppose the crystalizing effect of the attractive forces. In hydrogen the intermolecular forces are very much stronger, so it, too, is solid. In liquid ^{4}He there is a further transition at 2.2°K, the λ-transition, between two liquid states of different properties. A transition is expected (at 3.2°K) for such atoms if the interatomic forces are neglected, as Einstein [3] noticed. London [4] has argued that the λ-transition corresponds to the transition which occurs even in the ideal Einstein-Bose gas. The inter-atomic forces alter the temperature and, in a way as yet only imperfectly understood [5,6] the order of the transition, but qualitatively the reason for the transition is understood. We will concern ourselves here, only with the liquid He II, below the λ-point, and shall try to elucidate the qualitative reasons for some of its strange behavior. Also we explicitly limit our considerations to a liquid made purely of ^{4}He atoms so that the wave function must be symmetric for interchange of the atoms. We do not mean to imply anything about liquid helium ^{3}He, nor about superconductors, either by analogy nor by contrast. That is, we shall use the fact that the wave function is symmetric in many arguments without

stopping to inquire whether the symmetry is necessary part of the argument.

The central feature which dominates the properties of helium II is the scarcity of available low energy excited states in the Bose liquid [7, 8]. There do exist excited states of compression (i.e.: phonons) but states involving stirring or other internal motions which do not change the density cannot be excited without expenditure of an appreciable excitation energy. This is because, for quantum energies to be low, long wave lengths or long distances are necessary. But the wave function cannot depend on large scale modifications of the liquid's configuration. For a large scale motion, or stirring, which does not alter the density, only moves some atoms away to replace them by others *. It is essentially equivalent to a permutation of one atom for another ,and the wave function must remain unchanged by a permutation of atoms, because ^{4}He obeys the symmetrical statistics. The only wave functions available are those which change when atoms move in a way which is not reproducible by permutation, and therefore either, (1) movements accompanied by change in density (phonons), (2) movements over distances less than an atomic spacing, therefore of short wave length and high energy (rotons and more complex states), or (3) movements resulting in a change in the position of the containing walls (flow). We shall discuss these states in detail presently.

The scarcity of low energy excited states is the seat of many of the phenomena in the liquid. This has been known since the work of Landau who developed a theory of the liquid on the assumption of such scarcity. The specific heat is very low at low temperature and only rises rapidly above about 1°K when enough thermal energy is available to excite an appreciable number of the higher energy states (rotons). There are so few states excited that the excitations may be localized in the fluid like wave packets. These move about, collide with each other and the walls, and imitate the appearance that in the perfect background fluid there is another fluid or gas. This "gas" of excitations carries all the entropy of the liquid, may carry waves of number den-

* In the ideal gas the low excitations are those in which one or two atoms are excited to low states. These involve density changes and are more analagous to phonon states (but are even lower in energy than in the liquid because the ideal gas has infinite compressiblity, and therefore vanishing sound velocity). The interatomic forces in the liquid make it more imperative that if atoms are moved away from one point others move in to take their place, if high repulsive energies between nearby atoms are to be avoided.

sity (second sound, analagous to sound in an ordinary gas), finds it difficult to diffuse through long thin channels, tries to even up uneven velocity distributions among its roton "molecules" (viscosity), and acts in many ways as a normal fluid. Meanwhile the background in which the rotons travel, that is, the total body of fluid itself, can flow. It flows, at low velocity, without resistance through small cracks. The reason is that to have resistance, flow energy and momentum must go into heat, that is internal excitations (eg. rotons). The energy required to form a roton is not available (at the necessary momentum change) unless the fluid velocity is very high.

Actually it appears likely that helium in flow doesn't form rotons directly at all. Resistance sets in at a relatively low velocity (critical velocity) because apparently a kind of turbulence begins in the perfect fluid *. This cannot occur at lower velocities because energy is needed to create vorticity. And, if we accept Onsager's suggestion, the vorticity is quantized, the line integral of the momentum per atom (mass of atom times fluid velocity) around a closed circuit must be a multiple of h. Below the critical velocity not enough kinetic energy is available in the fluid to produce the minimum vortex lines.

We shall discuss first the way that the scarcity of states accounts for many of the properties of the liquid. Here we are summarizing work of many others, particularly Landau. It is thought best to reemphasize this viewpoint, since it is the one which is directly supported by quantum mechanics. Furthermore, in this way we are starting over the more familiar ground. Next we discuss the quantum mechanical view of the reason for the scarcity of states. Finally in the second part of the paper we discuss the quantized vortex lines proposed by Onsager.

3. Landau's Interpretation of the Two Fluid Model

One of the most fruitful ideas in interpreting the behavior of the liquid is the two fluid model. It was developed by Tisza [9] from analogy to the structure of an ideal Bose gas. It is often spoken of as a vague association of two penetrating fluids. Landau [10] has interpreted it in a definite manner. We review his interpretation here, although an excellent review by Dingle [11] already exists. He has strongly emphasized the fact that one might picture the helium as a background fluid in which excitations move. At absolute zero one has a perfect

* The author now considers his statement (reference 7) that the reason for flow resistance "cannot very well be a kind of turbulence", to be in error.

ideal fluid which may flow frictionlessly with potential flow. If heated, the heat energy excites the liquid. This it does by creating here and there within it excitations of some sort. These excitations can make their way from one place to another, collide with the walls and with each other, and give to helium some properties associated with the so-called normal fluid component, such as viscosity. Landau as a result of his study of quantum hydrodynamicas was lead to suppose the excitations to be of two kinds. Of lowest energy are the phonons, or quantized sound waves, whose energy E equals pc where p is the momentum and c the speed of sound. Above these separated by an energy gap Δ are those of another kind, called rotons. At first he supposed the energy of these to be given by $\Delta + p^2/2\mu$ if they have momentum p, where μ is an effective mass. Later he found that this did not agree with the experiments of Peskhov on second sound, and he proposed instead the formula

$$E_{\text{rot}} = \Delta + (p - p_0)^2/2\mu \tag{1}$$

where p_0 is some constant. He went further and suggested that all these excitations really are of the same class and differ only in momen-

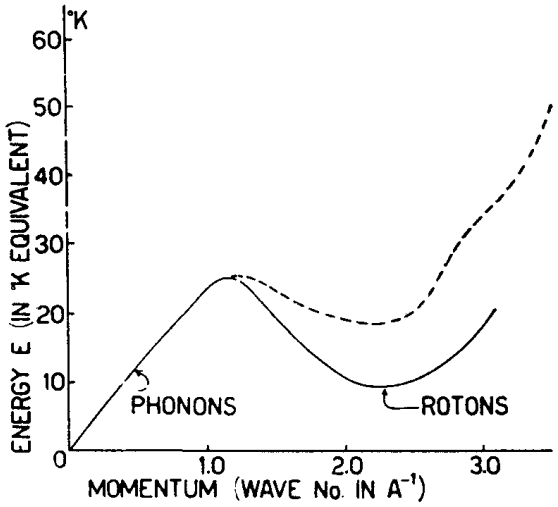


Fig. 1. The energy of excitations as a function of their momentum. Solid line as envisaged by Landau with parameters set to fit specific heat data; dotted line, an approximate curve derived from quantum mechanics.
Excitations in linear section for low momentum correspond to phonons. Those near the minimum of the curve are called rotons.

tum. The energy $E(p)$ of those of momentum p, depends only on the magnitude of p, rising at first linearly as pc, but later falling to a minimum at p_0 and rising again, as in Fig. 1, solid curve. The curve in the vicinity of the minimum is given by (1). At the low temperatures encountered in He II only the states near $p = 0$, and those close to the minimum are excited. Therefore we do not have to know the rest of the curve accurately. Furthermore, the only important excitations are one of the two classes, phonons and rotons.

Supposing the excitations to obey Bose statistics the number, at temperature T, of momentum in the range $d^3\boldsymbol{p}$ is, according to statistical mechanics,

$$n_{\boldsymbol{p}} = (\exp \beta E - 1)^{-1} d^3\boldsymbol{p}(2\pi\hbar)^3 \tag{2}$$

with $\beta = (kT)^{-1}$ and $E = E(p)$. From this the average energy $E(p)$ and the specific heat can be calculated. In agreement with experiment it begins at low temperature as T^{+3} as expected, according to Debye, since only phonons are excited. At higher temperatures the higher energy roton excitations become excited, and the specific heat rises much more rapidly. The thermodynamic properties are in excellent agreement with the theory if [12]

$$\begin{aligned} c &= 240 \text{ meters/sec} \\ \Delta/k &= 9.6^\circ\text{K} \\ p_0/\hbar &= 2.0 \text{ Å}^{-1} \\ \mu &= 0.77\, m \end{aligned}$$

where m is the atomic mass of helium.

The hydrodynamic equations of the two fluid model arise as follows. Suppose the fluid at absolute zero has density ϱ_0 and velocity $\boldsymbol{v}_s$. In the first part of this paper we shall take $\boldsymbol{v}_s$ to be irrotational $\nabla \times \boldsymbol{v}_s = 0$. Later we discuss the problem of local circulation. The mass current density is $\varrho_0\boldsymbol{v}_s$ and the kinetic energy is $\frac{1}{2}\varrho_0\boldsymbol{v}_s^2$. Suppose that as a result of a rise in temperature a limited number of excitations are formed in the fluid. Landau has shown that the energy to form excitations in a moving fluid is not $E(p)$ but is

$$E = E(p) + \boldsymbol{p} \cdot \boldsymbol{v}_s \tag{3}$$

This results from simple considerations of the relations in moving and still frames of reference. The mass current density equals the momentum density of the fluid since all of the atoms have the same mass. It now is

$$\boldsymbol{j} = \varrho_o \boldsymbol{v}_s + <\boldsymbol{p}> \tag{4}$$

where $<\boldsymbol{p}>$ is the mean momentum of the excitations per unit volume. Now the mean $<\boldsymbol{p}>$ depends on how the excitations drift. If they are in equilibrium with the fixed walls of the vessel the mean $\boldsymbol{p}$ is *not* zero. The energy is $E(p) + \boldsymbol{p} \cdot \boldsymbol{v}_s$. It is lower than $E(p)$ for those excitations, whose momentum is directed oppositely to $\boldsymbol{v}_s$. Therefore in equilibrium more excitations align oppositely to $\boldsymbol{v}_s$ than parallel to it. For this reason the mean $\boldsymbol{p}$ is directed oppositely to $\boldsymbol{v}_s$ and for small $\boldsymbol{v}_s$ is proportional to it, let us say $<\boldsymbol{p}> = -\varrho_n \boldsymbol{v}_s$. This defines ϱ_n. If ϱ_s is defined as $\varrho_o - \varrho_n$ we have a total current $\varrho_s \boldsymbol{v}_s$ in a situation in which the excitations are in equilibrium with fixed walls. The equilibrium is established by collisions of the excitations with the walls and with each other.

The number of excitations of momentum $\boldsymbol{p}$ is again determined by (2) but now with E given by (3) so that the average $\boldsymbol{p}$ is

$$<\boldsymbol{p}> = \int \boldsymbol{p}(\exp \beta(E(\boldsymbol{p}) + \boldsymbol{p} \cdot \boldsymbol{v}_s) - 1)^{-1} d^3\boldsymbol{p}(2\pi\hbar)^{-3}$$

or expanding to first order in $\boldsymbol{v}_s$, find $<\boldsymbol{p}> = -\varrho_n \boldsymbol{v}_s$ where

$$\varrho_n = -\frac{\beta}{3} \int p^2(\exp(\beta E(p)) - 1)^{-2} \exp \beta E(p)\, d^3\boldsymbol{p}\, (2\pi\hbar)^{-3} \tag{5}$$

The density ϱ_n determined from experiments in second sound is in reasonable agreement with this expression (evaluated with the constants given above it fits above 1°K, but below 1°K the values $p_o/\hbar = 2.3$ Å^{-1} and $\mu = 0.40\, m$ fit better, and do not alter the good fit to the thermodynamic data). This explicitly shows that ϱ_n is a derived concept, and does not represent the density of anything which has microscopic meaning.

The excitations can drift also. The distribution for equilibrium in a drifting gas is, according to statistical mechanics,

$$n(E) = (\exp \beta(E - \boldsymbol{p} \cdot \boldsymbol{u}) - 1)^{-1} \tag{6}$$

where $\boldsymbol{u}$ is a parameter. In this case the mean momentum is

$$<\boldsymbol{p}> = -\varrho_n(\boldsymbol{v}_s - \boldsymbol{u}).$$

If we write $\boldsymbol{u} = \boldsymbol{v}_n$ we have for the current

$$\boldsymbol{j} = \varrho_o \boldsymbol{v}_s - \varrho_n(\boldsymbol{v}_s - \boldsymbol{u}) = \varrho_s \boldsymbol{v}_s + \varrho_n \boldsymbol{v}_n \tag{7}$$

This can be interpreted macroscopically as saying that the current is like that in a mixture of two fluids, one of density ϱ_s moving at velocity $\boldsymbol{v}_s$, the other of density ϱ_n and velocity $\boldsymbol{v}_n$.

Actually (6) is not an equilibrium distribution unless the walls move at velocity $\boldsymbol{u}$, and furthermore $\boldsymbol{u}$ is constant throughout the liquid. It is generally taken as a good approximation in the case that $\boldsymbol{u}$, that is, $\boldsymbol{v}_n$, is not constant. The lack of equilibrium in this case produces irreversible effects, such as viscosity, which can be associated with the "normal fluid component". The distribution is in equilibrium even if $\boldsymbol{v}_s$ is not constant.

The entropy of the system is that of the excitations. It is easily verified that the mean group velocity of the excitations (the mean of $\delta\,(E(\boldsymbol{p}) + \boldsymbol{p} \cdot \boldsymbol{v}_s)/\delta \boldsymbol{p}$) is just $\boldsymbol{v}_n$. The entropy can therefore be considered as flowing with the "normal fluid".

It is also possible to work out the expected value of the energy of the system. If one calculates the internal energy and subtracts the internal energy the system would have at the same entropy but with $\boldsymbol{v}_s = \boldsymbol{v}_n = 0$ the excess expanded to the second order in the velocities can be written $\frac{1}{2}\varrho_s \boldsymbol{v}_s^2 + \frac{1}{2}\varrho_n \boldsymbol{v}_n^2$. This is just what the two fluid model would expect.

Therefore Landau shows that a liquid system with excitations as described will behave in many ways like a mixture of two fluids.

Furthermore, considerable progress has been made by Landau and Khalatnikov [13] in the interpretation of many irreversible phenomena, such as viscosity, attenuation of second sound, etc. from the kinetic theory implied by such excitations. It is not possible as yet to find the crosssection for collision, say between two rotons, from first principles. But if a few such quantities are considered as unknown parameters, a great deal can be said. The number of rotons varies very rapidly with temperature, in the manner given by (2). For this reason the mean free path for collision and the resultant viscosity resulting from roton-roton collisions has a known temperature dependence. In a similar way the contribution of collisions between rotons and phonons or between phonons can be worked out. There are also collisions in which the number of excitations change. The results are often in excellent agreement with experiment.

There is, therefore, little doubt that in liquid helium there are such excitations, with the energy spectrum that Landau suggests, and that this picture supplies the complete interpretation of the two fluid model for helium II.

4. The Reason for the Scarcity of Low Energy States

The next question that concerns us is to try to see from first principles why the excitations of the helium fluid have these characteristics.

Landau has, in fact, tried to obtain some justification for the spectrum from a study of quantum hydrodynamics. This is not a completely detailed atomic approach. One attempts to describe the liquid by a few quantities such as density and current, or velocity. Then one makes these quantities operators with reasonable commutation relations, and tries to find the excitation energies of the fluid. The problem has not been analyzed in sufficient detail to establish the energy spectrum (1). Such an approach cannot give us an ultimate detailed understanding for two reasons. First, the numerical values of Δ, p_o, μ show these quantities to be characteristic of the atomic structure of the liquid. A theory which describes the fluid simply by average variables and which therefore cannot represent the fact that the liquid does in fact have atomic structure cannot lead to definite values for excitation energy. A more serious problem is this. It is necessary to show not only that the excitations $E(p)$ exist, but that there are not a host of other possible excitations lying lower. If we describe the liquid with average variables we have no assurance that there are no excitations at a level below the coarsness of our averages. Possibly excitations exist which represent no gross density variation and no mean current. If many other lower excitations exist they dominate the specific heat curve and the properties of the fluid. (Perhaps in ^{3}He we have an example of a system capable of excitations at an atomic level which are not describable by the variables used in quantum hydrodynamics).

However it is possible from first principles to see why there are no other excitations but those supposed by Landau and why the energy spectrum of these excitations has, qualitatively, the form which he supposed. [7, 8]

In order to do so, we should, rigorously, have to solve the Schrödinger equation for the system.

$$H\psi = -\frac{\hbar^2}{2m} \Sigma_i \nabla_i^2 \psi + \Sigma_{ij} V(\boldsymbol{R}_{ij})\psi = E\psi$$

where m is the atomic mass, $V(\boldsymbol{R}_{ij})$ is the mutual potential of two atoms separated by the distance $\boldsymbol{R}_{ij} = \boldsymbol{R}_i - \boldsymbol{R}_j$, and ∇_i^2 represents the Laplacian with respect to the coordinates $\boldsymbol{R}_i$ of the i^{th} atom.

The sums must be taken over all of the N atoms in the liquid. We cannot solve this equation directly but we can make surprising headway in guessing the characteristics of the wave functions ψ which satisfy it.

We shall have to picture the wave function ψ. It differs from one state to another. But we will consider its value for only one state at a time. Then it is a definite but complicated function $\psi(\boldsymbol{R}_1, \boldsymbol{R}_2 \ldots \boldsymbol{R}_N)$ of the $3N$ variables $\boldsymbol{R}_i$. To picture it we must have a scheme by which we clearly represent it in our minds. Now such a function is a number associated with every set $\boldsymbol{R}^N$ of values $\boldsymbol{R}_i$, or, as we shall say, with every configuration of the atoms. We can represent a configuration $\boldsymbol{R}^N$ by imagining each of the N atoms in the vessel containing the liquid to be located with its center at one of the $\boldsymbol{R}_i$. That is, each configuration is represented, as classically, as a particular definite location for each of the atoms. Then $\psi(\boldsymbol{R}^N)$ is a number associated with each such arrangement of the atoms. We can call it the amplitude of the configuration. For a given state, this amplitude for some atomic arrangements is large – these arrangements then have large probability – for others small and the configuration is unlikely. When we wish to speak of how the amplitude changes as the values of $\boldsymbol{R}_i$ change, we shall use the more vivid language of asking how the amplitude changes as the atoms are "moved" about. Such motions are not directly related to any real classical motions, of course. In fact we cannot describe classical motions directly. All such classical ideas must be interpreted in terms of the mathematical behavior of ψ, if we are to be consistent with quantum mechanical principles. Most of our task, therefore, is trying to describe the ψ functions which correspond to the various kinds of states of energy, or motion, of which the liquid is capable.

Start by considering the ground state wave function which we shall call Φ .We use the intuition which we have acquired from knowing the solutions of the Schrödinger equations for simpler systems. For stationary states, ψ can be taken to be a real number. The lowest state always has no nodes (except for the exclusion principle, which does not operate here). Therefore Φ is everywhere positive. It is symmetrical, that is, Φ depends only on where atoms are, not on which is which. The energy $V(\boldsymbol{R})$ of interaction of two helium atoms, as worked out by Slater and Kirkwood,[14] for example, consists of a very weak attraction at large distances, but a powerful repulsion inside of 2.7 Å (see Fig. 2). The atoms in liquid helium at the normal density have

a volume of 45 cubic Ångströms each so they are not tightly squeezed together. If one wishes a rough approximation, consider the atoms as impenetrable spheres of 2.7 Å diameter, and forget the attraction, whose effect is, after all, mainly just to hold the liquid together at the normal density even if the external pressure falls to zero. Then configurations of atoms in which some overlap each other, that is, are closer than about 2.7 Å, are of very small amplitude. In the most likely configurations the atoms are well spaced. As for a particle in a box whose wave function bows highest in the center and falls gradually to zero at the walls, we may imagine the amplitude highest for good separation and falling toward zero if a pair of atoms approach

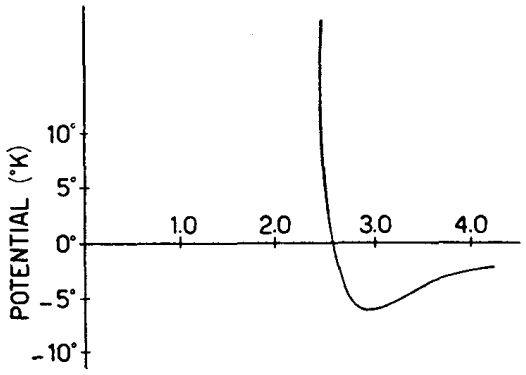


Fig. 2. The potential of interaction of two helium atoms as function of their separation as worked out from quantum mechanics by Slater and Kirkwood.

too closely. Our structure is a liquid, as a consequence of the zero-point energy, so that no particular lattice arrangement is strongly preferred. All configurations for which the spacing is ample have high probability. We can get from one arrangement to another without ever crossing a forbidden configuration of overlapping atoms because of the large spacing (cube root of atomic volume is 3.6 Å). Although not crystalline, there is a little local order induced by the tendency of atoms to stay apart, so that X-ray or neutron scattering experiments show a structure very similar to that of other simple liquids like liquid argon.

For the configurations of high amplitude the density is fairly uniform, at least until we look over such small volumes that we can see the fine grain atomic structure. If the density in a region is raised the atoms come closer together so that the "bow" on the wave function

which occurs as one atom is moved from contact with a neighbour on one side to one of the other side, is confined to a smaller space. The increased curvature represents increased kinetic energy and it is not as likely to find a configuration in which such an energy barrier is penetrated. As a matter of fact, this feature is easily analyzed quantitatively. Long range density fluctuations are sound waves. The rise of energy on compression is described by the compressibility coefficient, or equivalently by the speed of sound. Classically, standing density waves oscillating as a normal mode behave as an harmonic oscillator. Likewise, in quantum mechanics these are quantum oscillators and have zero point motions, although the most likely configuration is that of uniform constant density. The wave function for the zero-point motion of an oscillator is a gaussian so that the amplitude Φ for a given kind of density fluctuation falls off exponentially with the square of the fluctuation. To summarize, the ground state function is large for any configuration in which the atoms are well spaced from one another at nearly constant average density. If falls off if these conditions are violated.

Next we turn to the excited states. Right away one obvious excitation is that of the standing sound wave. If the classical frequency is ω the quantum excitation energy of such a mode is $\hbar\omega$. Usually one prefers by linear combinations to make states of running waves, or phonons. If the wave number is k, the energy is $\hbar kc$ if c is the sound velocity.

We may readily obtain the wave function for such a phonon excitation. If the density is $\varrho(\boldsymbol{R})$ the classical normal coordinate going with such a mode is

$$q_{\boldsymbol{k}} = \int \varrho(\boldsymbol{R}) \exp(i\boldsymbol{k} \cdot \boldsymbol{R})\, d^3\boldsymbol{R} \tag{8}$$

Quantum mechanically for an oscillator the wave function for the ground state is a gaussian, and the first excited state is just the coordinate times this gaussian (the first hermite polynomial $H_1(x)$ is just x). Hence the wave function is

$$\psi_{\text{phonon}} = q_{\boldsymbol{k}}\Phi \tag{9}$$

if Φ is the ground state wave function of the system, which we have described in the preceding paragraphs. We have not bothered to normalize our function. The liquid consists of many atoms so if $\boldsymbol{R}_i$ is the position of the i^{th}, the density in any configuration is

$$\varrho(\boldsymbol{R}) = \Sigma_i \delta(\boldsymbol{R} - \boldsymbol{R}_i) \tag{10}$$

the sum extending over all the atoms. Putting this in (8) and then (9) we find

$$\psi_{\text{phonon}} = (\Sigma_i \exp. i\boldsymbol{k} \cdot \boldsymbol{R}_i)\, \Phi \tag{11}$$

This is valid if the wave length $(2\pi/k)$ is much larger than the atomic spacing, for then our description by compressional waves is adequate. The state energy is $\hbar kc$. Since $\hbar\boldsymbol{k}$ is the momentum $\boldsymbol{p}$ of the state, this means $E = pc$. Since the wave length can be very long this energy can be exceedingly low.

The central problem is to see why no states other than these phonons can have such low energies. We try to construct the wave function ψ of an excitation which should be as low in energy as possible and yet not represent a phonon. We must associate a number which may now be positive or negative with each configuration. In fact, since ψ must be orthogonal to the ground state Φ which is everywhere plus, ψ must be plus for half the configurations and minus for the other half. Furthermore, ψ must be orthogonal to all the phonon states. This simply means that ψ must vary from plus to minus for changes in the configurations which do not appreciably alter the large scale density. Configurations can alter without variation of mean density by simply stirring the atoms about. Of course, since ψ must represent as low an energy as possible we must give low amplitude to configurations in which atoms seriously overlap, just as in the ground state Φ.

The function ψ takes on its maximum positive value for some configuration of the atoms. Let us call this configuration A, and the particular locations of the atoms α-positions. We said that the α-positions must be well spaced so that the atoms do not overlap, and further that they are, on a large scale, at roughly uniform density. Equally, call configuration B, with atomic positions β, that for which ψ has its largest negative value. Now we want B to be as different as possible from A. We want it to require as much readjustment over as long distances as possible to change A to B. Otherwise ψ changes too rapidly and easily from plus to minus, our wave function has a high gradient, and the energy of the state is not as low as possible.

Try to arrange things so that A *requires* a large displacement to be turned into B. At first you might suppose it is easy. For example (see Fig. 3) in A take some atom in the left side of the box containing the liquid and move it way over to the other side of the vessel, and

call the resulting configuration B. One objection to this is that an atom is moved from one side to another, so a hole remains at the left and an extra atom is at the right. This represents a density variation. To avoid this we may imagine that another atom has been moved at the same time from right to left, and the various holes and tight squeezes have been ironed out by some minor adjustments of several of the neighbouring atoms. This movement of two atoms each a distance of the size of the vessel, one from left to right and the other from right to left, is certainly a long displacement, so B and A are very different. But they are not.

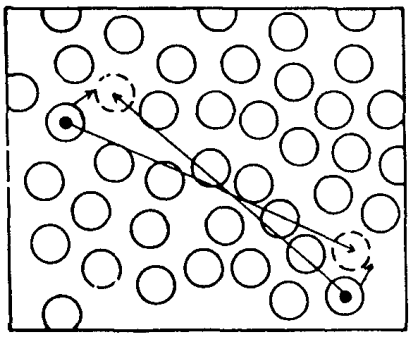


Fig. 3. Two configurations (solid and dotted) that result from large displacements (long arrows) of the atoms, can actually be accomplished by much smaller adjustments (short arrows) because of the identity of the atoms

The atoms must be considered as identical, the amplitude must not depend on which atom is which. One cannot allow ψ to change if one simply permutes atoms. The long displacements can be accomplished in two steps. In the first step permute the atoms you wish to move to those α-positions closest to the ultimate position they are to occupy in the final configuration B. This step does not change ψ because all the atoms are still in the same configuration of α-positions. Then the change to the B configuration is made by small readjustments, no atom moving more than half the atomic separation. In this minor motion ψ must change quickly from plus to minus and the energy cannot be low. For the reason that the wave function is unchanged by permutation of the atoms it is imposible to get a B configuration very far from the A configuration. No very low energy excitations can appear (other than phonons) at all.

In the phonon case we consider configurations in which, as ψ changes sign, the density distribution changes. A change in density cannot be accomplished by permuting atoms. That is why the Bose statistics does not affect phonon states. But it leaves them isolated as the lowest states of the system, so the specific heat approaches zero as T approaches zero according to Debyes T^3 law. This is the key argument for the understanding of the properties of liquid helium. It is given in

somewhat more detail in reference 7. Since it is a negative argument, attempting to prove that a low energy state does *not* exist, it is difficult to convey conviction in a few words. The reader should try to invent wave functions of low energy for himself. After a few attempts he will see much more clearly what we have tried to explain here.

5. Rotons

The qualitative argument is complete in itself. Nevertheless it is gratifying that it may be pushed even further to produce a quantitative estimate of the energy of these other excited states. We give only a summary of the considerations here (see reference 8 for details). We try to clarify our picture of the wave function ψ, until we can write a mathematical expression for it. This expression put into the energy integral $\int \psi^* H \psi d^N V / \int \psi^* \psi d^N V$ will give us an estimate for the energy.

As we said, in order to get the energy as low as possible we wish the gradients of ψ to be small. Therefore the configuration B (where ψ is maximum negative) must be as far as possible from configuration A. Yet we noted that no β-site is more than half the atomic spacing from an α-site. The two configurations are generally nearly the same. They are furthest from each other if as many atoms as possible must be moved. That is accomplished when, as illustrated in Fig. 4, all the β-sites are between α-sites, so every atom must move. To completely specify ψ, of course, we must give its value for all configurations, not only for A and B when all atoms are on α-sites, or all on β-sites. The lowest energy results if the transition from plus to minus (hence A to B) is as gradual as possible. First for configurations in which each of the atoms is either on an α or on a β-position, this is most naturally accomplished if ψ is proportional to the number on α-sites minus the number on β-sites. This difference passes smoothly from plus to minus. It can be expressed mathematically this way: Consider a function, $f(\boldsymbol{R})$, of position, which is $+1$ if $\boldsymbol{R}$ is at an α-

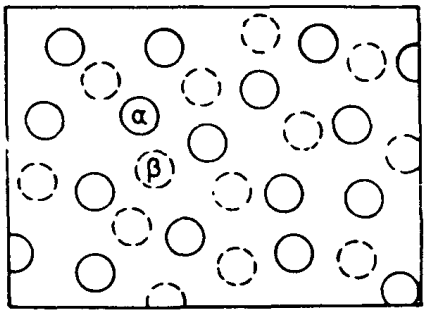


Fig. 4. The excited state wave function must be positive for one configuration, solid circles (α-positions) and negative for another. They are separated as far as possible if the negative configuration leaves no atom unmoved, dotted circles (β-positions)

site and — 1 if $\boldsymbol{R}$ is at a β-site. Then $\Sigma_i f(\boldsymbol{R}_i)$ summed on all the atoms is just the desired number on α-sites minus number on β-sites. For intermediate positions ψ will vary as smoothly as possible if $f(\boldsymbol{R})$ is taken to vary in some smooth way between its extreme values of + 1 and — 1, which it takes on at α and β-sites. This suggests that we take ψ to be of the form

$$\psi = \Sigma_i f(\boldsymbol{R}_i).$$

But this is incomplete for we tacitly assumed that in all the configurations the atoms did not overlap, the mean density did not vary very much and so on, just as in the ground state. This feature can be taken into account if we take instead

$$\psi_{\text{roton}} = \Sigma_i f(\boldsymbol{R}_i)\, \Phi \tag{12}$$

where Φ is the ground state function. [15] Then ψ will fall rapidly if the atoms overlap, etc. We actually do not know what the function $f(\boldsymbol{R})$ is but we expect it to vary rapidly, so that if expanded in a fourier integral the dominant wave lengths would be the atomic spacing.

According to the variational principle the best wave function is that which minimizes the energy integral. In this way, by variation of $f(\boldsymbol{R})$ it is readily found (see reference 8) that the minimum results if the function is

$$f(\boldsymbol{R}) = \exp(i\boldsymbol{k} \cdot \boldsymbol{R}) \tag{13}$$

and that the corresponding energy is

$$E(k) = \hbar^2 k^2 / 2mS(k) \tag{14}$$

where m is the atomic mass. The function $S(k)$ is the form factor for the scattering of neutrons from the liquid. That is, it is the Fourier transform of the function $p(R)$ which gives the probability per unit volume of finding an atom at a distance R from a given atom in the liquid in the ground state.

The local partial order of the liquid in the ground state shows up as in other liquids as a ring in the diffraction pattern (of neutrons, or X-rays). That is to say, there is a maximum in the function $S(k)$, which occurs when k represents a wavelength near the nearest neighbour spacing. The maximum in $S(k)$ represents a minimum in $E(k)$ here. This confirms the expectation that the low excitation would have wave numbers in this vicinity.

The state (12) and (13) has the momentum $\boldsymbol{p} = \hbar\boldsymbol{k}$. Ordinarily not

every value of a parameter in a wave function has significance in the variational method. But states of different momenta are orthogonal, and the energies (14) are significant not only for k near the minimum, but also in the neighbourhood of this value. The range of values for which (14) is useful is limited only by the range for which (12) can be expected to be a good wave function. For small k, (12) is identical to the wave function (11) representing phonon excitation, and (14) can be shown to give $\hbar kc$ in that region. Therefore the expression should be reasonable not only for k near the reciprocal atomic spacing, but for low k as well. It predicts a spectrum at first linear in p $(= \hbar k)$ then falling to a minimum, just as anticipated by Landau, and in agreement with experiment.

The curve $S(k)$ taken from neutron data of Henshaw and Hurst [16], or from the X-ray scattering data of Reekie [17] agree. The $E(k)$ which results is shown in Fig. 1 by the dashed line. The general behavior and minimum are clearly shown.

The actual value of the energy at the minimum is twice too high to agree with the experimental value (solid line) for Δ. The theoretical value lies above the true value, as it should according to the variational principle.

The inaccuracy of the wave function (12) prevents us from giving a complete description of what the roton wave function must look like. The function (12) does not satisfy the conservation of current. It appears as though a more accurate function would represent a current distribution large and unidirectional in one region, with a field of return currents surrounding it, somewhat in the nature of a smoke ring. These and other arguments suggest a trial function of the type

$$\psi = \Sigma_i \exp i\boldsymbol{k} \,.\, \boldsymbol{R}_i \,.\, \exp i \Sigma_j g(\boldsymbol{R}_j - \boldsymbol{R}_i)\, \Phi \qquad (15)$$

with the g, representing the back flow, to be determined. It is very hard to perform the integrals required in the variation problem with (15), so it has not been verified whether (15) represents a substantial improvement.

One way to understand the low energy for k near the reciprocal atomic spacing is this. One might consider these as sound waves of very short wave length. To obtain a density variation of long wave length is hard. To make the compression work must be done against opposing forces. For wave lengths closer to atomic spacing, however,

such density variations are easier to arrange. In fact, one can create variations of wave length equal to the atomic spacing simply by arranging the atoms, doing no appreciable work against repulsions, the energy being purely kinetic $\hbar^2k^2/2m$. Actually the energy is even lower ($S(k)$ at maximum is 1.3) for there is a positive tendency in the liquid to have such variations; if some atoms are correctly arranged the others are more likely to be also satisfactory because of the local order. Therefore the energy does not continue to rise as $\hbar kc$ but falls lower for wave lengths near the atomic separation.

It is easy to verify that these excitations behave in just the way that has been assumed in developing the statistical mechanics and the two fluid model. To represent a state with two excitations, say with momenta $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ one has the approximate wave function

$$\psi = (\Sigma_i \exp i\boldsymbol{k}_1 \cdot \boldsymbol{R}_i)\ (\Sigma_j \exp (i\boldsymbol{k}_2 \cdot \boldsymbol{R}_j))\ \Phi$$

and so on. Since the order of the factors is irrelevant this is the same state if $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ are reversed.The excitations obey the Bose statistics. In moving fluid the energy of the excitations can be shown to be (3).

6. Irrotational Superfluid Flow

So far we have only described the wave function for states representing internal excitation. We turn next to a description of the wave function which represents the state of the fluid when macroscopically we say it is flowing. We will assume that the flow velocity does not vary appreciably over distances of the order of an atomic spacing.

It is not dfficult to represent by wave functions states which represent the motion of the superfluid. Suppose the system is at absolute zero so there are no excitations. If the entire system moves forward as a body, since the center of gravity coordinate can be separated out from Schrödinger's equation, the wave function is

$$\psi = (\exp i\boldsymbol{k} \cdot (\Sigma_i \boldsymbol{R}_i))\ \Phi$$

where $N\hbar\boldsymbol{k}$ is the momentum of the system, if there are N atoms. In case the velocity is not uniform we can construct a wave function somewhat as follows: If the velocity varies only slowly from place to place, those atoms temporarily in a macroscopic region where the velocity is, say, $\boldsymbol{v}$ must surely have a wave function very much the same as though the liquid in the region were isolated and moving at a uniform velocity. This suggests that the phase contains a term

$\hbar^{-1}m\boldsymbol{v} \cdot \Sigma_i \boldsymbol{R}_i$, the sum being taken only over those atoms in the region. Other regions where $\boldsymbol{v}$ differs make similar contributions to the phase so the total phase is $m\Sigma_i \boldsymbol{v}(\boldsymbol{R}_i) \cdot \boldsymbol{R}_i$ where $\boldsymbol{v}(\boldsymbol{R})$ is the velocity at $\boldsymbol{R}$. This suggests a wave function of the form

$$\psi_{\text{flow}} = [\exp i(\Sigma_i s(\boldsymbol{R}_i))]\, \Phi \tag{16}$$

where $s(\boldsymbol{R})$ is a function which varies only very little over distances as small as the atomic spacing. We have suggested that it is $\hbar^{-1}m\boldsymbol{v}(\boldsymbol{R}) \cdot \boldsymbol{R}$, but as is usual for waves whose wave length varies with position, the momentum is the gradient of the phase, not the coefficient of $\boldsymbol{R}$. Thus (16) does represent the helium flowing, but the velocity is given by

$$\boldsymbol{v} = \hbar m^{-1} \nabla s. \tag{17}$$

It is readily verified that the current density is $\varrho_0 \boldsymbol{v}$, and the energy (from the variational integral) is $\frac{1}{2}\varrho_0 \boldsymbol{v}^2$, as expected classically. There is no change in density, as in (16) we have not allowed these small effects to be represented.

If excitations exist in the moving fluid the wave function is (16) multiplied by the factor $\Sigma_i f(\boldsymbol{R}_i)$ in (12). The excitation energy turns out to be (3) as expected, interpreting $\boldsymbol{v}$ as the superfluid velocity $\boldsymbol{v}_s$.

Equation (17) implies that the motion is irrotational, that is, $\nabla \times \boldsymbol{v}_s = 0$. In a simply connected region this has only one solution for given motion of the boundaries. For fixed boundaries it is $\boldsymbol{v}_s = 0$. In a multiply connected region the situation is different. Since $\nabla \times \boldsymbol{v}_s = 0$, the circulation about any closed curve which can be shrunk to a point is zero. On the other hand, in the case of a toroidal region, if the curve encloses the hole the circulation need not vanish. Although the wave function must be single valued, s may be of the nature of the azimuthal angle, increasing by 2π, or a multiple thereof if one goes around the hole. That is, for a circuit enclosing a hole (into which liquid may not freely flow) the circulation must be an integral multiple n of a quantized unit $2\pi\hbar/m$,

$$\oint \boldsymbol{v}_s \cdot d\boldsymbol{s} = 2\pi\hbar m^{-1} \cdot n = 2\pi n \cdot 1.5 \times 10^{-4}\ \text{cm}^2/\text{sec} \tag{18}$$

These states do not influence the previous statistical mechanical argument. There are too few of them. The velocity may be considered as a macroscopic variable, such as density. For a macroscopic torus even the lowest of the states given by (18) is very much higher than

a roton energy Δ. Thus if the torus area is A, radius R, the mass moving is $mA \cdot (2\pi R)/d^3$ where d^3 is the atomic volume. It moves at velocity given by $v_s \cdot 2\pi R = 2\pi\hbar m^{-1}$, from (18), so the kinetic energy is $(\hbar^2/2md^2) \cdot (2\pi A/Rd)$. The factor $\hbar^2/2md^2$ is an energy of the order of a roton, but the second factor is very large, being the torus dimension over the atomic spacing. Incidentally the total angular momentum is $\hbar$ per atom.

If the fluid must flow irrotationally, at first sight, it cannot lose energy, unless it is moving very rapidly. This has been pointed out by Landau. If a body of fluid is moving at velocity v, and loses a small energy δE, it must do so (to keep the flow irrotational) by the entire fluid changing its velocity. Let the change in v be δv. If M is the effective fluid mass the momentum change δp is $M\delta v$ and $\delta E = Mv\delta v = v\delta p$. Now this energy loss must go into heat; that is, into internal excitations of rotons. But if the momentum transferred to excitations is δp the energy cannot be small. It must be at least about $(\delta p/p_0)\Delta$ where Δ and p_0 are the energy and momentum of an individual roton. That is, δE must be at least $(\Delta/p_0)\delta p$ and energy cannot be lost unless v exceeds Δ/p_0, about 70 meters per second. (More accurately v must be high enough that a line drawn from the origin at slope v can cut the $E(p)$ *vs* p curve). This suggests the reason for the frictionless flow of superfluid. But we have proved too much, for in actuality the resistance sets in at velocities a few hundred times smaller.

The only way that gross slowing down can occur for lower velocities is for small parts of the fluid to stop or slow down without the entire fluid having to slow down at once. That is, energy loss must be accompanied by flow which is not irrotational; that is, flow which involves local circulation. To understand such effects we must add a new element to our picture of phonons, rotons and potential flow. These are the quantized vortex lines suggested by Onsager.[1] We proceed to describe them.

7. Rotation of the Superfluid

The problem which now faces us is to extend (16) so that we can also represent states for which $\nabla \times \boldsymbol{v}$ does not vanish, or at least where there is circulation in the superfluid. We analyze the situation at absolute zero for simplicity. We must present ourselves a problem in which such circulation is necessary and try to find the lowest energy state. The situation first considered by the author was the slip-stream

between two regions of fluid moving at different velocity, but it is easier to arrive at the result by considering the problem of helium with high angular momentum in a cylindrical vessel. Suppose, for example, the helium at absolute zero is initially under such pressure that it is solid and is set into rotation, then the pressure is released so that it liquifies. What is the final state of the helium? We ask then for the lowest state of a quantity of helium which has a definite, macroscopically high, total angular momentum.

For a system of given angular momentum the kinetic energy is least if the angular velocity ω is a constant throughout the liquid. This motion is not rotation free for $\nabla \times \boldsymbol{v} = 2\omega$. But it is very difficult for helium to manage a state of local circulation. In fact, without high excitation energy, local circulation is impossible. At first one might find it hard to see why the liquid cannot simply rotate as a rigid body. The energy is then low. But a liquid is not a rigid body. A part of it can turn independently of the whole. In a rough way of speaking the liquid may be thought of as made up of many quasi-independent units of nearly atomic dimensions. Any motion of the body can be compounded of motions of the tiny parts. But to set any small part into a rotational state requires a high energy because the moment of inertia is so small. If only a limited energy is available nearly all the "parts" must be frozen out in their ground states.That is, nearly everywhere the local angular momentum is zero, i.e., $\nabla \times \boldsymbol{v}_s = 0$. It takes energy to create circulation and, furthermore, we can expect this circulation not to be distributed uniformly throughout the fluid. The rigid body type of rotation where $\nabla \times \boldsymbol{v}_s \neq 0$ everywhere is not possible, or if at all, only with an enormous expenditure of energy, an expenditure far higher than that gained by the uniform distribution of angular velocity.

Another possibility that suggests itself is that the liquid, if the angular momentum is high, is not free of excitations like rotons and phonons even though the temperature is at absolute zero. These excitations could carry the angular momentum. That is, in the language of the two fluid model, perhaps there is at $T = 0$ a mixture of superfluid and normal fluid, with the superfluid component not rotating, and with the normal fluid carrying all of the angular momentum. The energy to maintain the normal fluid being sustained by the fact that if less normal fluid were present, for given angular momentum the kinetic energy would have to be larger. This turns out, for vessels of

centimeter dimensions turning at about one radian per second, to be a state of nearly 10^4 times the energy of a rigid body rotating at the same angular velocity. Surely nature can find some lower state for the helium.

We know (see 18) that if there is a hole in the liquid, circulation can exist. Therefore another solution suggests itself. The liquid circulates around a hole with constant circulation as in a free vortex (familiar from rotation of water around an emptying drain). The velocity varies inversely as the radius, rising to such heights near the center as to be able to maintain the hole free of liquid by centrifugal force. Such a solution would be easy to verify in a striking manner by looking at the surface of the liquid. Instead of the usual parabola it would be the curve of the surface of a free vortex. The energy is still quite a bit higher than the rigid body case, because the velocity instead of being distributed proportionally to the radius, actually falls as the radius increases. Nevertheless it is orders of magnitude below the mixture of normal fluid suggested above.

However, this is still not the lowest possible energy state, and the striking experiment will not succeed. To show this we construct a lower state. Suppose that the liquid has not only one vortex at the center, but several vortices. For example, suppose beside the central one there were a number distributed about the circle of radius $R/2$, half that of the vessel R, and all turning the same way. Viewed grossly this is like a vortex sheet so the tangential velocity can jump as we pass from inside $R/2$ to outside. Then the velocity can be arranged a little more like the linear curve by two sections, each of which is a $1/r$ curve. The gain in energy resulting from this improved distribution may more than compensate the energy needed to make the additional holes (and, further, the central vortex need not now be so large and energetic).

Continuing in this way with ever more vortices it soon becomes apparent that the energy can always be reduced if more vortices form. However there is a limit. Due to the quantization (18) of the vortex strength the smallest vortex has circulation $2\pi\hbar m^{-1}$. The lowest energy results if a large number of minimum strength vortex lines (which we shall call unit lines) form throughout the fluid at nearly uniform density. The lines are all parallel to the axis of rotation. Since the curl of the velocity is the circulation per unit area, and the curl is 2ω, there will be

$$2m\omega/2\pi\hbar = 2.1 \times 10^3\omega \text{ lines per cm}^2 \tag{19}$$

with ω in radians per second. For $\omega = 1$ rad per second the lines are about 0.2 mm apart so that the velocity distribution is practically uniform.

Such weak lines will not form actual macroscopic holes. In fact, if one neglects atomic structure and assumes a classical continuous liquid with surface tension, a unit line makes a hole opposed by surface tension which figures out to be only 0.4 Å in radians. That means that there is no real hole in the liquid. Around such a unit line, for example a straight one along the z-axis, the wave function off the axis is roughly

$$\psi = (\exp i\Sigma_i\varphi_i)\,\Phi \tag{20}$$

where φ_i is the angle about the z-axis. This does not hold close to the axis. On the axis exp $i\varphi$ is meaningless, and close to it has enormous gradients. A particle on the axis cannot have angular momentum, yet (20) implies that each atom has angular momentum $\hbar$, nor can there be exceptions because the Bose statistics implies that they are equivalent. Therefore a more accurate expression than (20) would be this expression multiplied by a factor which is unity except if any one of the atoms comes very close to the axis, in which case it falls rapidly to zero. The density of fluid falls to zero on the axis. This is the remnant of the classical hole. Actually quantum mechanically the line will not remain perfectly straight in one spot but will have some zero point motion of wandering and waving to and fro.

It is not hard to get a reasonable estimate of the energy contained in these lines. First consider an isolated unit line along the axis of a cylinder of length L, radius b. The velocity at radius r is $\hbar/mr$ and if ϱ_o is the fluid density in atoms per cc ($\varrho_o = 1/45$ Å^3) the kinetic energy is the integral

$$K.\,E. = \tfrac{1}{2}\int \varrho_o m(\hbar/mr)^2 \,.\, 2\pi r\, dr \,.\, L.$$

The upper limit of the integral is b. It diverges at the lower limit, but within about the atom spacing the velocity formula is meaningless. Furthermore, inside this radius some of the energy is potential, required to keep to density down near the axis (that is, to make the partial "hole"). Therefore the energy needed to form such a line, per unit length, is

$$\begin{aligned}\text{Line energy per unit length} &= \varrho_o\pi\hbar^2m^{-1}ln(b/a)\\ &= 10^{-8}ln(b/a) \text{ ergs/cm.}\end{aligned} \tag{21}$$

Here a is a length of order of the atomic spacing. Its exact determination would require solving the difficult quantum mechanical problem. In almost all applications the ratio b/a will be very large, and the logarithm large enough to be insensitive to the exact value of a. For this reason we will not attempt a detailed evaluation, but simply choose a to be close to the atomic spacing. We arbitrarily take $a =$ 4.0 Å. In more complicated geometrical situations the lower limit will be the same, but the upper limit b will be some other characteristic dimension of the apparatus, or more usually the spacing between vortex lines, etc. It can be found by integrating the velocity distribution as determined for the given distriubution of singular vortex lines.

For a cylinder of liquid rotating at angular velocity $\omega = 1$ rad/sec the vortices are about 0.02 cm apart. This is 0.5×10^6 times a if $a = 4$ Å, so we can take the $ln(b/a)$ in this case to be about $ln(0.5 \times 10^6)$ or 14. Neglecting the variation of this logarithm with ω we find for the energy of all of the lines:

Total line energy per unit volume $= \varrho_0 \omega \hbar \, . \, ln(b/a)$

where we have estimated $ln(b/a)$ as 14. The ratio of this to the kinetic energy for a rigid body is $4\hbar m^{-1} R^{-2} \omega^{-1} ln(b/a)$ if the cylinder radius is R. For $R = 1$ cm, $\omega = 1$ rad/sec this ratio is 10^{-2}. For macroscopic laboratory dimensions the excess energy to form the lines is small. They would form if rotating solid helium is melted by releasing the pressure, the angular velocity distribution would differ imperceptibly from uniformity, and the surface should appear parabolic.

It is not self-evident that there is no state of appreciably lower energy, and that the energy of the rotating liquid is correctly estimated. This subject has not yet been analyzed any more deeply than is reported here. Therefore this part of the paper is not on as firm a foundation as the rest. We must therefore still consider it conjectural whether the considerations on rotational flow reported here are actually correct. It is interesting that all the conclusions were arrived at independently by the author without knowledge of Onsager's previous work (with which they are in exact concordance).

8. Properties of Vortex Lines

In a situation more general than uniform rotation, in which the curl of the velocity is not constant, we can imagine a similar situation. We have a situation instantaneously with many vortex lines. Some are

closed on themselves in rings, and others terminate with their ends on the fluid boundaries. Viewed from a continuum approximation in which atomic structure is neglected, a velocity $\boldsymbol{v}_s$ can be defined at every point. The curl of this is zero everywhere, except at one of the vortex lines where it is infinite. These lines are real quantized vortex lines. The circulation around a small circuit surrounding only one line is $2\pi\hbar m^{-1}$. The lines have a sense depending on the direction of rotation. The circulation about any curve whatosever is given by

$$\oint \boldsymbol{v}_s \,.\, ds = 2\pi\hbar m^{-1} n$$

where n is always an integer, being the net number of lines linked by the circuit, account being taken to the sign of each.

If $\nabla \times \boldsymbol{v}_s$ is averaged over a large enough region that many lines are included, the number of lines per cm^2 must be at least $<\nabla \times \boldsymbol{v}_s> m/2\pi\hbar$ and the energy of these lines per unit volume is at least

$$\tfrac{1}{2} |<\nabla \times \boldsymbol{v}_s>| \varrho_0 \hbar . ln(b/a) \tag{22}$$

where b is the spacing between lines, $1/b^2 = <\nabla \times \boldsymbol{v}_s> m/2\pi\hbar$. This shows that in our liquid it takes energy to create circulation. Actually in real, complex situations the energy might exceed greatly the value in (22). There may be great complex activity with many lines twisting and turning so that several lines of opposite senses are close together. In this case, the case of developed turbulence, the number of lines present may be bigger than the average $\nabla \times \boldsymbol{v}_s$ would indicate. (Probably in such a case it would be hard to define the average $\nabla \times \boldsymbol{v}_s$ because the result may depend on the size of the region over which the average is taken).

The discussion of the rotating cylinder of liquid with which we introduced the lines is rather special. We shall try to give a more complete and general description of the state of the superfluid with circulation. We continue to study the case at absolute zero. Let us try to characterize the state of a fluid in which we desire two things (which, it will turn out, are mutually incompatable). We want (a) the liquid to be flowing with a velocity $\boldsymbol{v}_s$ which is a smooth function of position without singularities (on a scale of distances large compared to atomic dimensions) and (b) we want $\nabla \times \boldsymbol{v}_s$ not to vanish.

Suppose the liquid in an element of volume ΔV (large compared to the atomic volume) is moving at velocity $\boldsymbol{v}$. Then as we have seen the wave function should depend on the position of the atoms, if they are

within ΔV as exp $(im\boldsymbol{v} \cdot \Sigma_i \boldsymbol{R}_i)$ Φ. That is, if a number of atoms in the region are displaced, each by $\Delta \boldsymbol{R}_i$ from one allowed (by Φ) configuration to another allowed one, the main effect is that the wave function must change phase by

$$\Sigma_i (m\boldsymbol{v} \cdot \Delta \boldsymbol{R}_i)\hbar^{-1} \tag{23}$$

This can also be seen in another way. If a region of fluid can be considered to have a velocity $\boldsymbol{v}$ it has a momentum density $\varrho_0 m\boldsymbol{v}$. It is characteristic of momentum in quantum mechanics, that if the center of mass is changed the wave function changes phase by an amount proportional to the momentum and to the displacement of the center of mass. Now if the atoms are displaced by $\Delta \boldsymbol{R}_i$ the center of mass moves so the phase change (23) results. This is true at least if the displacement makes no other important change in the wave function. We will suppose that both before and after the displacement the atoms are well spaced and there are no gross density fluctuations, etc. such that in case the liquid were not in motion both configurations would have essentially the same amplitude.

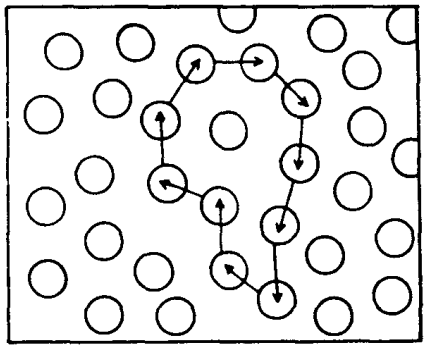


Fig. 5. The wave function must not change as a result of a permutation. If all the atoms are displaced around a ring, as shown, the phase change must be a multiple of 2π

The same argument goes for atoms in other regions, etc. so the phase shifts accumulate to a sum in (23) over displacements of atoms all over the liquid, if $\boldsymbol{v}$ is now considered as a function of $\boldsymbol{R}_i$. The displacements $\Delta \boldsymbol{R}$ must be small compared to the distances over which $\boldsymbol{v}$ varies. We shall apply the formula in a case in which $\Delta \boldsymbol{R}$ is the separation between atoms.

Select, in a given configuration, a very long closed chain of atoms each of which is a nearest neighbor of the next in line (see Fig. 5). The last should have the first as nearest neighbor. The chain may consist of very large numbers of atoms and may even be so long that it passes through regions of varying velocity. Consider a displacement of each atom to its nearest neighbor next in line. The wave function cannot change, for it is simply a special permutation of the atoms. Further we will suppose that if all the displacements are made together

a little at a time, each intermediate configuration is allowed. This sliding of the chain along itself is not prevented by potential barriers, especially if we allow small temporary displacement of other atoms adjacent to the ring to permit passage in tight places. In the final configuration all atoms have returned to their original positions, except those of the ring which have moved one over. We suppose, because of the ease in which the displacement can be made that we can assume the wave function does not vanish for any intermediate position during the displacement. Then its phase shift is given by (23), but this must represent no change in the wave function. It is therefore necessarily an integral multiple of 2π. We conclude that

$$\oint \boldsymbol{v}_s \cdot d\boldsymbol{s} = 2\pi\hbar m^{-1} n \tag{24}$$

where n is an integer and the integral is taken over any path which goes from one atom to the next neighbor, etc. If $\boldsymbol{v}_s$ is now assumed continuous at an atomic scale, the path can be smoothed out to any continuous curve. Of course, it is impossible that (24) holds for all continuous paths if n is an integer (depending on the path) if $\boldsymbol{v}_s$ is free of singularities and continuous unless $n = 0$ (in a simply connected region). Because any path can be deformed continuously into an infinitesimal path, the left side changing continuously to zero. The right side cannot change continuously so it must be zero for all paths. Likewise for a toroidal region n must be the same for all paths which surround the hole.

We see therefore that $\boldsymbol{v}_s$ cannot be continuous if we are to have circulation. There must be places where $\boldsymbol{v}_s$ is discontinuous, and places in the fluid where a displacement of an atom to its neighbor may not be possible without passing through a node in wave function. In the neighborhood of such a node the probability of finding an atom is reduced. This decrease in density requires energy to maintain it. We shall therefore try to arrange conditions so that such places are as infrequent as possible. Under those conditions, for nearly every conceivable ring of atoms the atoms can be moved over to the next adjacent atom without the wave function vanishing. Its phase change must be a multiple of 2π. If two adjacent rings have a phase change which is different, differing by 2π say, then between them somewhere must lie a very small ring of three or four atoms for which the circulation is $2\pi\hbar/m$. For example, suppose for a certain ring A the phase is zero, but for a nearby ring B it is 2π. Then shift ring B by a few

atoms at a time until it gets as close to ring A as possible, but still has phase shift 2π. Likewise shift A until it is as close to B as possible but so that it has still shift 0. Then A and B will contain many atoms in common and only differ by a few, as illustrated in Fig. 6. Then

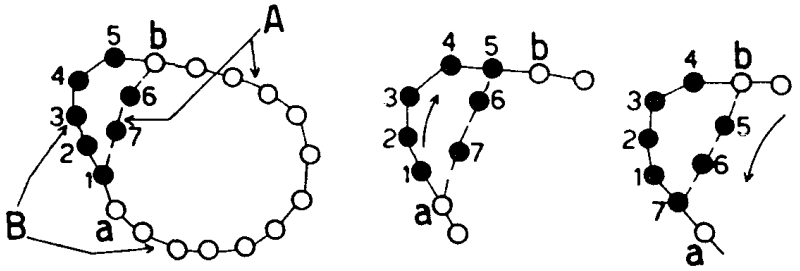


Fig. 6. A displacement along ring B followed by a reverse displacement along an adjacent ring A with many atoms in common is equivalent to a displacement around the ring C, indicated by black circles (except for an inconsequential permutation 5–b).

consider a permutation consisting of shifting B forward, then shifting A backward. It is readily verified that this change is the same as a shift of atoms around the very small ring C consisting of those parts of A and B which are not common, plus one of the common atoms. But the change in phase is 2π when B shifts and 0 when A shifts back, so that it must be 2π for the very small ring C. * This represents a highly concentrated angular momentum. Somewhere in the middle of ring C is a nodal point. It is readily appreciated geometrically that these nodal points must essentially form lines through the fluid. They are quantized vortex lines. It must be admitted that this argument is far from complete. We should consider states in which the location of the vortex line is uncertain, that is, a superposition of states with various locations for the line. Such a state would have a lower energy. Possibly we make a serious error in imagining that the velocity can be defined right up to atomic distance from the axes, if this axis itself does not have a definite location. Onsager has remarked, in private communication, on the possibility that these quantum effects might lower the energy to such an extent that the logarithm in (24) should

* The change in phase cannot be determined only from the initial and final configuration, but requires a description of the amplitude for intermediate configurations as well. Therefore this argument is not complete unless it is also assumed that partial rotations of C consisting of displacements of less than one atom spacing can also be roughly imitated by partial displacements of B and A.

be absent. At any rate, although our energy estimates may be incorrect, quantized vortex lines probably exist. We continue our discussion of the consequences of this assumption.

On a large scale according to the theorem of Helmholtz, vorticity moves with the fluid in such a way that the strength of a vortex filament remains constant. This means that if the fluid drifts the lines drift with it, maintaining their quantized strength. This is true, at least, if no forces act directly on the vortex line. In general the force per unit length on a vortex line equals the density, $\varrho_0 m$, times the vector cross product of the circulation, $2\pi\hbar m^{-1}$, and the velocity of fluid where the vortex is.

9. Critical Velocity and Flow Resistance

We next turn to the role such vortex lines may play in the resistance to flow found at sufficiently high velocities. We have suggested that this resistance cannot be understood in terms of a direct creation of rotons, the superfluid otherwise being in perfect flow. Let us consider what would happen if liquid is flowing out of an orifice, or tube, into a reservoir of fluid at rest. In Fig. 7 is illustrated the distribution of flow for irrotational motion. A very high velocity develops near the corners and large accelerations develop there. An ordinary fluid, such as water, flows in a complicated manner such as illustrated in Fig. 8 (a few moments after flow starts). The water shoots out straight into the nearly still fluid in the reservoir, forming a vortex sheet, which is unstable and curls around, eventually in an extremely complex manner. Let us see how helium might try to imitate some of the features of the type of flow illustrated in Fig. 8. Just for rough orientation and estimate suppose the fluid tries to go out in a jet, let us say at first of the same width and velocity as in the tube. Take the case that the tube is a long slot perpendicular to the paper, and the flow is roughly two dimensional.

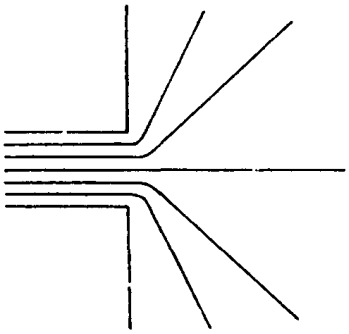


Fig. 7. Ideal potential flow from an orfice

Then circulation is implied for the velocity is v in the jet and 0 outside. This requires the formation of vortex lines, perhaps as illustrated in Fig. 9. The spacing is x and if this is small compared to d,

the slot width, the velocity distribution is roughly uniform inside the the jet. Taking a line integral along the jet for unit distance, and returning outside the jet, the circulation is v so the number of lines per centimeter is

$$\frac{1}{x} = v/2\pi\hbar m^{-1}.$$

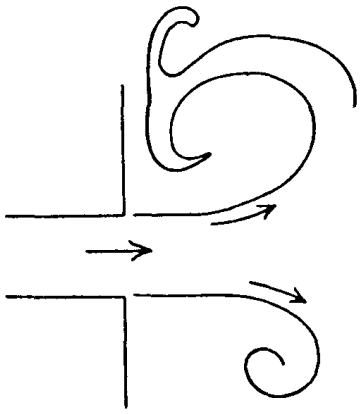


Fig. 8. Real flow from an orfice for ordinary liquids, producing an unstable vortex sheet

It takes energy to form these lines. If there is not enough kinetic energy in the fluid to supply the energy to make the lines, no resistance will appear. Once the lines can be formed they are, in a manner we shall soon discuss, ultimately dissipated as heat and a resistance appears. Let us see what order of critical velocity we would estimate in this way. The lines move out at the velocity of the fluid at their own location, which is $v/2$. Another way to see the necessity for this is to realize that as the fluid passes from inside to outside the pipe vorticity is created, so new lines must continually come rolling out of the ends of the orifice. In our case v/x lines are created per second. The energy needed to create these is (per unit length of slot)

$$\frac{v^2}{2\pi\hbar m^{-1}} \varrho_o \pi\hbar^2 m^{-1} ln(d/a)$$

where the argument in the logarithm is only approximate. The total kinetic energy available per cc of fluid is $\frac{m\varrho_o v^2}{2}$, so that per second $vd\,\frac{m\varrho_o v^2}{2}$ is available. If we define v_o as that velocity for which the energy available is just large enough to create the vortices we find

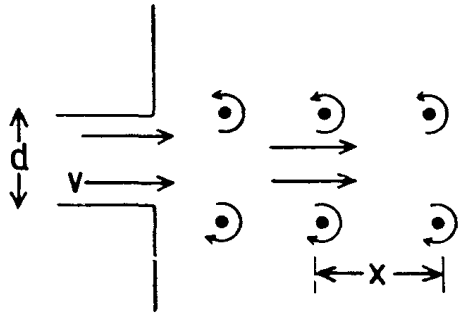


Fig. 9. Idealization of supposed vortex rings formed when superfluid helium issues at high speed from an orfice.

$$v_o d = \hbar m^{-1} ln(d/a).$$

For example, for a slit of width $d = 10^{-5}$ cm, (which is about three

times the width of a Rollin film at a height of 1 cm) this gives $v_0 =$ 100 cm/sec, if $ln(d/a)$ is taken as 6. This is somewhat higher than the critical velocities observed. The calculation is only meant as an estimate because the actual situation must be complicated. For one thing, near the critical condition x comes out about $3d$ so our picture of a uniform jet is poor. Further, the velocity in the jet must of course be reduced as a result of the energy needed to form the vortex line. Actually probably the situation near the critical point must be very complicated and irregular. The flow for short momentary periods may be much like Fig. 7 but irregularly vortex lines peel off of the edges of the slit, probably starting at one point along the slit and progressing to other places, or perhaps if the hole is circular, one or two vortex lines is fed out continuously in a form roughly like a helix. It is predicted that very close to the critical velocity when loss just begins, the resistance will be irregular and show fluctuations. These fluctuations are very small however and would be hard to detect. Possibly some sound may be generated by the flow irregularities. It is difficult to estimate its intensity. When helium is driven, just above critical velocity, through an emery powder superleak, some noise should be generated as the various vortex lines suddenly form and pass into the stream. The irregularities are a result of the unpredictable quantum transitions between states of no vortex line and one with a section of line.

Another possible source of vortex lines is the contact between flowing liquid and the walls. It is not necessary that all the loss occurs at the exit end of the tube. The walls of the pipe are irregular. Vortex lines may be created inside the pipe also.

It is difficult to go beyond this order of magnitude calculation in describing the conditions controlling the production of vortex lines. For example, if one studies the example given there are serious difficulties. As a particular vortex line leaves the end of the tube there are very great forces trying to pull it back resulting from its image in the tube wall. Let us imagine a line a distance b above the wall in a tube in which the velocity of flow is v_0. It is readily shown that the forces acting on the line are these. First a force pulling away from the surface of strength $2\pi\hbar\varrho_0 v_0$. Second, from the image, an attraction to the wall of strength $\pi\hbar^2\varrho_0/mb$. A vortex line responds to forces by moving through the liquid to reduce the net force to zero. In this case it would drift upstream if the attraction is highest. But a vortex line

will interact with the wall, especially at its ends which go into the wall surface. Suppose this results in a frictional force which keeps the line from moving upstream. Then the response is to move closer to the wall. The vortex only moves away from the wall if $2\pi\hbar\varrho_0 v$ exceeds $\pi\hbar^2\varrho_0/mb$. Even if v is 100 cm/sec this requires b to exceed 10^{-6} cm or 20 atomic spacings. We might expect a vortex line to fluctuate away from the surface by a few atomic diameters. But how can we expect to penetrate the enormous potential barrier, to create a line so far away from the surface that the flow velocity can pull it further out and create eventual vorticity and energy loss?

More likely a line gets started somehow and has its ends tied on the wall. Then the forces of the fluid on the rest of the line cause it to wander about in such a way that more and more vortex line is fed out. It is not necessary to create bodily at one instant a complete section of line. For example, for the case of liquid issuing from a tube perhaps the vortex lines are helices with contact points at the edge of the hole which turn round and round while the helix moves outward. Similar things could happen inside tubes. If the tubes are very narrow the line will hit the other surface easily and be attracted by the walls. It can never get very far from a wall. Even if started somehow it will fall back into the tube walls unless the velocity v_0 suffices to keep it in the stream. Therefore the smallest tubes have the highest critical velocities.

10. Turbulence

The patterns of vortex lines which we have studied are well known to be unstable. In the case of the rotating cylinder this is not true if the cylindrical vessel containing the helium rotates also. But if the container is stopped the situation is altered. There are forces between the wall and vortex lines. (This is because the fluid density is altered near the line axis, so the interaction with the wall is not the same as the average for the rest of the helium). The lines at the outside drag past the stationary wall and as a result get distorted from their original vertical line position. This twists others, etc. Lines fall into the wall and others twist about each other in a complex way. It would be interesting to study this experimentally, to see how fast, and in what manner, the liquid eventually slows down.

In ordinary fluids flowing rapidly and with very low viscosity the phenomena of turbulence sets in. A motion involving vorticity is

unstable. The vortex lines twist about in an ever more complex fashion, increasing their length at the expense of the kinetic energy of the main stream. That is, if a liquid is flowing at a uniform velocity and a vortex line is started somewhere upstream, this line is twisted into a long complex tangle further down stream. To the uniform velocity is added a complex irregular velocity field. The energy for this is supplied by pressure head.

We may imagine that similar things happen in the helium. Except for distances of a few ångstroms from the core of the vortex, the laws obeyed are those of classical hydrodynamics. A single line playing out from points in the wall upstream (both ends of the line terminate on the wall, of course) can soon fill the tube with a tangle of line. The energy needed to form the extra length of line is supplied by a pressure head. (The force that the pressure head exerts on the lines acts eventually on the walls through the interaction of the lines with the walls). The resistance to flow somewhat above critical velocity must be the analogue in superfluid helium of turbulence, and a close analogue at that.

There are some ways, however, in which the two cases differ. In a classical fluid there is a thin boundary layer near the wall of the pipe in which viscosity controls the situation. In this boundary layer there is a large vorticity, but it escapes into the stream to be amplified, only from the edge of the layer. Inside it is damped by viscosity. As the stream velocity falls the boundary layer thickens, for the amplification is less and the damping overpowers it ever further from the wall. Below a critical velocity the turbulence ceases altogether and the flow is laminar, but with vorticity, the viscosity keeping the vorticity from amplifying itself. That is, viscosity is the mechanism which determines whether vorticity will be amplified or not, and therefore whether turbulence is produced. If the viscosity goes to zero as a limit (and no other physical phenomena are added) a classical ideal liquid would exhibit turbulence at any velocity, no matter how small.

Superfluid helium is an ideal fluid of zero viscosity. It does not exhibit turbulence at low velocity because of another, quantum mechanical, effect. The vorticity is quantized and cannot begin at as low amount as desired. One must supply energy enough to get the first one or two vortex lines started before the amplification process of turbulence can take over. There will not be a boundary layer with a structure analogous to that in classical flow (although near the walls the flow will

be somewhat different because of the dragging forces between the moving vortex lines and the wall).

In a classical fluid ,if the turbulent stream empties into a reservoir, the turbulent motion continues for a while, but as a result of the viscosity, it gradually slows up and dies out, the energy appearing eventually as heat.

What happens to a turbulent mass of superfluid left to itself? If there is normal fluid present the rotons and phonons will collide with the vortex lines and take energy from them, gradually turning this energy gain into more rotons and phonons (as a result of collisions among rotons the number of these may change). But an interesting question arises if the experiment is imagined at absolute zero. What can eventually become of the kinetic energy of the vortex lines?

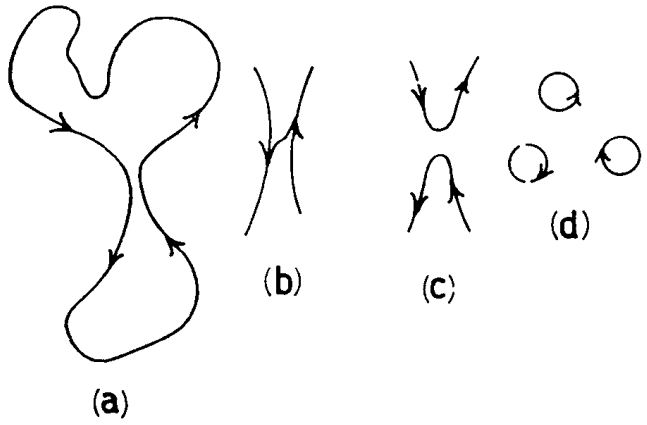


Fig. 10. A vortex ring (a) can break up into smaller rings if the transition between states (b) and (c) is allowed when the separation of vortex lines becomes of atomic dimensions. The eventual small rings (d) may be identical to rotons.

One possibility that suggests itself is this. Consider a large distorted ring vortex (Fig. 10a). If, in a place, two oppositely directed sections of line approach closely, the situation is unstable, and the lines twist about each other in a complicated fashion, eventually coming very close; in places, nearly within an atomic spacing. Consider two such lines (Fig. 10b). With a small rearrangement, the lines (which are under tension) may snap together and join connections a new way to form two loops (Fig. 10c). Energy released this way goes into further twisting and winding of the new loops. This continues until the single loop has become chopped into a very large number of small loops (Fig. 10d).

The smallest ring vortex that can exist must have a radius about half the atomic spacing. Let us guess that this is in fact a roton. Then all the energy of the vortex will eventually end by forming large numbers of rotons, that is, heat. Perhaps eventually it will be easier to understand the details of the complete transformation of organized flow energy into disorganized heat energy for liquid helium than for other substances.

11. Rotons as Ring Vortices

It is not unreasonable to guess that these smallest vortices are rotons. The velocity distribution around a roton, which is found by analytic means (ref. 8) is similar to that around a vortex ring. It is quite reasonable that a vortex ring can be only so small. To increase the curvature of a vortex line beyond that of radius roughly a may take energy. Let us imagine a roton to be the circular quantized vortex of lowest energy. A large circular vortex has (from (21)) energy $E = 2\pi R \cdot \frac{\pi\hbar^2}{m} \varrho_o ln\, R/a$. It carries momentum $p = \pi R^2 \cdot 2\pi\hbar\varrho_o$. This momentum is that of a roton, p_o, if $R = 2.2$ Å. The energy is the right order (it corresponds to replacing ln by 1.6).

One might object that such a vortex drifts through the fluid, at velocity $v = (\hbar/2mR) ln\, R/a$, so one would expect rotons not to have a zero group velocity. Actually this drift, of a large vortex, has its seat in the force tending to shrink the vortex to decrease the energy of the line. The response to the radially directed force is a perpendicular motion. It is analagous to the ornery response of a gyroscope. In fact, if a vortex line were a thin flexible mechanical tube with inertia, and were started with zero forward motion, it would first fall in a bit and then move forward in a halting fashion, like the nutation of a gyroscopic, or the motion of an electron in crossed magnetic and electric fields. In a roton we imagine that the forces tending to contract the ring are already opposed by a kind of stiffness of the ring. It is already as small as possible. No drift motion results. In fact forward drift would expand it and raise the energy, while reverse drift would try to compress it to smaller size, again raising the energy. The lowest energy is at zero drift velocity. We may notice in passing that they can only drift in a direction perpendicular to their plane, that is, along, or opposite, the direction of the momentum. This agrees with a property derived for rotons from their energy-momentum relation

(1), that the group velocity $\delta E(p)/\delta p$ is in the direction of the momentum, (or opposite).

Having travelled so far making one unverified conjecture upon another we may have strayed very far from the truth. However imprudent it may be, there is one further observation we would like to make. A detailed picture is not available which describes physically just what goes on as the transition is approached from below. The free energy expression arising from (6) does not of itself describe the transition. The transition occurs when the number of rotons is very large. Some sort of interaction may occur between them, or there may be some limitation to the degrees of freedom. [18] There is no doubt that it is the analogue of the transition in the ideal gas, but it would be nice if we could get a less mathematical and formal description of the events. Of the following I am not sure, but it does seem to be an interesting possibility.

If rotons are the smallest ring vortices, and those of lowest energy, Δ, then there are states of higher energy corresponding to larger rings. For example, a ring of twice the diameter may have twice the energy more or less. The relative number of these will be expected to be very low, however. Since Δ is 9.6°K, at the transition $\exp(-\Delta/kT)$ is 10^{-2}, so very few larger vortices will be expected in equilibrium. Certainly none whose length is 10^2 or 10^3 atoms! This neglects an important feature, however. For a long line there are an enormous number of shapes and orientations available. Such a line is not infinitely flexible, of course, for the curvature cannot well exceed a^{-1}. It may be likened to a chain of a finite number of links. Adding one link requires an energy ε, say of order Δ, but increases the number of orientations by some factor, asymptotically, say s. In equilibrium then, the number of chains of $n+1$ links is a factor $s \exp(-\varepsilon/kT)$ times the number with n links. For low temperatures this is less than unity. No long chains are important. The excitations consist of rotons and a few other rings of slightly larger size. As the temperature rises, however, there comes a time when the factor $s \exp(-\varepsilon/kT)$ exceeds unity. Then suddenly the rings of very largest length are of importance. The state with one vortex line (or a very few) which winds and winds throughout the liquid like a near approximation to a Jordon curve, is no longer of negligible weight. The superfluid is pierced through and through with vortex line. We are describing the disorder of Helium I. At first the curve doesn't make full use of all of its orientations and higher

entropy. But as the temperature rises a little more it squeezes into the last corners and pockets of superfluid until it has no more degrees of flexibility available. The specific heat curve drops off from the transition to a smooth curve and the memory of the possiblity the helium can exhibit quantum properties in a unique way is lost in the perfusion of states and in disorder, as it is for more usual liquids.

REFERENCES

[1] L. Onsager, Nuov. Cim., **6**, Supp., **2**, 249, (1949).
[2] F. London, Nature, **141**, 643, (1938).
[3] A. Einstein, Ber. Berl. Akad., **3**, (1925).
[4] F. London, Phys. Rev., **54**, 947, (1938).
[5] R. P. Feynman, Phys. Rev., **91**, 1291, (1953).
[6] G. V. Chester, Phys. Rev., **94**, 246, (1954).
[7] R. P. Feynman, Phys. Rev., **91**, 1301, (1953).
[8] R. P. Feynman, Phys. Rev., **94**, 262, (1954).
[9] L. Tisza, Phys. Rev., 72, 838, (1947).
[10] L. Landau, J. Phys. U.S.S.R., **5**, 71, (1941).
[11] R. B. Dingle, Advances in Phys., **1**, 112, (1952).
[12] D. de Klerk, R. P. Hudson and J. R. Pellam, Phys. Rev., **93**, 28, (1954).
[13] L. Landau and I. Khalatnikov, J. Exper. Theor. Phys. (U.S.S.R.), **19**, 637, 709, (1949).
[14] J. C. Slater and J. G. Kirkwood, Phys. Rev., **37**, 682, (1932).
[15] A. Bijl, Leiden, Comm. Suppl., **90a**, Physica, **7**, 869, (1940).
[16] D. G. Henshaw and D. G. Hurst, Phys. Rev., **91**, 1222, (1953).
[17] J. Reekie and T. S. Hutchison, Phys. Rev., **92**, 827, (1953). We are grateful to Dr. Reekie for sending us the curve of S(k) derived from his data.
[18] R. P. Feynman, Phys. Rev. **94**, 262, (1954), especially the discussion preceding expression (28). There is a typographical error there. The relation (28) should read as an inequality, the left side not greater than the right side.

THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 102, No. 5 JUNE 1, 1956

Energy Spectrum of the Excitations in Liquid Helium*

R. P. FEYNMAN AND MICHAEL COHEN
California Institute of Technology, Pasadena, California

(Received February 27, 1956)

A wave function previously used to represent an excitation (phonon or roton) in liquid helium, inserted into a variational principle for the energy, gave an energy-momentum curve having the qualitative shape suggested by Landau; but the value computed for the minimum energy Δ of a roton was 19.1°K, while thermodynamic data require $\Delta=9.6$°K. A new wave function is proposed here. The new value computed for Δ is 11.5°K. Qualitatively, the wave function suggests that the roton is a kind of quantum-mechanical analog of a microscopic vortex ring, of diameter about equal to the atomic spacing. A forward motion of single atoms through the center of the ring is accompanied by a dipole distribution of returning flow far from the ring.

In the computation both the two-atom and three-atom correlation functions appear. The former is known from x-rays, while for the latter the Kirkwood approximation of a product of three two-atom correlation functions is used. A method is developed to estimate and correct for most of the error caused by this approximation, so that the residual uncertainty due to this source is negligible.

1. INTRODUCTION

LIQUID helium undergoes a thermodynamic transition at 2.19°K. Below this temperature, many of the properties of the liquid are explained by Tisza's phenomonological two-fluid model. Landau realized that the macroscopic properties of the liquid would resemble those of a mixture of two fluids, provided that a certain form is assumed for the energy-momentum curve of the elementary excitations in the liquid. Starting from first principles, one of the authors has recently computed an energy-momentum curve which is based on certain ideas about the nature of the wave functions representing the excitations.[1] The shape of the curve is in qualitative agreement with Landau's, but some serious quantitative discrepancies exist. The ideas of III are pursued further here, and a more complicated wave function is constructed to represent an excitation. The energy-momentum curve computed with this wave function will prove to be in better agreement with Landau's. In addition to the actual computations, we discuss some approximate methods which may be useful in other work of this sort.

2. LANDAU'S SPECTRUM

The energy momentum curve proposed by Landau[2,3] rises linearly for small p, passes through a maximum, falls to a minimum, and rises steeply for large p. (See Fig. 6.) The excitations in the linear region are quantized sound waves (phonons); their energy, measured relative to the ground-state energy, is

$$E(p)=cp, \tag{1}$$

where c is the velocity of sound (240 m/sec). Near its minimum, the spectrum can be approximated by a parabola,

$$E(p)=\Delta+(p-p_0)^2/2\mu. \tag{2}$$

Landau believed that excitations in this region represent some kind of rotation of the fluid, and called them "rotons." In the present paper we are led to the picture of a roton as the closest quantum-mechanical analog of a smoke ring. The remaining portions of the spectrum are not excited at low temperatures. For $T<2$°K the phonons and rotons are present in sufficiently small

* This paper is based on a Ph.D. dissertation submitted to California Institute of Technology in November, 1955. A preliminary report of this work has appeared [R. P. Feynman and Michael Cohen, Progr. Theoret. Phys. Japan 14, 261 (1955)].

[1] R. P. Feynman, Phys. Rev. 94, 262 (1954), henceforth referred to as III.

[2] L. Landau, J. Phys. (U.S.S.R.) 5, 71 (1941).

[3] L. Landau, J. Phys. (U.S.S.R.) 11, 91 (1947).

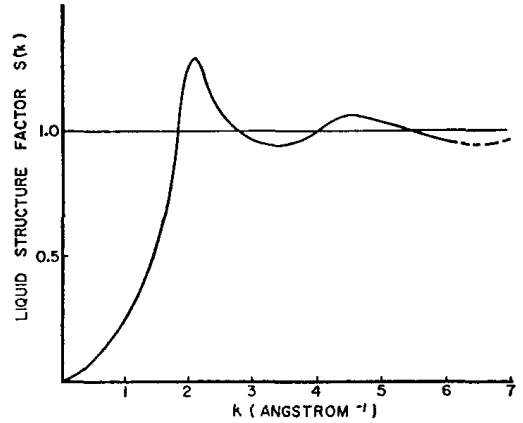


FIG. 1. The liquid structure factor $S(k)$, based on the x-ray scattering data of Reekie and Hutchison. The principal maximum corresponds to a wavelength equal to the nearest neighbor distance in helium. Appendix A describes modifications we have made in the data.

numbers to allow them to be treated for thermodynamic purposes as noninteracting. The thermodynamic functions can then be computed; Landau fitted the available (1947) data on specific heat and second-sound velocity with the values

$$\Delta/\kappa=9.6°\text{K}, \quad p_0/\hbar=1.95\ \text{A}^{-1}, \quad \mu=0.77\ m_{\text{He}}. \tag{3}$$

More recent measurements[4] of the velocity of second sound down to $T=0.015°$K suggest the values

$$\Delta/\kappa=9.6°\text{K}, \quad p_0/\hbar=2.30\ \text{A}^{-1}, \quad \mu=0.40\ m_{\text{He}}, \tag{4}$$

although the values (3) also fit the data quite well. The value of Δ/κ is quite well determined[5] by the thermodynamic data, since it enters formulas in the form $\exp(-\Delta/\kappa T)$. The differences between (3) and (4) are probably a fair measure of the uncertainty in our knowledge of p_0 and μ.

The reasoning which led Landau to the general form of the spectrum, and his method of deducing the two-fluid picture from the spectrum, will not be reviewed here. He did not attempt to compute the values of Δ, p_0, and μ from first principles.

3. A SIMPLE WAVE FUNCTION FOR THE EXCITATIONS

In III a wave function of the form $\psi=\varphi\sum f(\mathbf{r}_i)$ is proposed to represent an excitation. The physical reasons for this wave function will not be reviewed here. The sum runs over all the atoms in the liquid, and φ is the wave function for the liquid in its ground state. The requirement that ψ be an eigenfunction of the total momentum operator[6] $\mathbf{P}=-i\hbar\sum\nabla_i$ corresponding to the eigenvalue $\hbar\mathbf{k}$ implies that $f(r)=e^{i\mathbf{k}\cdot\mathbf{r}}$, and thus

$$\psi=\varphi\sum\exp(i\mathbf{k}\cdot\mathbf{r}_i). \tag{5}$$

Substitution of (5) into the variational principle

$$E=\mathcal{E}/\mathcal{I}, \tag{6}$$

where

$$\mathcal{E}=\int\psi^*H\psi d^N\mathbf{r} \tag{7}$$

and

$$\mathcal{I}=\int\psi^*\psi d^N\mathbf{r}, \tag{8}$$

gives an upper limit[7] for the energy of the lowest excitation of momentum $\hbar\mathbf{k}$. The result is

$$E(k)=\hbar^2k^2/2mS(k), \tag{9}$$

where $S(k)$ is the Fourier transform of the zero-temperature two-atom correlation function $p(r)$,

$$S(k)=\int e^{i\mathbf{k}\cdot\mathbf{r}}p(r)d\mathbf{r}. \tag{10}$$

The data which we have used for $S(k)$ are given in Fig. 1 and are essentially those obtained from x-ray diffraction by Reekie and Hutchison.[8,9] Figure 2 is the corresponding curve for $p(r)$. $S(k)$ exhibits a sharp maximum near $k=2\ \text{A}^{-1}$, which corresponds to a wavelength equal to the nearest neighbor distance in the liquid. Accordingly, the spectrum (9) exhibits a minimum at approximately the correct wave number (see curve B of Fig. 6). It is shown in III that the wave function (5) is exact for phonons (small k) and that the limiting form of (9) is $E(k)=\hbar ck$. The occurrence of a minimum at $k=2\ \text{A}^{-1}$ is in qualitative agreement with Landau's predictions, but the value of Δ/κ computed from (9) is 19.1°K, which is twice the value given by experiment.

4. ARGUMENTS FOR A NEW WAVE FUNCTION

The excitation (5) can be localized in a definite region by the formation of a wave packet. If $h(\mathbf{r})$ is a function, like a Gaussian, which is peaked about some

[4] deKlerk, Hudson, and Pellam, Phys. Rev. 93, 28 (1954).

[5] Dr. J. R. Pellam (private communication) estimates the uncertainty in Δ/κ to be less than 0.2°.

[6] If the liquid were confined to a box of side L, with fixed walls, then the walls could absorb momentum and the energy eigenstates would not be momentum eigenstates. Instead, we control the density by requiring the wave function to be periodic in all variables with period L. With this boundary condition, P commutes with H and the energy eigenstates can also be taken as momentum eigenstates.

[7] Eigenfunctions of $\mathbf{P}$ belonging to different eigenvalues $\hbar\mathbf{k}$ are orthogonal. Hence, for different $\mathbf{k}$, the trial functions (5) are orthogonal to each other and also to the true wave functions which minimize (6). The entire spectrum $E(k)$ therefore lies above the true spectrum. In footnote 3 of III it is mentioned that the wave function $\varphi\exp(iN^{-1}\mathbf{k}\cdot\sum\mathbf{r}_i)$, which represents translational motion of the whole liquid, has momentum $\hbar\mathbf{k}$ and energy $\hbar^2k^2/2mN$, which is certainly lower than any energy we shall compute from (5). The periodic boundary condition, however, rules out such states unless k is as large as $N^{\frac{1}{3}}$.

[8] J. Reekie and T. S. Hutchison, Phys. Rev. 92, 827 (1953). Their paper contains a curve for $r^2p(r)$, but does not include their data on $S(k)$. We are indebted to Dr. Reekie for sending us the data, which are now generally available in reference 9. Appendix A contains a discussion of some changes which we have made in the data.

[9] L. Goldstein and J. Reekie, Phys. Rev. 98, 857 (1955).

location in the liquid and falls off smoothly in a distance large compared with $2\pi/k$ but small compared with the size of the box, then the wave function

$$\psi=\sum h(\mathbf{r}_i)\exp(i\mathbf{k}\cdot\mathbf{r}_i)\varphi \tag{11}$$

represents a localized excitation. The packet will spread in time, and will drift with velocity $\hbar^{-1}\nabla_k E(k)$. The current and density associated with (11) were computed in III. The number density is very close to the average density ρ_0, even in the region of the packet, and the current at a point $\mathbf{a}$ is $\mathbf{j}(\mathbf{a})=\hbar m^{-1}|h(\mathbf{a})|^2$. The wave function (11) therefore leads to the picture of a total current $\hbar\mathbf{k}m^{-1}$ (assume $\int|h(\mathbf{a})|^2d\mathbf{a}=1$) distributed over a small region and having everywhere the same direction, with no appreciable change in the number density anywhere. Such a picture clearly cannot represent anything like a stationary state, since in a stationary state the current is divergence-free and there would necessarily be a return flow directed oppositely to $\mathbf{k}$.

One way to incorporate such a backflow into (11) is to multiply the wave function by $\exp[i\sum g(\mathbf{r}_i)]$, obtaining

$$\psi=\varphi\exp[i\sum g(\mathbf{r}_i)]\sum h(\mathbf{r}_i)\exp(i\mathbf{k}\cdot\mathbf{r}_i). \tag{12}$$

Application of the velocity operator $-i\hbar m^{-1}\nabla_i$ shows that, in addition to whatever velocity it had in (11), the ith atom now has an extra velocity $\hbar m^{-1}\nabla g(\mathbf{r}_i)$. Substitution of (12) into (6) shows that the energy is minimized if $g(r)$ satisfies

$$\nabla\cdot(\mathbf{j}+\hbar\rho_0 m^{-1}\nabla g)=0, \tag{13}$$

where $\mathbf{j}$ is the current computed from the old wave function (11). Furthermore, the current arising from (12) is $\mathbf{J}=\mathbf{j}+\hbar\rho_0 m^{-1}\nabla g$, so that (13) states that the best backflow g is that which conserves current. Equation (13), with the physically reasonable boundary condition that $g\to 0$ as $r\to\infty$, completely determines g. At large distances g has the form of the velocity potential for dipole flow, namely $\boldsymbol{\mu}\cdot\mathbf{r}/r^3$; the dipole moment is

$$\boldsymbol{\mu}=m(4\pi\hbar\rho_0)^{-1}\int\mathbf{a}[\nabla\cdot\mathbf{j}(\mathbf{a})]d\mathbf{a}$$

$$=-m(4\pi\hbar\rho_0)^{-1}\int\mathbf{j}(\mathbf{a})d\mathbf{a}=-(4\pi\rho_0)^{-1}\mathbf{k}. \tag{14}$$

The negative sign of $\boldsymbol{\mu}$ indicates that the direction of the backflow is opposite to that of $\mathbf{k}$, as expected. We shall refer to the value of $\boldsymbol{\mu}$ given by (14) as the "classical value," since it is derivable from the equation of conservation of current plus the assumption that the momentum density is equal to the current density times the mass. The energy of (12) is only slightly lower than that of (11), the difference being of the order of the reciprocal of the volume of the packet. The important point to be learned from this calculation is that the energy is lowered if the wave function conserves current.

The solution of a somewhat different problem tends to support the same idea. Suppose we want to find the energy of a state in which a foreign atom moves through the liquid with momentum $\hbar\mathbf{k}$. The foreign atom is assumed to have the same mass as He atoms, and also to experience the same forces, but it is not subject to Bose statistics. The energy of this situation was computed in III. The simplest trial wave function is

$$\psi=\varphi\exp(i\mathbf{k}\cdot\mathbf{r}_A); \tag{15}$$

$\mathbf{r}_A$ is the coordinate of the foreign atom, and φ is the wave function for the ground state of the entire system (which is the same as if all the atoms obeyed Bose statistics). With this wave function, Eq. (6) gives $E=\hbar^2k^2/2m$. A possible way of lowering the energy would be to let the neighbors of the moving atom execute some pattern of flow around it, leaving space in front of it and filling in the hole behind it. Some such pattern is already contained in (15), since the ground-state wave function φ prohibits atoms from overlapping. But in the ground state, readjustments are made by pushing a few immediate neighbors of the foreign atom out of the way; these neighbors are crowded into less than their usual volumes, causing (15) to have a high kinetic energy. If, instead, room could be made for the moving atom by the simultaneous motion of many atoms, each being crowded only slightly, the kinetic energy of the state would be lower. In fact, there is no reason why the crowding cannot be eliminated entirely, since the amount of matter in the system remains constant. Roughly speaking, the requirement of no crowding means that the current is divergence-free, and the no-crowding argument shows physically why it is energetically advantageous to conserve current. The argument is vague, however, and the exact form of the backflow will be determined by more accurate methods.

A wave function of momentum $\hbar\mathbf{k}$ which includes a

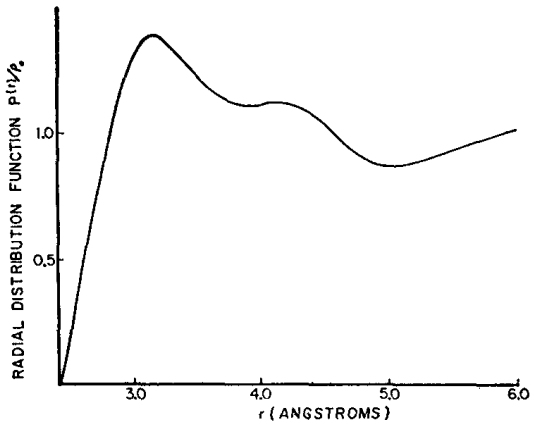


FIG. 2. The radial distribution function $p(r)$, based on the data of Reekie and Hutchison.

pattern of backflow around the foreign atom is

$$\psi = \varphi \exp(i\mathbf{k}\cdot\mathbf{r}_A) \exp[i \sum_{i\neq A} g(\mathbf{r}_i - \mathbf{r}_A)]. \quad (16)$$

When (16) is substituted into the expression (6) for the energy, minimization of E leads to a differential equation which determines $g(\mathbf{r})$. The solution at large $\mathbf{r}$ is proportional to $\mathbf{k}\cdot\mathbf{r}/r^3$. Accurate numerical solution is simple, but uncertain because of uncertainty in the values of $p(r)$; since (6) is a variational principle, we may take $g(\mathbf{r}) = A\mathbf{k}\cdot\mathbf{r}/r^3$. Substitution of (16) into (6) gives

$$E = (2m)^{-1}\hbar^2k^2[1 + I_1A + (I_4 + I_{5a})A^2], \quad (17)$$

where I_1 and I_4 are integrals defined by Eq. (25) and I_{5a} is an integral defined by (57). The integrals are evaluated further on; only the answer interests us here. Equation (17) becomes

$$E = (2m)^{-1}\hbar^2k^2(1 + 0.186A + 0.0246A^2), \quad (18)$$

with A measured in A^3. The energy is minimum when $A = -3.8$ A^3; the "classical" value predicted by (14) is $A = -(4\pi\rho_0)^{-1} = -3.6$ A^3. The close agreement of the two values seems to indicate that the reduction in energy is due to the physical effects we have mentioned, and is not simply the result of allowing an extra degree of freedom in the wave function. The improved value for the energy is[10]

$$E = 0.648\hbar^2k^2/2m. \quad (19)$$

Since the wave function (5) for a phonon or roton is just what would result for symmetrizing (15), one might hope to lower the energy of (5) by adding terms to represent a backflow around each moving atom. The resulting wave function would be the symmetrization of (16), i.e.,

$$\psi = \varphi \sum \exp(i\mathbf{k}\cdot\mathbf{r}_i) \exp[i \sum_{j\neq i} g(\mathbf{r}_{ji})]. \quad (20)$$

For large k, when this wave function is substituted into the energy and normalization integrals, there is little interference between terms with different i; the energy is therefore given by (19) and is a definite improvement over (9). For small k, (20) cannot lead to a lower energy than (5), because (5) is exact for phonons. At intermediate k, one might thus expect to lower the energy by a factor between 1.00 and 0.65. In fact, we do better than this.

The attempt to find the function $g(\mathbf{r})$ which gives the lowest energy when (20) is substituted into (6) leads to an intractable equation. We therefore take $g(\mathbf{r}) = A\mathbf{k}\cdot\mathbf{r}/r^3$, where A will be chosen to minimize the energy. The difficulty of handling integrals which involve e^{ig} leads one to consider the possibility of replacing $\exp(i\sum g)$ by $1 + i\sum g$. The average value of $\sum_{j\neq i} g(\mathbf{r}_{ji})$ is $\int_0^\infty p(r)g(\mathbf{r})d\mathbf{r}$, which is zero because $g(\mathbf{r})$ is an odd function. The mean square value of $\sum_{j\neq i} g(\mathbf{r}_{ji})$ is $k^2A^2I_3$, where the integral I_3 is defined and evaluated further on. With the classical value for A (which is close to the optimal value throughout the interesting range of k) the root-mean-square value of $\sum g(\mathbf{r}_{ji})$ turns out to be $0.25k$, where k is measured in inverse angstroms. Even with $k = 2$ A^{-1}, replacement of $\exp(i\sum g)$ by $1 + i\sum g$ is not unreasonable,[11] and we shall work with the wave function

$$\psi = \varphi \sum_i \exp(i\mathbf{k}\cdot\mathbf{r}_i)[1 + i \sum_{j\neq i} g(\mathbf{r}_{ji})], \quad (21)$$

where $g(\mathbf{r}) = A\mathbf{k}\cdot\mathbf{r}/r^3$. This wave function is still an eigenfunction of the total momentum operator $\mathbf{P}$, with eigenvalue $\hbar\mathbf{k}$.

The roton state represented by the function (21) can be described roughly classically as a vortex ring of such small radius that only one atom can pass through the center. Outside the ring there is a slow drift of atoms returning for another passage through the ring. There are at least three ways that the classical picture is modified. (1) The momentum of atoms passing through the center cannot be made smaller because the wave function must return to its original value when, after one moves through, another stands in its old place. The wavelength must be the atomic spacing. (2) The ring does not drift forward as a large smoke ring, because as it is as small as possible there is no force tending to shrink it; such a force in a classical ring is balanced as a consequence of the forward drift. (3) The location of the ring is not definable. In typical quantum-mechanical fashion the lowest energy state corresponds to superposition of amplitude to find the ring anywhere in the liquid. The energy is *less* than the kinetic energy $\hbar^2k_0^2/2m$ of one atom with momentum $\hbar k_0$ because there is a correlated motion of many atoms moving together so the effective inertia is higher (the energy Δ/κ corresponds to 2.5 atoms moving together at total momentum $\hbar\mathbf{k}_0$).

[10] This is somewhat higher than the value obtained in III, where a rather inaccurate approximation was used for I_4. With the new value for I_4 we find that the effective mass of a He^3 atom moving through He^4 is 5.0 atomic mass units, instead of 5.8. In the calculation it is assumed that the distribution of atoms around the He^3 atom is the same as that around an He^4 atom. The higher zero-point motion of the lighter atom actually pushes its neighbors further away. This effect will increase the mass, but probably by only a small fraction of a mass unit.

[11] Of course, since a trial function is a free choice, it would be mathematically legitimate to insert $1 + i\sum g$ into the variational principle (6) even if $\sum g$ were not small, but there would be little physical reason to expect a good answer. If $\exp[i\sum g(\mathbf{r}_i - \mathbf{r}_A)]$ is replaced by $1 + i\sum g(\mathbf{r}_i - \mathbf{r}_A)$ in the foreign atom problem, the resulting integrals are among the ones defined and evaluated further on. The energy is given by

$$E = \frac{\hbar^2k^2}{2m} \frac{1 + 0.186A + (0.0217 + 0.0049k^2)A^2}{1 + 0.0049k^2A^2}.$$

When $k = 2$ A^{-1}, the fraction has the minimum value 0.689, which is 6% higher than the value given by (19). The associated value of A is -3.4 A^3. When $k = 2.5$ A^{-1}, the fraction is 0.716, corresponding to $A = -3.1$ A^3. We conclude that for $k < 2$ A^{-1}, replacement of $\exp(i\sum g)$ by $1 + i\sum g$ does not seriously raise the energy.

5. COMPUTATIONS WITH THE NEW WAVE FUNCTION

(a) Definitions

If the wave function for an excited state is $\psi=F\varphi$, it is easily shown (see III) that

$$\mathcal{E}=\int\psi^*H\psi d^N\mathbf{r}=(\hbar^2/2m)\sum_i\int\nabla_iF\cdot\nabla_iF^*\varphi^2d^N\mathbf{r}. \quad (22)$$

The only memory of the potentials is in the ground-state wave function φ; the information which we need about φ will be taken from experiment, since our main interest here is to test some ideas about the nature of excited states and not to develop a detailed theory of the ground state.[12] Substitution of (21) into (22) and (8) gives $E=\mathcal{E}/\mathcal{I}$ where

$$2m\mathcal{E}/N\hbar^2=k^2[1+A(I_1+I_2)+A^2(k^2I_3+I_4+I_5+kI_6+I_7)], \quad (23)$$

$$\mathcal{I}/N=I_8+AkI_9+A^2k^2I_{10}, \quad (24)$$

and

$$I_1=-2\mathbf{k}(\rho_0k)^{-1}\cdot\int\nabla g_1(\mathbf{r}_{21})\rho_2(1,2)d\mathbf{r}_{21},$$

$$I_2=2(\rho_0k)^{-1}\int\exp(i\mathbf{k}\cdot\mathbf{r}_{12})\,\mathbf{k}\cdot\nabla g_1(\mathbf{r}_{12})\rho_2(1,2)d\mathbf{r}_{21},$$

$$I_3=\rho_0^{-1}\int g_1(\mathbf{r}_{21})g_1(\mathbf{r}_{31})\rho_3(1,2,3)d\mathbf{r}_{21}d\mathbf{r}_{31},$$

$$I_4=\rho_0^{-1}\int\nabla g_1(\mathbf{r}_{21})\cdot\nabla g_1(\mathbf{r}_{31})\rho_3(1,2,3)d\mathbf{r}_{21}d\mathbf{r}_{31},$$

$$I_5=\rho_0^{-1}\int\exp(i\mathbf{k}\cdot\mathbf{r}_{23})\nabla g_1(\mathbf{r}_{12})\cdot\nabla g_1(\mathbf{r}_{13})\rho_3(1,2,3)d\mathbf{r}_{21}d\mathbf{r}_{31}, \quad (25)$$

$$I_6=2i(\rho_0k)^{-1}\int\exp(i\mathbf{k}\cdot\mathbf{r}_{13})g_1(\mathbf{r}_{21})\mathbf{k}\cdot\nabla g_1(\mathbf{r}_{13})\rho_3(1,2,3)d\mathbf{r}_{21}d\mathbf{r}_{31},$$

$$I_7=-2\rho_0^{-1}\int\exp(i\mathbf{k}\cdot\mathbf{r}_{13})\nabla g_1(\mathbf{r}_{21})\cdot\nabla g_1(\mathbf{r}_{13})\rho_3(1,2,3)d\mathbf{r}_{21}d\mathbf{r}_{31},$$

$$I_8=\int e^{i\mathbf{k}\cdot\mathbf{r}}p(r)d\mathbf{r}=S(k)\quad[\text{See Eq. (10)}],$$

$$I_9=-2i\rho_0^{-1}\int\exp(i\mathbf{k}\cdot\mathbf{r}_{12})g_1(\mathbf{r}_{32})\rho_3(1,2,3)d\mathbf{r}_{21}d\mathbf{r}_{31},$$

$$I_{10}=\rho_0^{-1}\int\exp(i\mathbf{k}\cdot\mathbf{r}_{12})g_1(\mathbf{r}_{31})g_1(\mathbf{r}_{42})\times\rho_4(1,2,3,4)d\mathbf{r}_{21}d\mathbf{r}_{31}d\mathbf{r}_{41}. \quad (25)$$

[12] R. M. Mazo and J. G. Kirkwood [Proc. Natl. Acad. Sci. 41, 204 (1955)] have computed $p(r)$ theoretically by solving an approximate integral equation.

We have written $g(\mathbf{r})=Akg_1(\mathbf{r})$. The mean density of atoms is $\rho_0=N/V$. The probability in the ground state that atoms are located at $\mathbf{r}_1$ and $\mathbf{r}_2$ is $\rho_2(\mathbf{r}_1,\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2$. Except in the negligible region near the surface of the liquid, we have $\rho_2(\mathbf{r}_1,\mathbf{r}_2)=\rho_0p(r_{12})$. In writing (25) we have made use of the fact that certain integrals like $\int g(\mathbf{r}_{12})\rho_2(\mathbf{r}_1,\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2$ vanish because g is odd. A term $\rho_0\delta(\mathbf{r}_{12})$ is contained in $\rho_2(\mathbf{r}_1,\mathbf{r}_2)$, and hence $p(r)$ contains a δ-function at the origin. We define p_1 and p_2 by

$$p(r)=\delta(\mathbf{r})+p_1(r), \quad (26)$$

$$p_1(r)=\rho_0[1+p_2(r)]; \quad (27)$$

these functions have no singularities and $p_2(r)\to0$ as $r\to\infty$. Strictly speaking, in the definition of I_1 we should replace $\rho_2(\mathbf{r}_1,\mathbf{r}_2)$ by $\rho_2(\mathbf{r}_1,\mathbf{r}_2)-\rho_0\delta(\mathbf{r}_{12})$ since g is always a function of the relative coordinates of two *distinct* atoms. To avoid unnecessary confusion, however, it is easier to think of $g(\mathbf{r})$ as becoming zero for sufficiently small r. Similar remarks apply to ρ_3 and ρ_4 when they occur in I_2, $\cdots$, I_{10}. If one does not wish to think of $g(\mathbf{r})$ as being modified near the origin, then the ρ's should be understood as containing delta functions of all coordinate differences except those which appear as arguments of g in the same integral. The probability in the ground state that atoms are at $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$ is $\rho_3(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3$. The nonsingular part of ρ_3, which we call ρ_3', is defined by

$$\rho_3(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)=\rho_3'(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)+\rho_0p_1(r_{12})\delta(\mathbf{r}_{23})+\rho_0p_1(r_{13})\delta(\mathbf{r}_{12})+\rho_0p_1(r_{23})\delta(\mathbf{r}_{13})+\rho_0\delta(\mathbf{r}_{12})\delta(\mathbf{r}_{23}). \quad (28)$$

No experimental data for ρ_3' are available. If any of the mutual distances, say r_{12}, is large, then

$$\rho_3'(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)=\rho_0p_1(r_{13})p_1(r_{23}).$$

If any of the interatomic distances becomes less than 2.4 A, then $\rho_3'=0$. The approximation[13]

$$\rho_3'(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)\simeq p_1(r_{12})p_1(r_{13})p_1(r_{23}) \quad (29)$$

has these correct limiting features. Much has been written about the validity of this approximation; for some, but not all purposes (29) is quite sufficient. We shall see that our answer is only slightly sensitive to the difference between the right and left sides of (29). Furthermore, we shall be able to estimate the magnitude and sign of the errors due to (29).

I_1, I_3, and I_4 are independent of k. In the other integrals it will prove possible to extract most of the k-dependence rather simply in the roton region, the remaining complicated terms being very small. This means that the computation of the entire roton spectrum will not be much more difficult than the computation of one point on it. We now discuss the evaluation of the various integrals.

[13] This is sometimes called the Kirkwood approximation or the superposition approximation.

(b) Evaluation of I_1, I_2, and I_4

I_1 can be done exactly. We integrate by parts over a volume bounded by two concentric surfaces, one lying inside the radius where $p(r)=0$ and the other very far from the origin. The inner surface contributes nothing, but the integrand $g_1(\mathbf{r})p_1(r)$ falls off only as r^{-2}, with the result that the outer surface makes a finite contribution, which is easily computed to be $-(8\pi\rho_0/3)$. We eliminate this contribution by redefining $g(\mathbf{r})$ to have a decay factor, say $e^{-\epsilon r}$, with very small ϵ, which makes surface terms vanish at infinity. This procedure is mathematically legitimate, since we are free to use any wave function we want in the variational principle, and is in accord with the physical idea that all the momentum of the backflow should be contained in a finite volume. It will generally not be necessary to represent ϵ explicitly; the convergence factor will be used only to justify certain operations. After the integration by parts, there remains

$$I_1=(2\mathbf{k}/k)\cdot\int g_1(\mathbf{r})\nabla p_1(r)d\mathbf{r}$$
$$=(8\pi/3)\int_0^\infty [dp_1(r)/dr]dr=8\pi\rho_0/3. \quad (30)$$

In the last integral, the integrand should really be $e^{-\epsilon r}(dp_1/dr)$, but if ϵ is small enough the convergence factor will be unity out to radii where dp_1/dr becomes negligible.

After performing the angular integrations in I_2, we find

$$I_2=16\pi\rho_0\int_{r_0}^\infty r^{-1}j_2(kr)[1+p_2(r)]dr$$
$$=16\pi\rho_0[(kr_0)^{-1}j_1(kr_0)+F(k)]=16\pi\rho_0 I_{2a}, \quad (31)$$

where r_0 is any radius inside the region where $p_1(r)=0$ [we take $r_0=2.4$ A, the radius where $p_1(r)$ first becomes positive] and

$$F(k)=\int_{r_0}^\infty r^{-1}j_2(kr)p_2(r)dr. \quad (32)$$

In order to do integrals like I_4 and I_5, we need to know the value of $I_{2a}(k)$ for all k. Using tabulated values for the spherical Bessel function j_2, $F(k)$ was evaluated by numerical integration for 23 values of k between 0 and 7 A^{-1}. Figure 3 gives the results for $I_{2a}(k)$. For $k<1.5$ A^{-1}, $F(k)$ is negligible compared with $(kr_0)^{-1}\times j_1(kr_0)$.

One might expect from (25) that $I_2\to -I_1$ as $k\to 0$. As k approaches zero, $F(k)$ approaches zero and $j_1(kr_0)/kr_0$ approaches $\frac{1}{3}$. Comparison of (30) and (31) thus shows that I_2 approaches $2I_1$ instead of $-I_1$. The reason for the discrepancy is that (31) is wrong when k is very small, of the same order of magnitude as ϵ; in this case we must take account of the term $e^{-\epsilon r}$ in g, and there will be a correction term which will cause I_2 to change from $16\pi\rho_0/3$ to $-8\pi\rho_0/3$ as k decreases from ϵ to 0.

I_2 can also be evaluated in momentum space, using data for $S(k)$ rather than $p(r)$. In momentum space the integrals converge best for small k rather than large k. The results are not very important because (31) is useful down to $k=0$; but they do provide a check of our numerical work and also of the consistency of the data for $p(r)$ with that for $S(k)$. $S(k)$ was defined by (10) as the Fourier transform of $p(r)$, where $p(r)$ includes a delta function at the origin and a constant term ρ_0 at infinity. Therefore $S(k)\to 1$ as $k\to\infty$ and $S(k)$ includes a term $(2\pi)^3\rho_0\delta(\mathbf{k})$. We define

$$S_1(k)=S(k)-1-(2\pi)^3\rho_0\delta(k). \quad (33)$$

It follows that

$$\rho_0 p_2(r)=(2\pi)^{-3}\int e^{i\mathbf{k}\cdot\mathbf{r}}S_1(k)d\mathbf{k}. \quad (34)$$

Taking the Fourier transform of $\nabla g_1(\mathbf{r})$, we obtain after the angular integrations

$$I_2=16\pi\rho_0/3+(2/\pi)\int_0^\infty S_1(k_1)k_1^2 b(k_1/k)dk_1, \quad (35)$$

where

$$b(x)=\frac{5}{6}-\frac{x^2}{2}-\frac{1}{4x}(1-x^2)^2\log\left|\frac{1-x}{1+x}\right|.$$

The numerical integral in (35) was evaluated for $k=0.5$, 1.0, 1.5, 2.0 A^{-1}. Convergence is good, and the values are accurate to within a few percent. Nevertheless, (35) does not give accurate values of I_2 when $k>1.5$ A^{-1}, because for large k the cancellation between the two terms of (35) is almost complete (as it must be because $I_2\to 0$ as $k\to\infty$), and hence a 3% error in the numerical integral may cause a 30% error in I_2. We take the volume per atom of liquid helium

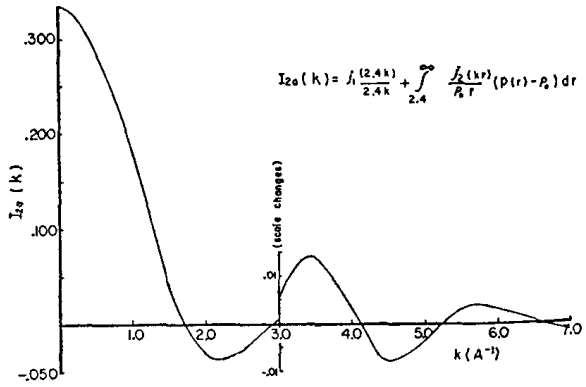


FIG. 3. $I_{2a}(k)$ is the Fourier transform of $p(r)$ times the velocity distribution in the pattern of backflow around a moving atom. In the most important region ($k<1.5$ A^{-1}), $I_{2a}(k)$ would be unchanged if we took $p(r)=0$ for $r<2.4$ A and $p(r)=\rho_0$ for $r>2.4$ A.

as 45 A^3.[14] The following table compares the values of $I_2(k)$ obtained from (31) with those obtained from (35).

k(A^{-1})		0	0.5	1.0	1.5	2.0	∞
$I_2(k)$ (A^{-3})	from (31)	0.372	0.322	0.194	0.054	−0.036	0
	from (35)	0.372	0.326	0.200	0.050	−0.060	0

The discrepancy at $k=2$ A^{-1} is not serious, for the reasons just mentioned, and the agreement elsewhere is sufficiently close for our purposes. The values derived from (31) are used throughout our work.

I_4 presents no problem if $I_{2a}(k)$ is known for all k. Using (28) and the approximation (29) for ρ_3', and Fourier analyzing $p_2(r_{23})$ with (34) we obtain

$$I_4=\int p_1(r)[\nabla g_1(\mathbf{r})]^2 d\mathbf{r}+\left[\int p_1(r)\nabla g_1(\mathbf{r})d\mathbf{r}\right]^2 +(2\pi)^{-3}(\rho_0)^{-1}\int d\mathbf{k}_1 S_1(\mathbf{k}_1)\left[\int \exp(i\mathbf{k}_1\cdot r)\times p_1(r)\nabla g_1(\mathbf{r})d\mathbf{r}\right]^2. \quad (36)$$

The integral in square brackets is a generalization of I_2 to the case where the $\mathbf{k}$ in the exponential has a different direction from the $\mathbf{k}$ in g_1. The angular integrations are easily performed, yielding

$$I_4=8\pi\rho_0\int_{r_0}^{\infty}[1+p_2(r)]r^{-4}dr+(4\pi\rho_0/3)^2 +16\rho_0\int_0^{\infty}k_1^2 S_1(k_1)[I_{2a}(k_1)]^2 dk_1$$
$$=0.01190+0.00867-0.00790=0.01267\text{A}^{-6}. \quad (37)$$

The value obtained for I_4 may be in error because of uncertainties in the values of $p(r)$ and $S(k)$, and also because the approximation (29) is not exact. Discussion of the error due to (29) is postponed until the evaluation of I_3. The uncertainty in $p_2(r)$ is unimportant because the magnitude of $\int_{r_0}^{\infty} p_2(r)r^{-4}dr$ is only 1/10 that of $\int_{r_0}^{\infty} r^{-4}dr$. Similarly, 90% of the contribution to $\int_0^{\infty} k_1^2 S_1(k_1)[I_{2a}(k_1)]^2 dk_1$ comes from the region $k_1<1.2$ A^{-1}. In this region $I_{2a}(k_1)$ is the same as would result if we took $p_1(r)=0$ for $r<2.4$ A and $p_1(r)=\rho_0$ for $r>2.4$ A, and $S_1(k_1)$ is largely determined by its value and slope at the origin, both of which are known theoretically. The important point to be learned from this discussion is that the values of I_4, and of the other integrals which contribute significantly to the coefficient of A^2 in (23), depend mainly on the gross features of $p(r)$ (i.e., its delta function at the origin, vanishing for $r<2.4$ A, and its quick approach to the asymptotic value ρ_0) and not on the details of its behavior. The coefficients of A are more sensitive to the detailed behavior of $p(r)$; it is the detailed behavior which determines the location of the minimum in the energy spectrum. The insensitivity of the quadratic coefficients to the exact form of $p(r)$ can be similarly verified in the computations which follow, and will not be pointed out explicitly.

(c) Approximate Methods

The value of I_4 and the size of the various terms which contribute to it can be understood fairly well in terms of some simple approximations for integrals involving the coordinates of three atoms. With the help of these approximations we can understand the sizes of all the remaining integrals; if we know that an integral is small, it will not be necessary to waste time in evaluating it very accurately.

Suppose we want to do an integral of the form

$$\int f(\mathbf{r}_{21})f(\mathbf{r}_{31})\rho_3(1,2,3)d\mathbf{r}_2 d\mathbf{r}_3$$

(in this integral we shall understand ρ_3 to include a delta function on coordinates 2 and 3, but not on any other pairs). If the positions of 1 and 2 are fixed and 3 is not too close to 2, then $\rho_3(1,2,3)$ can be approximated very closely by $\rho_0 p_1(r_{21})p_1(r_{31})$. We write

$$\rho_3(1,2,3)\simeq\rho_0 p_1(r_{21})p_1(r_{31}). \quad (38)$$

When 3 approaches 2, this is wrong because ρ_3 goes to zero but $p_1(r_{21})p_1(r_{31})$ keeps a finite value (assuming, of course, that $r_{21}>2.4$ A; otherwise both expressions are zero). When 2 and 3 coincide, however, ρ_3 exhibits a delta function and far exceeds $\rho_0 p_1(r_{21})p_1(r_{31})$. The strength of the delta function is such that if we integrate the difference between the two sides of (38) over the positions of 3, the result is exactly zero, i.e.,

$$\int[\rho_3(1,2,3)-\rho_0 p_1(r_{21})p_1(r_{31})]d\mathbf{r}_3=0. \quad (39)$$

We believe this equation not to be a relation among distribution functions in general, but to hold for the distribution functions for the liquid at absolute zero. We do not have a rigorous proof, but shall discuss our reasons for believing it in Appendix B.

If $f(\mathbf{r})$ is a slow-varying function, i.e., $f(\mathbf{r})$ does not change much when $\mathbf{r}$ changes by 2.4 A, then for a fixed value of $\mathbf{r}_{12}$ the value of $f(\mathbf{r}_{31})$ is almost constant over the region where the two sides of (38) differ appreciably. Using (39), we see that the integral

$$\int f(\mathbf{r}_{31})[\rho_3(1,2,3)-\rho_0 p_1(r_{21})p_1(r_{31})]d\mathbf{r}_3$$

[14] The atomic volume of liquid He under its own saturated vapor pressure at 0°K is 46 A^3, but 45 A^3 is closer to the value at 2.06°K, where the structure factor data was taken. Internal inconsistencies would develop if ρ_0 and $S(k)$ were taken at different temperatures. One might ask where the theory takes account of the external pressure. The pressure determines the values of ρ_0 and, more important, $S(k)$. An increase in pressure is expected to sharpen the maxima and minima of $p(r)$ and $S(k)$.

is very close to zero. We therefore find

$$\int f(\mathbf{r}_{31})\rho_3(1,2,3)d\mathbf{r}_3 \simeq \rho_0 p_1(r_{21})\int f(\mathbf{r}_{31})p_1(r_{31})d\mathbf{r}_3,$$

and finally,

$$\int f(\mathbf{r}_{21})f(\mathbf{r}_{31})\rho_3(1,2,3)d\mathbf{r}_2 d\mathbf{r}_3 \simeq \rho_0\left[\int f(\mathbf{r})p_1(r)d\mathbf{r}\right]^2. \quad (40)$$

Similarly, if f *or* g is slowly varying,

$$\int f(\mathbf{r}_{21})g(\mathbf{r}_{31})\rho_3(1,2,3)d\mathbf{r}_2 d\mathbf{r}_3 \simeq \rho_0\left(\int f(\mathbf{r})p_1(r)d\mathbf{r}\right)\left(\int g(\mathbf{r})p_1(r)d\mathbf{r}\right). \quad (41)$$

Actually, our criterion for a slowly varying function is too stringent. The behavior of $f(\mathbf{r})$ for $r<2.4$ A is of no importance, since $p_1(r)$ is zero in that range; hence f may be singular at the origin. The important question is, how much does $f(\mathbf{r}_1+\mathbf{r}_2)$ differ from $f(\mathbf{r}_1)$ when $\mathbf{r}_1$ and $\mathbf{r}_2$ are any two vectors of length 2.4 A? And even if the difference is large compared with $f(\mathbf{r}_1)$, (40) is still good if $f(\mathbf{r})$ is such that the major contribution to $\int f(\mathbf{r})p_1(r)d\mathbf{r}$ comes from $r>3$ or 4 A.

Another type of integral which interests us is

$$\int f(\mathbf{r}_{21})g(\mathbf{r}_{31})h(\mathbf{r}_{23})\rho_3(1,2,3)d\mathbf{r}_2 d\mathbf{r}_3,$$

where f and g are smooth and $h(\mathbf{r})$ oscillates so rapidly that it produces almost complete cancellation when integrated against $p_1(r)$. ρ_3 is still understood to contain a delta function on 2 and 3, and on no other pair. In this case, if 1 and 2 are held fixed and 3 is allowed to vary, the oscillation in $h(\mathbf{r}_{23})$ make the contribution to the integral small. The major contribution comes when 3 and 2 are tied together by the delta function and we find

$$\int f(\mathbf{r}_{21})g(\mathbf{r}_{31})h(\mathbf{r}_{23})\rho_3(1,2,3)d\mathbf{r}_2 d\mathbf{r}_3 \simeq \rho_0 h(0)\int f(\mathbf{r})g(\mathbf{r})p_1(r)d\mathbf{r}. \quad (42)$$

If $\nabla g_1(\mathbf{r})$ is sufficiently smooth, (40) can be used to estimate I_4. The answer thus obtained is $(4\pi\rho_0/3)^2 = 0.00867$ A^{-6}, which is the middle term of (37); if ∇g_1 were very smooth, the first and third terms would cancel completely. The first term (0.01190 A^{-6}) is larger than the third term (-0.00790 A^{-6}) because $\nabla g_1(\mathbf{r})$ is proportional to r^{-3} and therefore quite strongly peaked for small r; hence the delta function more than compensates for the "hole" in ρ_3. The answer given by (40) is $\frac{2}{3}$ the correct answer.

With the aid of (40)–(42) we can discuss the remaining integrals more intelligently. If Landau's energy spectrum is even qualitatively correct, then the most important points to compute are those in the neighborhood of the roton minimum. The phonon spectrum is guaranteed to be correct; and when the temperature is high enought to excite the portion of the spectrum lying appreciably above the roton minimum, then the picture of the liquid as a gas of independent excitations has broken down. Thus, even if we knew the exact form of the high part of the spectrum, we would not know how to do the thermodynamics. Furthermore, the high-momentum end of the spectrum computed with (20) or (21) is certainly wrong, since the slope dE/dp exceeds the velocity of sound when $k\geq 2.2$ A^{-1}; whenever $|dE(p)/dp|>c$, there obviously exist states with two excitations, one of which is a phonon, which have total momentum p but energy less than $E(p)$. We shall therefore compute the energy at several points in the region 1.6 A$^{-1}\leq k\leq 2.4$ A^{-1}, and also at $k=1.2$ A^{-1} in order to estimate the height of the hump between the phonon and roton regions.

(d) Evaluation of I_3 and Correction to the Kirkwood Approximation

Since g_1 is smoother than ∇g_1, I_3 is a good candidate for the approximation (40), which predicts $I_3=0$ because $\int p_1(r)g_1(\mathbf{r})d\mathbf{r}=0$. We infer that I_3 is small; but it is important to know *how* small, because the factor k^2 which multiplies I_3 in (23) is fairly large. The exact value of I_3 [i.e., no approximations beyond (29)] can be computed by the method used for I_4. The result is

$$\begin{aligned} I_3 &= I_{3a}+I_{3b} \\ &= \int [g_1(\mathbf{r})]^2 p_1(r)d\mathbf{r} + \int g_1(\mathbf{r}_{21})g_1(\mathbf{r}_{31})p_1(r_{21}) \\ &\qquad \times p_1(r_{31})p_2(r_{23})d\mathbf{r}_{21}d\mathbf{r}_{31} \\ &= \rho_0\Big\{(4\pi/3)\int_0^\infty [1+p_2(r)]r^{-2}dr \\ &\qquad + (8/3)\int_0^\infty S_1(k)[I_{9a}(k)]^2 dk\Big\} \\ &= (1/45)(1.707-1.470) = 0.0053\text{A}^{-4}. \end{aligned}$$

The integral $I_{9a}(k)$ is defined by Eq. (49). The approximation (40) is based on the idea that I_{3a} and I_{3b} should cancel each other. Since $I_{3a}-I_{3b}$ is only 14% of I_{3a}, the idea behond (40) is good, but (40) tells us nothing about the size of I_3 because $\int g_1(\mathbf{r})p_1(r)d\mathbf{r}=0$.

If $I_3=0.0053$ A^{-4}, then in the roton region the term $k^2 I_3$ contributes about half of the total coefficient of A^2 in (23). Any possibility of serious error in I_3 ought therefore to be investigated carefully. The idea that I_3 is almost zero is based on the approximation (40), which in turn is based on the identity (39). Actually,

the approximate form which we have used for ρ_3 does not satisfy (39) exactly. Slight departures from (39) ordinarily would not affect the validity of (40), were it not for the fact that $\int p_1(r)g_1(\mathbf{r})d\mathbf{r}=0$. In this case the question arises; how much of the failure of I_3 to vanish is real, and how much is due to the fact that the approximate ρ_3 does not satisfy (39)? An exact expression for I_3 is

$$I_3=0.0053-\rho_0^{-1}\int g_1(\mathbf{r}_{21})g_1(\mathbf{r}_{31})[p_1(r_{21})p_1(r_{31})p_1(r_{23}) -\rho_3'(1,2,3)]d\mathbf{r}_{21}d\mathbf{r}_{31} =0.0053-I_{3c}. \quad (43)$$

If any one the mutual distances is less than 2.4 A or more than about 4 A, then $p_1(r_{21})p_1(r_{31})p_1(r_{23})-\rho_3'(1,2,3)$ is very close to zero. Consequently, the integrand of I_{3c} is appreciable only if the three atoms are at the corners of a triangle, each of whose legs may vary in length from 2.4 to 4 A. Therefore, if the spatial variation of $g_1(\mathbf{r})$ were slow, the replacement of $g_1(\mathbf{r}_{21})g_1(\mathbf{r}_{31})$ by $[g_1(\mathbf{r}_{31})]^2$ would not greatly alter the value of I_{3c}. The resulting integral is then easily evaluated. We can, however, find an even better approximation to I_{3c} by taking the angular variation of $g_1(\mathbf{r})$ into account. Since I_{3c} is independent of the direction of $\mathbf{k}$, we can average the integrand over the directions of $\mathbf{k}$. The average of $(\mathbf{k}\cdot\mathbf{r}_{12})(\mathbf{k}\cdot\mathbf{r}_{31})$ is $\frac{1}{3}k^2\mathbf{r}_{21}\cdot\mathbf{r}_{31}$, and in the important configurations the three atoms almost form an equilateral triangle; therefore, the average over these configurations of the cosine of the angle between $\mathbf{r}_{21}$ and $\mathbf{r}_{31}$ is very close to $\frac{1}{2}$. Most of the angular dependence of the integrand is therefore correctly accounted for if we replace $\mathbf{r}_{21}\cdot\mathbf{r}_{31}$ by $\frac{1}{2}r_{21}r_{31}$; at this stage we note that the radii r_{21} and r_{31} are almost equal in the important region, and we take r_{21} and r_{31} to be the same in the integrand. This approximation differs from the preceding one through the presence of the factor $\frac{1}{2}$. We obtain

$$I_{3c}\simeq I_{3d},$$

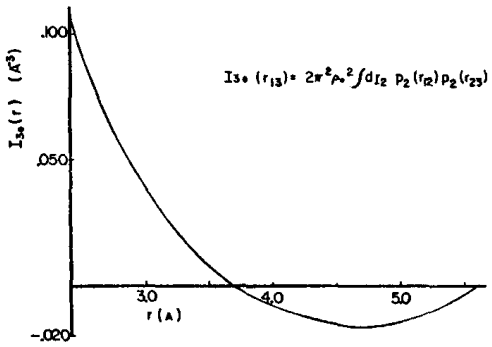


FIG. 4. $I_{3c}(r)$ measures the error in the Kirkwood approximation, and would vanish for all $r>2.4$ A if the approximation were exact. The rapid decrease of $I_{3c}(r)$ accounts for the high accuracy of the Kirkwood approximation in this computation.

where

$$I_{3d}=(6\rho_0)^{-1}\int r_{31}^{-4}[p_1(r_{21})p_1(r_{31})p_1(r_{23}) -\rho_3'(1,2,3)]d\mathbf{r}_{21}d\mathbf{r}_{31}.$$

The identities (39) and (70) imply that

$$\int d\mathbf{r}_2[p_1(r_{21})p_1(r_{31})p_1(r_{23})-\rho_3'(1,2,3)] =\rho_0^2p_1(r_{13})\int p_2(r_{23})p_2(r_{12})d\mathbf{r}_2.$$

We find

$$I_{3d}=(3\pi\rho_0)^{-1}\int_{2.4}^{\infty}r^{-2}p_1(r)I_{3e}(r)dr, \quad (44)$$

where

$$I_{3e}(r_{13})=2\pi^2\rho_0^2\int d\mathbf{r}_2p_2(r_{12})p_2(r_{23}) \quad (45a)$$

$$=\int_0^{\infty}(kr_{13})^{-1}\sin(kr_{13})[S_1(k)]^2k^2dk. \quad (45b)$$

$I_{3e}(r)$ was computed from (45b) for 19 values of r between 2.4 A and 5.6 A. The numerical integrals converge well, and the results are shown in Fig. 4. Performing the final numerical integration in (44), we find $I_{3d}=0.00040$ A^{-4} and finally

$$I_3=0.0049\text{ A}^{-4}. \quad (46)$$

The smallness of the correction to I_3 shows that the slight failure of (39) does not cause a significant error in I_3. This fact was not intuitively obvious, however, and needed verification. It should be emphasized that we have gone beyond the Kirkwood approximation. We have written an exact expression I_{3c} for the error due to the Kirkwood approximation, and we have estimated I_{3c} quite accurately by an integral I_{3d} which is easily evaluated. We believe the inaccuracy in the approximation $I_{3c}\simeq I_{3d}$ to be about 25%, and therefore our lack of knowledge of ρ_3 causes a residual uncertainty of 0.0001 A^{-4} in the value of I_3.

By exactly the same method, one can estimate the error in I_4 caused by the Kirkwood approximation. The answer is 0.0001 A^{-6}, which is negligible compared with the value given by (37).

(e) Evaluation of Remaining Integrals

I_9 occurs in (24) as a coefficient of A, rather than A^2, and ought therefore to be treated as accurately as is possible. In I_9, ρ_3 includes delta functions on $\mathbf{r}_{12}$ and $\mathbf{r}_{13}$. Using (29), and noting that several terms are zero because g_1 is odd, we obtain

$$I_9=-2i\int\exp(i\mathbf{k}\cdot\mathbf{r}_{12})g_1(\mathbf{r}_{32})p_2(r_{13})p_1(r_{12}) \times p_1(r_{23})d\mathbf{r}_{21}d\mathbf{r}_{31}-2i\int\exp(i\mathbf{k}\cdot\mathbf{r}_{12})g_1(\mathbf{r}_{12}) \times p_1(r_{12})d\mathbf{r}_{21}. \quad (47)$$

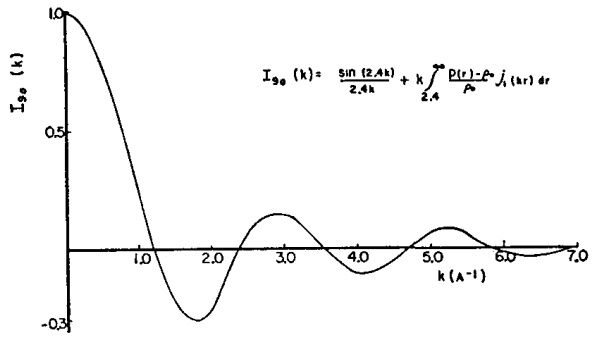


FIG. 5. $I_{9a}(k)$ is the Fourier transform of $p(r)$ times the velocity potential for the backflow pattern. Like $I_{2a}(k)$ it is determined in in the most important region by the gross features of $p(r)$.

First we consider the second integral, since we must know its value for all $\mathbf{k}$ in order to do the first integral. The integral, like I_2, can be performed in coordinate or momentum space; after the angular integrations are done, the result in coordinate space is

$$\int \exp(i\mathbf{k}_1\cdot\mathbf{r})g_1(\mathbf{r})p_1(r)d\mathbf{r}=(\mathbf{k}_1\cdot k/k_1^2k)4\pi\rho_0 iI_{9a}(k_1), \tag{48}$$

where[15]

$$I_{9a}(k)=(2.4k)^{-1}\sin(2.4k)+k\int_{2.4}^{\infty}p_2(r)j_1(kr)dr. \tag{49}$$

As before, the coordinate space formula proves sufficient over the entire range of k. For small k, when the numerical integral cannot be done accurately, its value is so small as to be unimportant compared with $(2.4k)^{-1}\sin(2.4k)$. Figure 5 gives the values of $I_{9a}(k)$. As in the case of $I_2(k)$, some points were also computed in momentum space, using data for $S(k)$ rather than $p(r)$. The results were in good agreement with the coordinate space computations.

Since $p_1(r_{12})=\rho_0[1+p_2(r_{12})]$, the first integral in (47) becomes

$$\rho_0\int \exp(i\mathbf{k}\cdot\mathbf{r}_{12})g_1(\mathbf{r}_{32})p_1(r_{32})[1+p_2(r_{12})]p_2(r_{13})d\mathbf{r}_{21}d\mathbf{r}_{31}$$

$$=\rho_0\int \exp(i\mathbf{k}\cdot\mathbf{r}_{12})g_1(\mathbf{r}_{32})p_1(r_{32})p_2(r_{13})d\mathbf{r}_{21}d\mathbf{r}_{31}$$

$$+\rho_0\int \exp(i\mathbf{k}\cdot\mathbf{r}_{12})g_1(\mathbf{r}_{32})p_1(r_{32})\times p_2(r_{12})p_2(r_{13})d\mathbf{r}_{21}d\mathbf{r}_{31}. \tag{50}$$

[15] Since $\int g_1(\mathbf{r})p_1(r)d\mathbf{r}=0$, one might expect the right side of (48) to approach zero as k_1 becomes small. But $I_{9a}(k_1)$ approaches unity for small k_1, and consequently the right side of (48) approaches $\pm\infty$, depending on the angle between $\mathbf{k}$ and $\mathbf{k}_1$. The trouble, as before, is resolved by noting that (48) and (49) are wrong for $k_1<\epsilon$ (g_1 should really gave a factor $e^{-\epsilon r}$ in it). In the correct version of (49) the term $(2.4k)^{-1}\sin(2.4k)$ is replaced by zero when $k\ll\epsilon$; hence $I_{9a}(k_1)$ goes as k_1^2 when $k_1\ll\epsilon$, and the right side of (49) approaches zero. The "error" in (48) and (49) has no effect on our computations, but is worth mentioning lest the reader discover it and develop a distrust of the formulas.

The first integral on the right in (50) can be evaluated by writing $\exp(i\mathbf{k}\cdot\mathbf{r}_{12})=\exp(i\mathbf{k}\cdot\mathbf{r}_{13})\exp(i\mathbf{k}\cdot\mathbf{r}_{32})$ and using the new integration variables $\mathbf{r}_{31}$ and $\mathbf{r}_{32}$. The result is

$$\int \exp(i\mathbf{k}\cdot\mathbf{r}_{12})g_1(\mathbf{r}_{32})p_1(r_{32})p_2(r_{13})d\mathbf{r}_{21}d\mathbf{r}_{31}=(4\pi i/k)S_1(k)I_{9a}(k). \tag{51}$$

In the second integral on the right in (50) we use the integration variables $\mathbf{r}_{21}$ and $\mathbf{r}_{32}$; Fourier-analyzing $p_2(r_{13})$, we obtain

$$\int \exp(i\mathbf{k}\cdot\mathbf{r}_{12})g_1(\mathbf{r}_{32})p_1(r_{23})p_2(r_{12})p_2(r_{13})d\mathbf{r}_{21}d\mathbf{r}_{31}$$

$$=(i\pi\rho_0)^{-1}\int_0^{\infty}\int_0^{\pi}dk_1d\theta k_1S_1(k_1)S_1(|\mathbf{k}+\mathbf{k}_1|)\times I_{9a}(k_1)\cos\theta\sin\theta, \tag{52}$$

where θ is the angle between $\mathbf{k}$ and $\mathbf{k}_1$. If we let

$$u=|\mathbf{k}+\mathbf{k}_1|=(k^2+k_1^2+2kk_1\cos\theta)^{\frac{1}{2}}$$

the right side of (52) becomes

$$(2\pi\rho_0ik^2)^{-1}\int_0^{\infty}dk_1k_1^{-1}I_{9a}(k_1)S_1(k_1)\int_{|k-k_1|}^{k+k_1}\times S_1(u)(u^2-k^2-k_1^2)udu. \tag{53}$$

Since

$$\int_0^{\infty}dk_1f(k_1)\int_{|k-k_1|}^{k+k_1}g(u)du=\int_0^{\infty}dug(u)\int_{|k-u|}^{k+u}f(k_1)dk_1,$$

a numerical integral like (53) can be done in several different ways. We look for the way which converges fastest, is least sensitive to information which we do not have [like the value of $S_1(u)$ for large u], and does not involve small differences of big terms. For example, we should avoid dealing with the indefinite integral of $S_1(u)u^3$, since it oscillates badly for large u; hence (53) is not convenient to use as it stands. Probably the best form of (53) is

$$(2\pi\rho_0ik^2)^{-1}\Big[\int_0^{\infty}uS_1(u)(u^2-k^2)du\times\int_{|u-k|}^{u+k}dk_1I_{9a}(k_1)S_1(k_1)/k_1-\int_0^{\infty}uS_1(u)du\int_{|u-k|}^{u+k}dk_1k_1I_{9a}(k_1)S_1(k_1)\Big]=(2\pi\rho_0ik^2)^{-1}[I_{9b}(k)-I_{9c}(k)]. \tag{54}$$

In this form, the inside integral acts as a convergence factor for the integrand of the outside integral, and the answer is not sensitive to the values of $S_1(u)$ for large u.

The inside integral can be tabulated once and for all as an indefinite integral; thus, the evaluation of (54) involves only a single numerical integration for each value of k.

If (41) were used to estimate $I_9(k)$, the result would be

$$I_9(k)\simeq -2i\int e^{i\mathbf{k}\cdot\mathbf{r}}p_1(r)d\mathbf{r}\int g_1(\mathbf{r})p_1(r)d\mathbf{r}=0.$$

As in the case of I_3, the question arises: how much of the failure of $I_9(k)$ to vanish is real, and how much is due to the failure of our approximate ρ_3 to satisfy (39)? The analysis proceeds exactly as with I_3, and we find that the quantity

$$I_{9d}(k)=(2/\pi\rho_0)\int_0^\infty p_1(r)j_1(kr)I_{3e}(r)dr \qquad (55)$$

should be subtracted from (47). Combining Eqs. (47)–(55), we obtain

$$I_9(k)=(8\pi\rho_0/k)I_{9a}(k)S(k)+(\pi k^2)^{-1}[I_{9c}(k)-I_{9b}(k)]-I_{9d}(k). \qquad (56)$$

Table I gives values of $I_9(k)$, $I_{9b}(k)$, $I_{9c}(k)$, and $I_{9d}(k)$. In the roton region the correction I_{9d} is about one-tenth as large as I_9 (except near $k=2.4$ A^{-1}, where I_9 is negligible anyway). Since we believe that I_{9d} estimates, within an accuracy of 25%, the error due to the Kirkwood approximation, the residual error in I_9 due to this source is probably only 2 or 3%.

Using (29), we can write I_5 as

$$I_5=\int_0^\infty p_1(r)[\nabla g_1(\mathbf{r})]^2d\mathbf{r}+\left[\int_0^\infty e^{i\mathbf{k}\cdot\mathbf{r}}p_1(r)\nabla g_1(\mathbf{r})d\mathbf{r}\right]^2$$
$$+\int \exp(i\mathbf{k}\cdot\mathbf{r}_{23})\nabla g_1(\mathbf{r}_{21})\cdot\nabla g_1(\mathbf{r}_{31})p_1(r_{21})\times p_1(r_{31})p_2(r_{23})d\mathbf{r}_{21}d\mathbf{r}_{31}$$
$$=I_{5a}+I_{5b}+I_{5c}. \qquad (57)$$

The oscillatory factor $\exp(i\mathbf{k}\cdot\mathbf{r}_{23})$ makes I_5 a likely candidate for (42), which says[16] $I_5\simeq I_{5a}$. At the cost of considerable labor we have computed $I_{5b}+I_{5c}$ when $k=2$ A^{-1} and when $k=1.2$ A^{-1}, and verified that it could indeed have been neglected.

I_{5a} has been evaluated in connection with I_4. From (31) and (25) we obtain

$$I_{5b}(k)=[8\pi\rho_0 I_{2a}(k)]^2.$$

$I_{5c}(k)$ can be evaluated by the same methods used for I_9. The resulting expression is similar in form to (54), and will not be exhibited here. Laborious computations give

$$I_{5b}(2\text{ A}^{-1})+I_{5c}(2\text{ A}^{-1})=-0.0010\text{ A}^{-6}$$

and

$$I_{5b}(1.2\text{ A}^{-1})+I_{5c}(1.2\text{ A}^{-1})=0.0010\text{ A}^{-6}.$$

[16] Compare (42) with the definition of I_5 in (25).

TABLE I. Values of numerical integrals involved in $I_9(k)$.

k(A^{-1})	$I_{9b}(k)$(A^{-4}) [See (54)]	$I_{9c}(k)$(A^{-4}) [See (54)]	$I_{9d}(k)$(A^{-2}) [See (55)]	$I_9(k)$(A^{-2}) [See (56)]
1.2	−0.019	0.172	0.0076	0.0444
1.6	−0.387	0.089	−0.0002	−0.0054
1.8	−0.406	0.031	−0.0035	−0.0386
2.0	−0.364	−0.017	−0.0047	−0.0518
2.2	−0.086	−0.048	−0.0039	−0.0307
2.4	0.155	−0.054	−0.0020	−0.0044

Since $I_{5a}=0.0119$ A^{-6}, the complete omission of I_{5b} and I_{5c} would not cause a serious error in the roton spectrum. We shall omit these terms while locating the minimum of the spectrum, and shall reinstate them in the final computation of Δ.

I_6 is estimated by (41) as zero. As in the case of I_3, it is important to find out whether I_6 is really small enough to be neglected. To obtain a more accurate estimate, we write

$$I_6=I_{6a}+I_{6b},$$

where

$$I_{6a}=(2i/k)\int e^{-i\mathbf{k}\cdot\mathbf{r}}g_1(\mathbf{r})\mathbf{k}\cdot\nabla g_1(\mathbf{r})p_1(r)d\mathbf{r},$$

$$I_{6b}=(2i/k)\int \exp(i\mathbf{k}\cdot\mathbf{r}_{13})g_1(\mathbf{r}_{21})\mathbf{k}\cdot\nabla g_1(\mathbf{r}_{13})p_1(r_{31})p_1(r_{21})p_2(r_{23})d\mathbf{r}_{31}d\mathbf{r}_{21}.$$

According to the discussion preceding (41), I_{6a}, and I_{6b} will cancel each other almost completely, so I_6 is some fraction (probably about one-fourth) of I_{6a}. Performing the angular integrations in I_{6a}, we obtain

$$I_{6a}=8\pi\int_0^\infty dr r^{-3}p_1(r)[-2(kr)^{-1}\cos(kr)+8(kr)^{-2}\sin(kr)+18(kr)^{-3}\cos(kr)-18(kr)^{-4}\sin(kr)].$$

A rough numerical integration gives

$$I_{6a}(2\text{ A}^{-1})\simeq -0.003\text{ A}^{-5}.$$

Hence $I_6(2\text{ A}^{-1})\simeq -0.001$ A^{-5}, and $kI_6\simeq -0.002$ A^{-6} when $k=2$ A^{-1}. Since $k^2I_3+I_4+I_5\simeq 0.040$ A^{-6} near $k=2$ A^{-1}, we can neglect kI_6 without much error.

Estimation of I_7 by (41) gives

$$I_7\simeq -2(-4\pi\rho_0/3)8\pi\rho_0 I_{2a}(k)=\tfrac{1}{3}(8\pi\rho_0)^2I_{2a}(k). \qquad (58)$$

Considerations similar to those used in estimating I_6 show that (58) is accurate to better than 0.001 A^{-6}.

When k is in the roton region, the major portion of I_{10} comes from the term $\delta(\mathbf{r}_{12})\rho_3(2,3,4)$, which is contained in $\rho_4(1,2,3,4)$. When $r_{12}\neq 0$, the oscillations of $e^{i\mathbf{k}\cdot\mathbf{r}_{12}}$

make the contribution to the integral very small.[17] If we neglect all of ρ_4 except $\delta(\mathbf{r}_{12})\rho_3(2,3,4)$, we are making essentially the approximation which was used in I_5 and was shown to be very accurate there. We then obtain

$$I_{10}\simeq I_3 \tag{59}$$

and our evaluation of the integrals in (25) is completed.

The oscillation argument which leads to (59) fails when k is very small. For any value of k the requirement that the normalization integral $\mathcal{I}$ have no roots when considered as a polynomial in A leads to the inequality

$$I_{10}>I_9^2/4I_8. \tag{60}$$

The failure of (59) for small k is most easily seen by noting that I_3 becomes much smaller than the right side of (60) as $k\rightarrow 0$.

For $k\geq 1.2\ \text{A}^{-1}$, the coefficient of A^2 in (23) is estimated well by

$$k^2I_3+I_4+I_{5a}+\tfrac{1}{3}(8\pi\rho_0)^2I_{2a}(k).$$

We have omitted I_{5b}, I_{5c}, and kI_6, and have approximated I_7 by (58). I_{5c} and kI_6 have both been shown to be very small, and are both difficult to compute; omission of these terms simplies the computation of the energy spectrum, and does not significantly change the location of the minimum. I_{5b} has been omitted for the sake of consistency, since it is even smaller than I_{5c}. After locating the minimum, we shall reinstate the omitted terms in our final computation of Δ. We estimate I_{10} by (59).

6. THE IMPROVED ENERGY SPECTRUM

With the omissions and approximations mentioned in the preceding paragraph, we obtain

$$\frac{E(k)}{E_1(k)}<\frac{1+A[I_1+I_2]+A^2[k^2I_3+I_4+I_{5a}+\frac{1}{3}(8\pi\rho_0)^2I_{2a}]}{1+A[kI_9/I_8]+A^2[k^2I_3/I_8]}$$

$$=\frac{1+A[0.186+1.117I_{2a}(k)]+A^2[0.0246+0.0049k^2+0.108I_{2a}(k)]}{1+A[kI_9(k)/S(k)]+A^2[0.0049k^2/S(k)]}$$

$$=E_2(k)/E_1(k). \tag{61}$$

$E(k)$ is the true lowest energy of a state having momentum $\hbar\mathbf{k}$; $E_1(k)$ is the energy computed with the wave function (5), i.e., $E_1(k)=\hbar^2k^2/2mS(k)$; $E_2(k)$ is the spectrum we have computed, subject to the omissions and approximations noted above.

For $k=2\ \text{A}^{-1}$, (61) becomes

$$\frac{E_2(k)}{E_1(k)}=\frac{1+0.149A+0.0406\ \text{A}^2}{1-0.0822A+0.0156\ \text{A}^2}. \tag{62}$$

The first attractive feature of (62) is that the coefficients of A in the numerator and denominator have opposite signs, so that the denominator increases while the numerator decreases. The optimal value of A is -3.5, which is very close to the classical value $A_{cl}=-3.6$. The minimum value of $E_2(2)/E_1(2)$ is 0.659, corresponding to $E_2(2\ \text{A}^{-1})/\kappa=12.6°\text{K}$.

Computation of the coefficients in (61) and minimization of the resulting expressions yield the results given in Table II. We estimate the minimum value of $E_2(k)/\kappa$ as 12.0°K, corresponding to $k=1.85\ \text{A}^{-1}$. If we estimate $I_{5c}(1.85\ \text{A}^{-1})$ and $I_6(1.85\ \text{A}^{-1})$ by the values of the corresponding integrals at $k=2\ \text{A}^{-1}$, we find that the coefficient of A^2 in the numerator of (61) should be diminished by $0.003\ \text{A}^{-6}$ when $k=1.85\ \text{A}^{-1}$. This change lowers the energy by 0.5°K and we obtain the following as the final result of this computation[18]:

$$p_0/\hbar=1.85\ \text{A}^{-1},\quad \Delta/\kappa=\mathbf{11.5°K}. \tag{63}$$

It is evident from Table II that $E_2(k)/E_1(k)$ passes through a minimum near $k=1.2\ \text{A}^{-1}$. In any correct theory $E_2(k)/E_1(k)$ must approach unity for very small k because we cannot lower the energy of a phonon. By studying the behavior of the integrals in (25) for very small k, we have verified that our spectrum does indeed

TABLE II. The energy spectrum $E_2(k)$ computed from (61).[a]

$k(\text{A}^{-1})$	$A_{\text{opt}}(\text{A}^3)$	$E_2(k)/E_1(k)$	$E_2(k)/\kappa(°\text{K})$
1.2	−3.6	0.569	14.08
1.6	−3.7	0.576	13.44
1.8	−3.6	0.594	12.00
2.0	−3.5	0.659	12.59
2.2	−3.0	0.730	16.86
2.4	−2.5	0.791	24.04

[a] $E_2(k)$ is essentially the spectrum computed here. Some further small corrections lower the minimum energy to 11.5°K. $E_1(k)$ is the spectrum previously computed with a simpler wave function. A_{opt} is the optimal value of the strength of the return flow in the wave function (21), and is chosen so as to minimize $E_2(k)$. The values of A_{opt} are close to the "classical" value $-3.6\ \text{A}^3$ computed from a current conservation argument.

[17] If the integral

$$J(\mathbf{r}_{12})=(\rho_0)^{-2}\int\rho_4(1,2,3,4)g_1(\mathbf{r}_{31})g_1(\mathbf{r}_{42})d\mathbf{r}_3d\mathbf{r}_4,$$

were to become large compared with $J(0)$ as r_{12} grows large, then the growth of J might offset the oscillations of $\exp(i\mathbf{k}\cdot\mathbf{r}_{12})$ and (58) would be wrong. It is easy to see, however, that as 1 and 2 go farther apart, $J(\mathbf{r}_{12})$ approaches $[\int\rho_1(r)g_1(\mathbf{r})d\mathbf{r}]^2$, which is zero. Since the factorization of J into a product of two integrals becomes more nearly exact as r_{12} increases, it is very plausible that J decreases with increasing r_{12} and is largest when 1 and 2 coincide. In the latter case, J is equal to I_3.

[18] A similar result has been obtained from perturbation theory by C. G. Kuper, Proc. Roy. Soc. (London) 233, 223 (1955). As he points out, the perturbation theory is not reliable because of the large size of the energy change.

have the correct limiting behavior.[19] A more direct way of seeing the result is to look at (21) [or (20)] when k is very small. The correlation term $g(\mathbf{r}_{ij})$ is significant only when atoms i and j are fairly close. But in this case $\exp(i\mathbf{k}\cdot\mathbf{r}_i)$ and $\exp(i\mathbf{k}\cdot\mathbf{r}_j)$ are almost equal because k is small, and hence the correlation terms cancel almost completely because g is odd. Thus, (21) is almost the same as (5) for small k, and leads to the same energy.

For high k, $E_2(k)/E_1(k)$ approaches unity because the approximation $\exp[i\sum g(\mathbf{r}_{ji})]\simeq 1+i\sum g(\mathbf{r}_{ji})$ fails badly. We noted earlier that if we could compute with the wave function (20), the interference between terms with different i would vanish when k is large. If $E_3(k)$ is the energy arising from (20), we should find that for large k,

$$E_3(k)/E_1(k)=0.65,$$

as in the foreign atom problem. It is amusing to conjecture on how much $E_3(k)$ might lie below $E_2(k)$ when $k=1.8$ A^{-1}. The accuracy of the approximation $\exp(i\sum g)=1+i\sum g$ in the foreign atom problem (see reference 11) suggests that E_3 may be 0.5° less.

The energy spectrum $E_2(k)$ is shown in Fig. 6 as curve A. We have also plotted B: $E_1(k)=\hbar^2k^2/2mS(k)$; C: de Klerk, Hudson, and Pellam's spectrum [Eq. (4)]; D: spectrum of the form (2), with $\Delta/\kappa=9.6°$, $p_0/\hbar=1.85$ A^{-1} and μ chosen so that $\mu^{\frac{1}{2}}p_0^2$ has the same value as in C. (The specific heat depends on μ and p_0 only through the product $\mu^{\frac{1}{2}}p_0^2$.) From the curvatures of A, C, and D it is clear that our spectrum $E_2(k)$ predicts too small a value of μ. In a computation of this sort, however, it is doubtful that the curvature has much significance.

Curve A brings out the fact that the "hump" between the phonon and roton regions is not nearly so high as one might expect from (1). Consequently, when computing the specific heat or normal fluid density at temperatures high enough to excite rotons, it is probably also necessary to take into account the deviations of the phonon spectrum from linearity (and also the deviations of the roton spectrum from pure parabolic behavior). Qualitatively, it appears that such corrections might improve the agreement between the theoretical spectrum and the specific heat and second sound data.

7. DISCUSSION OF ACCURACY

Initially, the major potential sources of error in this computation were (a) the absence of information about the true form of $\rho_3(1,2,3)$; (b) absence of information about $\rho_4(1,2,3,4)$; (c) uncertainties in the data for $S_1(k)$ at large k (see Appendix A).

The uncertainty caused by (a) has, we think, been minimized by the introduction of a correction to the Kirkwood approximation. The errors remaining in I_3 and I_9 after the correction are probably less than three percent; the resultant error in Δ/κ is less than 0.3°.

[19] If $g(\mathbf{r})$ falls off sharply at large r, the analysis is simple. In our case the analysis is complicated by the slowness with which $\mathbf{k}\cdot\mathbf{r}/r^3$ falls off, but the ultimate result is the desired one.

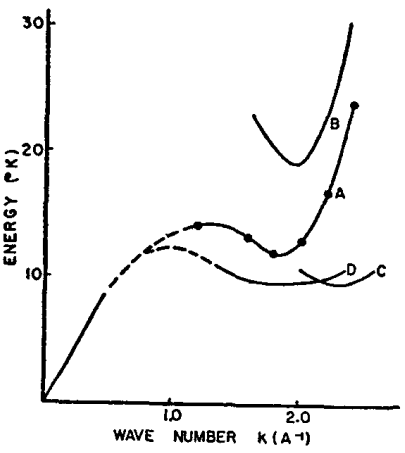


FIG. 6. The energy spectrum of excitations. Curve A is the spectrum $E_2(k)$ computed from Eq. (61). Curve B is the spectrum $E_1(k)$ computed with the simpler wave function (5). Curve C is the Landau-type spectrum used by de Klerk *et al.*[4] to fit the second sound and specific heat data. Curve D is a Landau-type spectrum with p_0 taken the same as in A, and μ and Δ chosen to fit the specific heat data. For small k, all curves are asymptotic to the line $E=\hbar ck$.

The approximation (59), which gives rise to the error (b), ought to be about as accurate as the approximation $I_6\simeq I_{6a}$, since both approximations are based on the same oscillation argument. The latter approximation was found accurate to better than 10% in the roton region. A ten percent change in I_{10} would alter the value of Δ/κ by 0.2°; we regard this number as a fair estimate of the error caused by (b).

Considerable pains were taken to arrange the numerical work in such a way that the answers are insensitive to the behavior of $S_1(k)$ for large k. The residual error due to (c) is found mainly in the coefficient of A^2 in the numerator of (61). This coefficient may be in error by 5%, and the resulting error in Δ/κ might be as much as 0.4°.

We consider the value $\Delta/\kappa=11.5°$ to be accurate within 0.6°, i.e., the lowest energy computable with the wave function (21) is between 10.9° and 12.1°.

A wave function which gives a good value of the energy may, of course, be inaccurate for calculation of other properties of the system. On the other hand this function was chosen by a physical argument, and achieved a very considerable increase in the accuracy of the energy, without in fact using any variable parameters. It might be argued that some of this increase should be associated simply with the fact that we have one extra parameter A to vary. But had we used the A determined by the physical argument (-3.6) we would have obtained practically the same energy as if we let it vary.

For this reason we believe that the wave function (20) [or for practical calculations (21)] not only gives the energy well but is a reasonably accurate physical description of the excitations. On the basis of this optimistic hope, (21) is currently being employed in the calculation of other properties of helium.

APPENDIX A. DATA USED FOR $S(k)$ AND $p(r)$

The curve for $S(k)$ given in Fig. 1 is essentially that obtained from x-ray scattering by Reekie and Hutchison.[8,9] The proper normalization of the data can, in principle, be determined from the fact that $S(k) \to 1$ as $k \to \infty$. According to Goldstein and Reekie,[9] "limitations inherent in the very low scattering cross section of liquid helium and the experimental technique have prevented effective exploration (of the range $k>6$ A^{-1})." Since $S(k)$ is still oscillating strongly at $k=6$ A^{-1}, the normalization of $S(k)$ is uncertain by a few percent. For $k \gtrsim 2.5$ A^{-1}, the percent error in $S(k)-1$ is large, and our computations would be totally unreliable if the integrals had not been set up in such a way as to be insensitive to the behavior of $S(k)-1$ for large k. We feel that $S(k)$ ought to oscillate about its asymptotic value, and have therefore taken $S(k)=1$ at a point whose ordinate is the average of the values of $S(k)$ at the minimum near 3.4 A^{-1} and the maximum near 4.6 A^{-1}. With Reekie's normalization, $S(k)$ is unity at an ordinate much nearer to the minima of the oscillations. Our normalization maximizes the cancellation at large k when we are performing integrals whose integrand contains $S(k)-1$ as a factor. Since $S(k)$ is the Fourier transform of $p(r)$, we find [see (34)]

$$-2\pi^2\rho_0=\int_{0^+}^{\infty} k^2[S(k)-1]dk. \tag{64}$$

The relation (64) might serve as a test of the normalization of $S(k)$, were it not for the fact that the numerical integral gives no sign of converging if we cut it off at $k=6$ A^{-1}. The left side of (64) is equal to -0.43 A^{-3}. With our normalization, integration of the right side out to $k=6$ A^{-1} gives $+0.44$ A^{-3}, but the integrand is still oscillating wildly and there is a chance of ultimately converging to a correct answer. With Reekie's normalization, integration of the right side out to $k=6$ A^{-1} gives a positive value much larger than $+0.44$ A^{-3}, and the contribution from $k>6$ A^{-1} will also be positive unless the successive minima of $S(k)$ cease to be closer and closer to the asymptotic value of unity. At any rate, the consistency of the results which we have obtained by performing the same integral in coordinate and momentum space convinces us that our $S(k)$, which is 0.97 times Reekie's, is sufficiently accurate for the present computations.

Most of the curve in Fig. 1 represents data taken at 2.06°K. According to reference 9, there is very little change in the values of $S(k)$ for $k>0.9$ A^{-1} as the temperature decreases from 2.5°K to 1.25°K. Therefore, in the range $k>0.9$ A^{-1}, it is probably safe to represent the zero-temperature structure factor $S(k)$ by the data taken at 2.06°K. For $k<0.9$ A^{-1}, the temperature dependence of $S(k)$ is more important, and it is necessary to extrapolate $S(k)$ linearly to zero by using (65). We have done this, using a slope about 20% higher than the theoretical value in order to join the experimental data smoothly. The error thus introduced is small.

Reekie and Hutchison[8] have computed $p(r)$ for $r \leq 6$ A by inverting their data for $S(k)$. The curve for $p(r)$ which we have given in Fig. 2 is obtained from one of their graphs[20] and, as has been previously mentioned, seems consistent with our curve for $S(k)$. The numerical inversion of diffraction data is not unambiguous, since the integrand of the relevant numerical integral is not small at the cut-off value $k=6$ A^{-1}. Furthermore, an arbitrary cutoff procedure must be used to make $p(r)$ vanish for $r<2.4$ A. More recently, Goldstein and Reekie[9] have employed an IBM 701 calculator to compute $p(r)$ out to 20 A, using the data of Reekie and Hutchison. Their article was not published until after the completion of the present calculation; the authors state that the results out to 6 A "fully confirm" the results of reference 9. Goldstein and Reekie apply the integral test (69) to their curves for $p(r)$ and find satisfactory results. Since the integrands do not become small until $r \gtrsim 13$ A, we found it impossible to apply the test to the curve in Fig. 2.

APPENDIX B. IDENTITIES SATISFIED BY $p(r)$ AND $S(k)$

To understand the behavior of $S(k)$ for small k, we note that as long as we are concerned with disturbances of long wavelength (small k) the liquid can be treated as a continuous compressible medium. If $\rho(r,t)$ is the number density in such a medium and we define the normal coordinates

$$q_{\mathbf{k}}=\int \rho(r,t)e^{i\mathbf{k}\cdot\mathbf{r}}d\mathbf{r},$$

then the energy is

$$E=\tfrac{1}{2}\sum_{\mathbf{k}} m_{\mathbf{k}}[\dot{q}_{\mathbf{k}}\dot{q}_{\mathbf{k}}^*+\omega_{\mathbf{k}}^2 q_{\mathbf{k}}q_{\mathbf{k}}^*],$$

where $\omega_{\mathbf{k}}=ck$ and $m_{\mathbf{k}}=m/Nk^2$. Quantum mechanically, $\rho(\mathbf{r})$ is replaced by the operator $\sum_i \delta(\mathbf{r}-\mathbf{r}_i)$ and $q_{\mathbf{k}}$ then goes over into the quantum-mechanical normal coordinate $q_{\mathbf{k}}=\sum \exp(i\mathbf{k}\cdot\mathbf{r}_i)$. $S(k)$ is just $1/N$ times the expectation value of $q_{\mathbf{k}}^2$. Since the average values of the potential and kinetic energies are equal for a harmonic oscillator, it follows that $S(k)=\langle E_{\mathbf{k}}\rangle/mc^2$, where $\langle E_{\mathbf{k}}\rangle$ is the average energy of the oscillator representing sound of wave number $\mathbf{k}$. When $T=0$, all the oscillators are in their ground states, and hence $\langle E_{\mathbf{k}}\rangle=\tfrac{1}{2}\hbar\omega_{\mathbf{k}}=\tfrac{1}{2}\hbar ck$ and

$$S(k)=\hbar k/2mc \quad \text{(small } k\text{)}. \tag{65}$$

When $T\neq 0$, the oscillator representing phonons of wave number $\mathbf{k}$ is no longer necessarily in its ground

[20] Figure 2 was obtained by dividing the data of reference 8, Fig. 1, by r^2. There are slight discrepancies between the resultant curve for $p(r)$ and the curve given in reference 9, Fig. 3. Errors of this magnitude in $p(r)$ would have a negligible effect on our results.

state, but may be in its nth excited state with probability proportional to $\exp(-E_n/\kappa T)$. It follows that $(\beta=1/\kappa T)$

$$S(k)=(\hbar k/2mc)\coth\tfrac{1}{2}\beta\hbar ck \tag{66}$$

$$=(\beta mc^2)^{-1}+(\beta\hbar^2/12m)k^2-\cdots. \tag{66a}$$

From (66) there follows immediately the famous formula

$$\lim_{k\to 0}S(k)=\rho_0\kappa T\chi_T, \tag{67}$$

where ρ_0 is the number density and χ_T the isothermal compressibility of the liquid. When $\hbar ck$ becomes greater than κT, (66) becomes essentially linear in k. Strictly speaking, however, $S(k)$ starts quadratically from a nonzero value except when $T=0$. The possibility of a linear behavior of $S(k)$ for small k, as predicted by (65) when $T=0$, has been sometimes questioned on the basis of (10). From (10) it follows that

$$S(k)-1=4\pi\int_{0^+}^{\infty}[p(r)-\rho_0](kr)^{-1}\sin(kr)r^2dr. \tag{68}$$

Since $p(r)-\rho_0$ approaches zero for large r, it is argued that it is legitimate to expand $(kr)^{-1}\sin kr$ as $1-(kr)^2/6+\cdots$. Integrating term by term, one finds

$$S(k)-1=C_1+C_2k^2+\cdots, \tag{68a}$$

where

$$C_1=4\pi\int_{0^+}^{\infty}[p(r)-\rho_0]r^2dr,$$

$$C_2=-\tfrac{2}{3}\pi\int_{0^+}^{\infty}[p(r)-\rho_0]r^4dr.$$

Hence it appears that $S(k)$ always starts quadratically in k. The fallacy in the argument lies in the fact that $p(r)$ may not approach its asymptotic value fast enough, and the expansion may be meaningless. For example, if $p(r)-\rho_0$ decreases as r^{-3} for large r, (68) converges perfectly well but C_1 and C_2 are infinite. When $T=0$, $p(r)-\rho_0$ falls off slowly enough to invalidate the expansion, and (65) is correct; at any finite temperature $p(r)-\rho_0$ ultimately falls off exponentially and the expansion (68a) is legitimate. One might think that all the coefficients of (68a) can be determined by comparison with (66a); this is incorrect because (66) is wrong for large k. Using (67) and (68), however, we do obtain the important result

$$1+4\pi\int_{0^+}^{\infty}[p(r)-\rho_0]r^2dr=\rho_0\kappa T\chi_T, \tag{69}$$

and when $T=0$

$$1+4\pi\int_{0^+}^{\infty}[p(r)-\rho_0]r^2dr=0. \tag{70}$$

The result (69) can also be obtained by rather simple classical arguments. It follows directly from the definition of $p(r)$ that the left side of (69) is $\langle(N-\bar{N})^2\rangle_{\text{Av}}/\bar{N}$, where N is the number of atoms in a large subvolume of the liquid, and the bar denotes "average," but statistical mechanics shows that $\langle(N-\bar{N})^2\rangle_{\text{Av}}/\bar{N}=\rho_0\kappa T\chi_T$, whence (69) follows.

One might think that (70) is a simple consequence of the definition of $p(r)$. For if an atom is known to be at $\mathbf{r}_1$, the probability that there is an atom at $\mathbf{r}_2$ is $\rho_2(\mathbf{r}_1,\mathbf{r}_2)/\rho_1(\mathbf{r}_1)$. If $\mathbf{r}_1$ is not near the surface of the liquid, then $\rho_1(\mathbf{r}_1)=\rho_0$; if $\mathbf{r}_1$ and $\mathbf{r}_2$ are both far from the surface, then $\rho_2(\mathbf{r}_1,\mathbf{r}_2)=\rho_0 p(r_{12})$. If we integrate $\rho_2(\mathbf{r}_1,\mathbf{r}_2)/\rho_1(\mathbf{r}_1)$ over all locations $\mathbf{r}_2$, excluding the point $\mathbf{r}_1$, the answer must be exactly $N-1$. But if we integrate $\rho_1(\mathbf{r})$ over all positions of $\mathbf{r}$, the answer is exactly N. Consequently

$$\int_{\mathbf{r}_2\neq\mathbf{r}_1}[\rho_2(\mathbf{r}_1,\mathbf{r}_2)/\rho_1(\mathbf{r}_1)-\rho_1(\mathbf{r}_2)]d\mathbf{r}_2=-1. \tag{71}$$

If we take $\mathbf{r}_1$ far from the surface, $\rho_1(\mathbf{r}_1)$ can be replaced by ρ_0. Furthermore, the integrand is appreciable only when $\mathbf{r}_2$ is near $\mathbf{r}_1$, in which case $\rho_2(\mathbf{r}_1,\mathbf{r}_2)/\rho_1(\mathbf{r}_1)-\rho_1(\mathbf{r}_2)=p(r_{12})-\rho_0$ (there are no complications at the surface of the liquid since the surface corrections to both terms of the integrand are identical). Then (71) reduces exactly to (70).

Something must be wrong with the preceding argument at finite temperatures, since (70) is false if $T\neq 0$. The difficulty lies in the fact that, at finite T, the limiting value of $\rho_2(\mathbf{r}_1,\mathbf{r}_2)/\rho_0$ for large $\mathbf{r}_{12}$ is not ρ_0, but is slightly lower by an amount of order $1/N$. Since $\mathbf{r}_2$ runs over a volume proportional to N, there is a finite negative contribution to (71) from the region of very large r_{12} (i.e., the region where $\rho_2(r_1,r_2)/\rho_0$ has reached its asymptotic value, which is not exactly equal to ρ_0). Since (71) is rigorously true, the integral

$$4\pi\int_{0^+}^{\infty}[p(r)-\rho_0]r^2dr,$$

which represents the contribution to the left side of (71) from the region where r_{12} is not very large, must be greater than -1; thus we arrive at (69) instead of (70).

The slight lowering of the density at infinity when an atom is known to be at the origin is not hard to understand, since the localization of one atom decreases by one the number of atoms eligible to occupy the site at infinity. In the classical perfect gas $\rho_0\kappa T\chi_T=1$; since the atoms are independent, the localization of one atom simply lowers the mean density by $1/V$ throughout the rest of the volume. In a real liquid, however, $\rho_0\kappa T\chi_T\to 0$ as $T\to 0$. Finally, when $T=0$, (70) implies that no influence propagates to infinity, even in order $1/N$, when an atom is localized at the origin. In this case, a

density excess at the origin is surrounded by a rarefaction slightly further away, so that no change occurs in the density at infinity.

Thus, the simple counting argument used to prove (70) is actually correct when $T=0$, because there is no change in the density far away when we localize an atom at the origin. For the same reason, we believe that any identity based on a counting argument becomes correct when $T=0$. We therefore believe in the truth of the identity

$$\int_{\mathbf{r}_3\neq\mathbf{r}_2,\mathbf{r}_1} d\mathbf{r}_3\left[\frac{\rho_3(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)}{\rho_0 p(r_{12})}-\rho_0\right]=-2, \qquad (72)$$

although we cannot give a rigorous proof of it. Equations (70) and (72) are easily combined to give Eq. (39), which we have used in our work (one must remember that, in (39), ρ_3 is defined to include a delta function on $\mathbf{r}_{23}$).

Equation (39) is easily understood for small or large values of r_{12}. If $r_{12}<2.4$ A, both sides of (39) are identically zero for all $\mathbf{r}_3$. If 1 and 2 are far apart, then ρ_3 can be written as

$$\rho_0 p_1(r_{31})p_1(r_{32})+\rho_0^2\delta(\mathbf{r}_{32})$$

and the right side becomes $\rho_0^2 p_1(r_{31})$. If 3 is far from 2, then both sides are equal. Hence the only contribution to the integral comes when 3 is near 2; but then we can set $p_1(r_{31})=\rho_0$ and we are left with

$$\int d\mathbf{r}_{32}\rho_0^2[p_1(r_{32})+\delta(\mathbf{r}_{32})-\rho_0],$$

which vanishes as a result of (70).

Even if (39) is not rigorously true for intermediate values of r_{12}, it cannot fail badly; for when r_{12} is greater than 2.4 A, but not very large, then for any fixed radius r_{32} the solid angle in which 3 interferes with 1 is small (less than one-quarter of the total solid angle available to $\mathbf{r}_{23}$).

PHYSICAL REVIEW VOLUME 107, NUMBER 1 JULY 1, 1957

Theory of Inelastic Scattering of Cold Neutrons from Liquid Helium

MICHAEL COHEN* AND RICHARD P. FEYNMAN
California Institute of Technology, Pasadena, California
(Received March 25, 1957)

A measurement of the energy losses of monoenergetic neutrons scattered from liquid He II would permit a determination of the energy-*versus*-momentum relation for the elementary excitations (phonons and rotons) in the liquid. A major part of the scattering at a fixed angle arises from production or annihilation of a single excitation and appears as sharp lines in the energy spectrum. From the position of these lines the energy-*versus*-momentum relation of the excitations can be inferred. Other processes, such as production or annihilation of multiple excitations, contribute a continuous background, and occur at a negligible rate if the incident neutrons are slow ($\lambda \gtrsim 4$A) and the helium cold ($T<2$°K). The total cross-section data can be accounted for by production of single excitations; the theoretical cross section, computed from a wave function previously proposed to represent excitations, agrees with experiment over the entire energy range, within 30%. Line widths in the discrete spectrum are negligible at 1°K because of the long lifetime of phonons and rotons.

I. INTRODUCTION

THE possibility of a direct experimental determination of the energy-*versus*-momentum relation for phonons in a solid was pointed out by Placzek and Van Hove.[1] They proposed to study the energy distribution of very slow neutrons scattered inelastically and coherently from the solid; if the incident neutron beam is monochromatic and if the scattering process involves only the production or annihilation of a single phonon, energy and momentum conservation imply that the neutrons emerging at a given angle can have only certain discrete energies. The energy-momentum relation for the phonons can be inferred from the angular variation of this discrete spectrum. Other processes, such as multiple phonon production or annihilation, contribute a continuous background above which the discrete spectrum is still observable.

The purpose of the present paper is to suggest that the same technique be used to determine directly the energy-*versus*-momentum curve for the excitations in liquid helium, and to predict some details of the experiment. A direct measurement of this curve would be of considerable interest, since the shape of the curve has already been predicted in some detail by indirect methods. Landau[2] argued on theoretical grounds that the energy $E(k)$ of an excitation momentum $\hbar\mathbf{k}$ should rise linearly with slope $\hbar c$ for small k (c=speed of sound =240 m/sec), pass through a maximum, drop to a local minimum at some value k_0, and rise again when $k>k_0$. For small k, the excitations are called phonons and may be thought of as quantized sound waves; the excitations with $k\sim k_0$ are called rotons, and seem to be the quantum-mechanical analog of smoke rings.[3,4] At low temperatures, only the linear portion of the curve and the portion near the minimum are excited; if the curve is represented near the minimum by $E(k)=\Delta+\hbar^2(k-k_0)^2/2\mu$, the specific heat and second sound data can be fitted best with the values[5]

$$\Delta/\kappa=9.6°\text{K}, \quad k_0=2.30\,A^{-1}, \quad \mu=0.40\,m_{\text{He}},$$

and almost as well with the values[2]

$$\Delta/\kappa=9.6°\text{K}, \quad k_0=1.95\ \text{A}^{-1}, \quad \mu=0.77\,m_{\text{He}}.$$

A Landau-type curve has recently been obtained from first principles by the substitution of a trial function into a variational principle for the energy.[3,4] The resulting curve is an upper limit to the true spectrum, and gives $\Delta/\kappa=11.5$°K, $k_0=1.85$ A^{-1}, $\mu\sim 0.20$

* Richard C. Tolman Fellow.

[1] G. Placzek and L. Van Hove, Phys. Rev. **93**, 1207 (1954). We have recently learned that some of the ideas in the present paper have been discussed by V. V. Tolmachev, Repts. Acad. Sci. U.S.S.R. **101**, No. 6 (1955).

[2] L. Landau, J. Phys. (U.S.S.R.) **5**, 71 (1941); **11**, 91 (1947).

[3] R. P. Feynman, Phys. Rev. **94**, 262 (1954).

[4] R. P. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956).

[5] deKlerk, Hudson, and Pellam, Phys. Rev. **93**, 28 (1954).

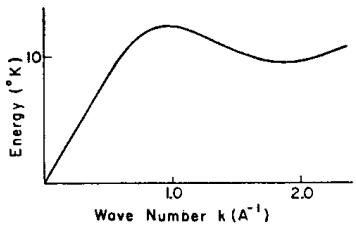


FIG. 1. A Landau-type energy-*versus*-momentum curve, with $\Delta=9.6°K$, $\mu=1.06\ m_{He}$, $k_0=1.85\ A^{-1}$.

m_{He}. In the rough computations of this paper we shall use the curve of Fig. 1, which has $\Delta/\kappa=9.6°K$, $k_0=1.85$ A^{-1}, $\mu=1.06\ m_{He}$. These values represent a compromise between theory and experiment, and also fit the specific heat data. Most of our numbers have only a qualitative significance, since the shape of the energy curve between the phonon and roton regions is highly uncertain.

II. GENERAL THEORY

Suppose the liquid is initially in state j, and we bombard it with neutrons of mass m and momentum $\hbar\mathbf{k}_i$; the cross section for a process in which the liquid is left in one of a group of final states F, and the neutron emerges with a momentum in some region G of $\mathbf{k}$ space, is given by the Born approximation[6,7] as

$$\sigma=\frac{2a^2}{k_i}\sum_{f\epsilon F}\int_{\mathbf{k}\epsilon G}|V_{fj}(\mathbf{k}-\mathbf{k}_i)|^2 \times\delta\left(k^2-k_i^2+\frac{2m}{\hbar^2}(E_f-E_j)\right)d\mathbf{k}, \quad (1)$$

where the sum and the integral extend over the regions F and G, respectively. The matrix elements V are given by

$$V_{fj}(\mathbf{q})=\int\psi_f^*(\mathbf{r}_1,\cdots\mathbf{r}_N)\sum_{i=1}^{N} \times\exp(i\mathbf{q}\cdot\mathbf{r}_i)\psi_j(\mathbf{r}_1,\cdots\mathbf{r}_N)d\mathbf{r}_1\cdots d\mathbf{r}_N,$$

where ψ_j and ψ_f are the wave functions for the liquid in states j and f. The scattering length a is independent of energy at low energies, and is related to the total cross section σ' of a bound He nucleus by $\sigma'=4\pi a^2$. McReynolds[8] found $\sigma'=1.1\pm0.15$ barns.

In elastic scattering, the final state of the liquid is the same as the initial state.[9] In the matrix element for elastic scattering, the integral of ψ_f^2 over all coordinates but one is equal to $1/V$ (V=volume of the liquid) except at points very close to the surface. Only the region near the surface contributes to the volume integral of $\exp(i\mathbf{q}\cdot\mathbf{r})$, and hence the only elastic scattering from the liquid is diffraction from the surface.[10] In a crystal, the integral of ψ_j^2 over all coordinates but one gives a function which is strongly peaked at the lattice points; hence for certain directions of $\mathbf{q}$ the matrix element $V_{fj}(\mathbf{q})$ becomes proportional to N, and elastic scattering occurs. Therefore, although neutrons are scattered elastically from solids, virtually no elastic scattering should occur from the liquid.

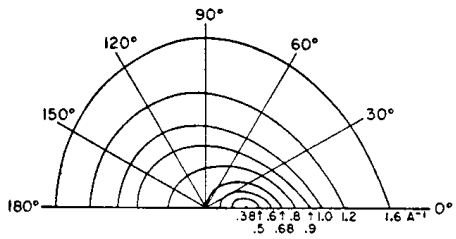


FIG. 2. Kinematics of production of single excitations. The curves give the wave number k_f of the exit neutron, as a function of k_i and θ; k_f is the distance from the origin. On each curve k_i is constant and has the value given on the $\theta=0$ axis. Note that $k_f=k_i$ when $\theta=0$ (and $k_i>0.68$ A^{-1}).

III. SCATTERING AT ZERO TEMPERATURE

If the liquid is at zero temperature, then the initial state is the ground state ψ_0, and the neutron must lose energy in the scattering. The simplest process which can occur is the creation of a single excitation of momentum $\hbar(\mathbf{k}_i-\mathbf{k}_f)$ in the liquid (if we impose periodic boundary conditions on the liquid, the stationary states may be taken as momentum eigenstates). Energy conservation requires

$$\hbar^2k_i^2/2m=\hbar^2k_f^2/2m+E(|\mathbf{k}_i-\mathbf{k}_f|). \quad (2)$$

If we fix $\mathbf{k}_i$ and the angle θ between $\mathbf{k}_i$ and $\mathbf{k}_f$, then if $k_i>0.68$ A^{-1} ($\lambda_i<9.25$ A) there is a unique k_f for each θ. When k_i is just less than 0.68 A^{-1}, (2) becomes insoluble for $\theta>90°$. As k_i decreases further, the region of solubility of (2) is a cone of decreasing aperture about the forward direction. For each direction θ in the cone there are two solutions for k_f. Finally, when $k_i<0.38$ A^{-1}, (2) becomes insoluble at any angle. The qualitative behavior of the solutions of (2) is shown in Fig. 2, but not too much significance should be attached to the numbers, which are based on the uncertain curve of Fig. 1.

The solutions of (2) should appear as lines in the

[6] G. Placzek, Phys. Rev. 86, 377 (1952).

[7] For the justification of the use of the Born approximation in this problem see E. Fermi, Ricerca sci. 7, 13 (1936); G. Breit, Phys. Rev. 71, 215 (1947).

[8] A. W. McReynolds, Phys. Rev. 84, 969 (1951).

[9] This is true if the liquid is confined in a fixed box. If the box is free to recoil, then for elastic scattering the final state is the same as the initial state except for a translational motion of the whole with momentum $\hbar(\mathbf{k}_i-\mathbf{k}_f)$ and infinitesimal energy $\hbar^2(\mathbf{k}_i-\mathbf{k}_f)^2/2Nm_{He}$. The wave function for the final state is then $\psi_f=\psi_i\exp[iN^{-1}(\mathbf{k}_i-\mathbf{k}_f)\cdot\Sigma\mathbf{r}_i]$. The remaining arguments are still valid, with each $\mathbf{r}_i$ being measured from the center of mass rather than from an origin determined by the location of the box.

[10] By "elastic scattering" we mean scattering processes in which the incident and exit neutrons have the same energy, and furthermore the state of the liquid is unchanged. If we relax the latter requirement, then at finite temperature some neutrons can scatter without energy loss by colliding with excitations in the liquid. However, neutrons which are scattered by collisions with excitations emerge with a continuous distribution of energies (i.e., the number in any energy range dE is proportional to dE), and there will not be any "group" of elastically scattered neutrons.

energy spectrum of the neutrons emerging at a given angle. The energy-*versus*-momentum curve for the excitations can be obtained from Eq. (2) by measuring k_f as a function of angle for fixed $\mathbf{k}_i$, or by looking at a fixed exit angle and varying k_i. Processes involving multiple excitations contribute a continuous background. When $k_i<0.38\,A^{-1}$, the right side of (2) is bigger than the left for any k_f; furthermore, since $E(k)$ is the energy of the *lowest* state of the liquid having momentum $\hbar\mathbf{k}$, production of multiple excitations will also be impossible when $k_i<0.38$ A^{-1}. Hence, when the liquid is at zero temperature, neutrons with $k_i<0.38$ A^{-1} ($\lambda>16.5$ A) should pass through with no scattering. This conclusion seems consistent with the data of Sommers, Dash, and Goldstein[11] on the transmission of neutrons by He.

The strengths of the lines are given by (1). We take the region G of k space as $k^2\Delta k d\Omega$, where Δk includes the line under study. Momentum conservation manifests itself in the vanishing of $V_{f0}(\mathbf{k}-\mathbf{k}_i)$ unless the state f has the momentum $\hbar(\mathbf{k}_i-\mathbf{k})$. Integration of (1) over G gives

$$\frac{d\sigma_1}{d\Omega}d\Omega=a^2\frac{k_f}{k_i}\frac{|V_{f0}(\mathbf{k}_f-\mathbf{k}_i)|^2 d\Omega}{\left|1+\frac{m}{\hbar^2}\frac{E'(|\mathbf{k}_f-\mathbf{k}_i|)}{|\mathbf{k}_f-\mathbf{k}_i|}\left(1-\frac{k_i}{k_f}\cos\theta\right)\right|} \tag{3}$$

as the cross section for scattering into $d\Omega$ with the production of one excitation of momentum $\hbar(\mathbf{k}_i-\mathbf{k}_f)$ (and final neutron momentum $\hbar\mathbf{k}_f$). In reference 3 the function

$$\psi_k=\mathcal{S}^{-\frac{1}{2}}\psi_0\sum\exp(i\mathbf{k}\cdot\mathbf{r}_i) \tag{4}$$

is proposed to represent a single excitation of momentum $\hbar\mathbf{k}$. This function is exact for very small k, and gives an energy spectrum qualitatively similar to Landau's, but with Δ twice too large. Normalization requires $\mathcal{S}=NS(k)$, where $S(k)$ is the Fourier transform of the zero-temperature radial distribution function $p(r)$,

$$S(k)=\int\exp(i\mathbf{k}\cdot\mathbf{r})p(r)d\mathbf{r}.$$

The resulting matrix element is

$$|V_{f0}(\mathbf{q})|^2=NS(q). \tag{5}$$

Actually, (5) is an over-estimate, as one can see from the exact sum rule

$$\sum_f|V_{f0}(\mathbf{q})|^2=[V(\mathbf{q})V^*(\mathbf{q})]_{00}=NS(q). \tag{6}$$

If (5) were exact, then (6) would imply that production of multiple excitations is impossible. A more accurate wave function for an excitation is given in reference 4, and leads to the matrix elements given in Fig. 3. In the roton region, these matrix elements are only ten to fifteen percent smaller than those given by (5), and

[11] Sommers, Dash, and Goldstein, Phys. Rev. **97**, 855 (1955).

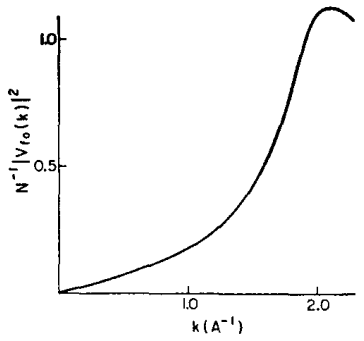


FIG. 3. Matrix elements for production of single excitations, computed from the wave functions of reference 4.

we infer that the most likely way for a neutron to lose a given amount of momentum is through the production of a single excitation.

Substituting the matrix elements of Fig. 3 into Eq. (3), we obtain the curves of Fig. 4, giving the angular variation of line strength for different values of k_i. The curves are given in units of a^2, which is the differential cross section per unit solid angle for scattering from a bound helium nucleus. When $k_i>0.68$ A^{-1}, $d\sigma_1/d\Omega$ vanishes at $\theta=0$ because the matrix element V_{f0} approaches zero when the momentum transfer is small. When $k_i<0.68$ A^{-1}, there are two curves for each k_i, corresponding to the two lines which are observed at each angle within the cone of solubility of (2). At the edges of this cone, $d\sigma_1/d\Omega$ becomes infinite because the denominator of (3) vanishes. The total cross section, however, is finite.

If we neglect the possibility of producing multiple excitations, the total cross section at zero temperature is obtained by integrating (3) over angles. The resulting cross sections are compared in Fig. 5 with the total cross sections measured[11] at 1.25°K (the temperature effect, which is negligible, is discussed in the next section). The agreement of theory and experiment within 30% is quite satisfactory in view of our incomplete knowledge of the wave function ψ_f and the curve $E(k)$. When k_i is large, Eq. (1) can be shown to lead to the total cross section $(16/25)4\pi a^2N$, which is just the cross section for free helium nuclei scattering incoherently. The theoretical curve has been extra-

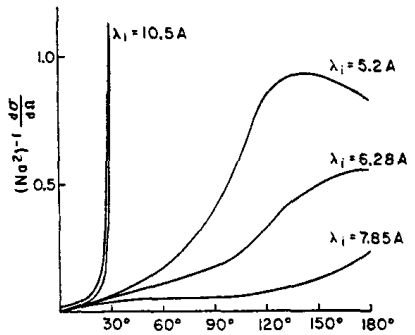


FIG. 4. Angular distribution of neutrons which have been scattered by the process of producing a single excitation, at zero temperature.

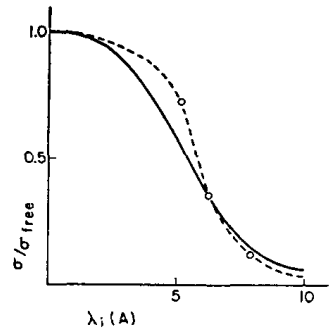


FIG. 5. The broken line is the total cross section computed here. Circles are computed points. The solid line represents the measurements of Sommers *et al.* at 1.25°K.

polated to this value. As λ decreases, the theoretical curve rises faster than the experimental one; the reason is probably that the matrix elements of Fig. 3 are too large when the momentum transfer is in the roton region. To see this more clearly, we note that there is an exact sum rule,

$$\sum_f |V_{0f}(\mathbf{k})|^2(E_f-E_0)=[V^*(\mathbf{k})(H-E_0)V(\mathbf{k})]_{00} =N\hbar^2k^2/2m, \quad (7)$$

which, in conjunction with Eq. (6), says that the "average" energy loss associated with momentum transfer k is $\hbar^2k^2/2mS(k)$. For small k, this "average" energy loss is the same as $E(k)$; hence the idea that multiple excitations are produced with negligible probability is correct. However, when $k=1.85$ A^{-1}, one finds that the "average" energy loss is 19.5°K, which is twice the size of $E(k)$. If the matrix elements of Fig. 3 are correct, then the sum rule (6) implies that there is only a 13% probability of producing multiple excitations when the momentum transfer is 1.85 A^{-1}. Such a small probability of multiple excitation seems hardly consistent with a mean energy loss corresponding to the production of two rotons. We conclude that the correct matrix elements for single roton production are almost certainly smaller that those given in Fig. 3.[12]

IV. EFFECTS OF FINITE TEMPERATURE

If the liquid is at a finite temperature, some phonons and rotons are always present, and the neutrons can gain energy by annihilating excitations. The energy spectrum of the neutrons emerging at a given angle will contain a line representing annihilation of single excitations as well as a line or lines arising from their production. At temperatures below the λ point, the production process is far more important than annihilation if the wavelength of the incident neutrons is less than 10 A. Since phonons and rotons obey Bose statistics, the rate of annihilation of excitations of momentum $\hbar\mathbf{k}$ is proportional to the number $n(\mathbf{k})$ of such excitations already present, while the rate of production is proportional to $n(\mathbf{k})+1$. At temperature T, we have $n(\mathbf{k})=\{\exp[E(\mathbf{k})/\kappa T]-1\}^{-1}$; for rotons at 2°, $E(k)/\kappa T\simeq 5$, and we see that the "spontaneous production" factor 1 is much greater than $n(\mathbf{k})$. Figure 2 shows that if the incident neutron wavelength is less than 10A ($k_i\gtrsim 0.6$ A^{-1}), most of the excitations produced have wave numbers greater than 0.4 A^{-1} and consequently energies large compared to 2°K. Furthermore, Fig. 3 shows that the matrix elements for the production of low-energy phonons are small. Thus we see that in the range $T\lesssim 2$°K, $\lambda\lesssim 10$ A, spontaneous production is much more important than "induced production" and annihilation, and the lack of temperature dependence of the total cross section is understood.

For incident neutrons of wavelength greater than $\sim$10 A, it is kinematically impossible to produce any except very low-energy phonons (and no excitations at all can be produced when $\lambda>16.5$ A). Hence the annihilation process is the most important one at long neutron wavelengths, and the total cross section in this region is strongly temperature-dependent. Ultimately, at very long incident neutron wavelengths, the kinematics becomes that of zero-energy incident neutrons. Figure 2 shows that a zero-energy incident neutron can annihilate only phonons with wave number $k=0.68$ A^{-1} and energy 11°K.[13] The total cross section ultimately depends on the temperature as $\exp(-11/T)$, and on the incident velocity as $1/v$ [arising from the factor $1/k_i$ in Eq. (1)].

V. RESOLUTION AND LINE WIDTH

In order to obtain even a moderately accurate measurement of the roton energy Δ, the velocities of the incident and exit neutrons must be known very accurately. Figure 2 shows that the slowest neutron which can produce a minimum energy roton has a wave number $k\sim 1.04$ A^{-1} (energy=25°); the exit neutron in this case has $k\sim 0.81$ A^{-1} (energy=15°). To measure Δ with an accuracy of one degree, the neutron energies must be accurate to 0.7°; for the incident neutrons, we need $\delta\lambda/\lambda=\delta E/2E=0.7/50=0.014$. Thus, to measure Δ with ten percent accuracy, the velocity spread of the incident neutron beam must be limited to about one percent. Nothing is gained by studying neutrons which have annihilated a roton; the slowest incident neutron which can annihilate a roton has $k\sim 0.81$ A^{-1}, and we need $\delta\lambda/\lambda=0.7/30=0.023$. The slight improvement in the resolution situation is far more than offset by the low rate of annihilation, as compared with production (see Sec. IV). The resolution situation is best when we observe neutrons scattered through 180°. If we study

[12] It is not entirely obvious that the total cross section is lowered by lowering the probability of producing single excitations, since the sum rule (6) implies that the probability of producing multiple excitations must be correspondingly increased. However, the density-of-states factor resulting from the delta function in Eq. (1) produces a decrease in the total cross section when probability is transferred from single to multiple excitations.

[13] This is not entirely correct, since annihilation of multiple excitations is possible, though unlikely at low temperatures. The total momentum of the excitations annihilated must be at least 0.68 A^{-1} and the total energy at least 11°K. At low temperatures and velocities, the cross section still varies as $v^{-1}\exp(-11/T)$.

the 90° scattering, the slowest allowed incident neutron has $k\sim1.5$ A^{-1} and we need $\delta\lambda/\lambda\sim0.007$.

If one looks at the energy distribution of the neutrons emerging at a particular angle, how broad is the line corresponding to those neutrons which have created a roton? We have studied this question in some detail; in the Appendix we compute the detailed line shape which would result if we make certain assumptions about the interaction between phonons and rotons. The assumptions prove to be unrealistic, but the method of computation is of some interest. The result which we obtain for the line width is what one would expect from the uncertainty principle; the width is $\hbar/\tau$, where τ is the lifetime of the roton until it collides with something. In our model, roton-roton interactions are neglected; hence the lifetime we compute is that for roton-phonon collisions. This lifetime is very long, and the resulting width is less than 10^{-6}°K (to be compared with a roton energy $\Delta=9.6$°K) when the helium is at a temperature of 1°K. Landau and Khalatnikov[14] have computed the lifetimes for phonon-phonon, phonon-roton, and roton-roton collisions. They find that at temperatures of 1° and higher, the roton-roton lifetime is much shorter than the roton-phonon lifetime. When $T=1°$, the width of a roton line is 0.006°K, which is still very small compared with the roton energy, but large compared with 10^{-6}°K. The width is proportional to $T^{\frac{1}{2}}\exp(-\Delta/\kappa T)$, which represents the temperature dependence of the number of rotons present; when $T=2$°K, the width is about 1°. Similarly, one can compute the width of a line arising from the production of phonons by neutrons; when $T=1$°K, the lifetimes of a phonon for scattering by a roton or by another phonon are comparable, both giving rise to widths of about 10^{-5}°K. We conclude that for all practical purposes the lines in the neutron spectrum will be true delta functions if the helium temperature is near 1°K.

To calculate the cross section for roton-roton collisions, Landau and Khalatnikov assume a delta-function interaction between rotons, the strength of the interaction being chosen to fit viscosity data. In the appendix we show that such a delta-function potential, with a strength close to that of Landau and Khalatnikov, arises from the possibility that roton I can emit a phonon which is subsequently absorbed by roton II. The result is only suggestive rather than conclusive, however, since we are also led to a velocity-dependent interaction between rotons, comparable in strength with the delta function.

ACKNOWLEDGMENTS

We wish to thank P. J. Bendt, M. Gell-Mann, W. J. Karzas, and H. Palevsky for valuable discussions.

[14] L. Landau and I. M. Khalatnikov, J. Exptl. Theoret. Phys. U.S.S.R. **19**, 637, 709 (1949).

APPENDIX

The problem of the breadth and shape of the lines in the neutron spectrum caused us some confusion, the details of which are not worth recounting. Finally, we constructed a "model" Hamiltonian for helium, including a phonon-roton interaction, for which the line shape can be computed very accurately. Analysis of this Hamiltonian not only resolved our private confusion, but also showed what the line shape is in the case of real helium. The important features of the answer can be obtained from perturbation theory, provided certain linear terms are interpreted as the beginning of exponentials. We regard the more accurate computation as sufficiently interesting to be presented here. It is analogous to the Weisskopf-Wigner method in the theory of optical spectra.

Suppose, for simplicity, that excitations with $k<k_c$ are phonons, with energy $E(k)=\hbar ck$, and excitations with $k>k_c$ are rotons with energy $E(k)=\Delta+\hbar^2(k-k_0)^2/2\mu$. If there were no interaction between phonons and rotons, the Hamiltonian for the liquid would be $H=\sum a_{\mathbf{k}}^*a_{\mathbf{k}}E(k)$, where $a_{\mathbf{k}}^*$ and $a_{\mathbf{k}}$ are the usual creation and destruction operators for excitations of momentum $\hbar\mathbf{k}$; the operator $a_{\mathbf{k}}^*a_{\mathbf{k}}$ has integral eigenvalues $n_{\mathbf{k}}$, which represent the number of excitations present with momentum $\hbar\mathbf{k}$. The matter density at $\mathbf{r}$ is given by

$$\rho(\mathbf{r})-\rho_0=(\rho_0\hbar/2Vc)^{\frac{1}{2}}\sum_{k<k_c}k^{\frac{1}{2}}\times[a_{\mathbf{k}}\exp(i\mathbf{k}\cdot\mathbf{r})+a_{\mathbf{k}}^*\exp(-i\mathbf{k}\cdot\mathbf{r})].$$

We are interested only in the average behavior of the matter density over a region of finite size (the size of the roton) and have therefore omitted wavelengths smaller than $2\pi/k_c$ in the representation of $\rho(\mathbf{r})$. We assume that if a roton of momentum $\hbar\mathbf{k}$ is at $\mathbf{r}$, its energy is $E(k)+(\partial\Delta/\partial\rho)[\rho(\mathbf{r})-\rho_0]$. Atkins and Edwards[15] have measured $\partial\Delta/\partial\rho$ and find $\partial\Delta/\partial\rho=-0.57\Delta/\rho$. The actual interaction between phonons and rotons involves further coupling terms, which will be discussed later.

It is convenient to use a mixed representation in which phonons are treated with creation and destruction operators, and rotons are represented as particles with coordinates and momenta. If the number of rotons present is m, the Hamiltonian is

$$H=\sum_{k<k_c}a_{\mathbf{k}}^*a_{\mathbf{k}}\hbar ck+\sum_{i=1}^{m}E(\mathbf{k}_i)+\frac{\partial\Delta}{\partial\rho}(\rho_0\hbar/2Vc)^{\frac{1}{2}}\sum_{i=1}^{m}\times\sum_{k<k_c}k^{\frac{1}{2}}[a_{\mathbf{k}}\exp(i\mathbf{k}\cdot\mathbf{r}_i)+a_{\mathbf{k}}^*\exp(-i\mathbf{k}\cdot\mathbf{r}_i)]. \quad (1')$$

The interaction term in (1′) creates and destroys single phonons, since the operators $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^*$ appear linearly.

[15] K. R. Atkins and M. H. Edwards, Phys. Rev. **97**, 1429 (1955).

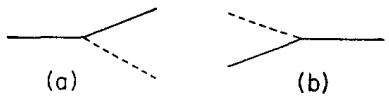


FIG. 6. Diagrams representing the phonon-roton interaction in the Hamiltonian (1′). The solid line is a roton, the broken line a phonon. Time increases from left to right. Diagram (a) represents emission of a phonon, and (b) represents absorption.

Rotons only change their momenta, however, since the interaction merely multiplies the roton wave function by a plane wave. If we represent rotons by solid lines and phonons by dotted lines, the two terms in the phonon-roton interaction can be represented by the diagrams of Fig. 6.

If the number of neutrons per second emerging in solid angle $d\Omega$ with an energy loss in the range $(E, E+dE)$ is $n(E)dEd\Omega$, the Fourier transform of $n(E)$ is given by Eq. (1) as

$$f(\eta)=\int_{-\infty}^{\infty}\exp(-i\eta E)n(E)dE$$
$$=\left(\frac{2a^2\hbar}{m}\right)\int k^2dk\sum_j\exp[-i\eta(E_f-E_j)]$$
$$\times|V_{fj}(\mathbf{k}-\mathbf{k}_i)|^2\delta\left(k^2-k_j^2+\frac{2m}{\hbar^2}(E_f-E_j)\right).$$

Since the energy needed to produce a roton is small compared with the energy of the incident neutrons, the term (E_f-E_j) in the argument of the delta function can be ignored with negligible error. This approximation gets better as the incident neutrons get faster.[16] Similarly, since the direction of $\mathbf{k}$ is fixed and its length is very close to k_i, we can replace $\mathbf{k}-\mathbf{k}_i$ by a constant vector $\mathbf{q}$, where $|\mathbf{q}|=2k_i\sin(\theta/2)$. Furthermore, the liquid may be in different initial states j with probability $\exp(-\beta E_j)/\sum\exp(-\beta E_j)$, where $\beta=1/\kappa T$. Thus we obtain

$$f(\eta)=(a^2\hbar k_i/m)[\textstyle\sum\exp(-\beta E_j)]^{-1}\sum_{i,f}$$
$$\times\exp\{-[\beta E_j+i\eta(E_f-E_j)]\}|V_{fj}(q)|^2$$
$$=(a^2\hbar k_i/m)[\mathrm{Tr}\exp(-\beta H)]^{-1}$$
$$\times\mathrm{Tr}[V\exp(-i\eta H)V^*\exp(-\beta H+i\eta H)]. \quad (2')$$

In Eq. (4) we have suggested that the wave function for a single roton of momentum $\mathbf{q}$ is the ground-state wave function multiplied by $\sum\exp(i\mathbf{q}\cdot\mathbf{r}_i)$. Since the wave function for an oscillator in its first excited state is just the ground-state wave function multiplied by the normal coordinate of the oscillator, it would follow that $\sum\exp(i\mathbf{q}\cdot\mathbf{r}_i)$ is the normal coordinate for rotons of momentum $\mathbf{q}$. In particular, if $|j\rangle$ is a state containing no rotons of momentum $\mathbf{q}$, then

$$V|j\rangle=\textstyle\sum\exp(i\mathbf{q}\cdot\mathbf{r}_i)|j\rangle$$
$$=[NS(q)/V]^{\frac{1}{2}}\exp(i\mathbf{q}\cdot\mathbf{r})|j\rangle. \quad (3')$$

The wave function $\exp(i\mathbf{q}\cdot\mathbf{r})|j\rangle$ represents whatever was present in the initial state j, plus a roton of momentum $\mathbf{q}$; $\mathbf{r}$ is the position coordinate of the roton. If $|j\rangle$ already contains rotons of momentum q, then $\sum\exp(i\mathbf{q}\cdot\mathbf{r}_i)$ both creates and destroys rotons. Actually, $\sum\exp(i\mathbf{q}\cdot\mathbf{r}_i)$ is not the exact normal coordinate for a roton, and consequently this factor can also create and destroy multiple excitations. In keeping with the spirit of this computation, we deal with a fictitious model in which direct production of multiple excitations by neutrons does not occur. This does not mean that no multiple excitations are produced; a neutron can produce a single virtual roton, which then breaks up into a real roton and a real phonon through the interaction term in H.

As a further simplification, we deal with a case slightly different from thermodynamic equilibrium. In the initial states j we allow an arbitrary number of phonons to be present, with the usual thermodynamic distribution; however, we consider only initial states in which no rotons are present. This picture is accurate at low temperatures. Consequently, for the initial states j the Hamiltonian is

$$H_0=\sum_{k<k_c}a_k^*a_k\hbar ck, \quad (4')$$

and for the final states f

$$H=H_0+E(\mathbf{k})+\partial\Delta/\partial\rho(\rho_0\hbar/2Vc)^{\frac{1}{2}}\sum_{k<k_c}k^{\frac{1}{2}}$$
$$\times[a_k\exp(i\mathbf{k}\cdot\mathbf{r})+a_k^*\exp(-i\mathbf{k}\cdot\mathbf{r})] \quad (5')$$
$$=H_0+H_1.$$

In determining the traces of (2′) we can use any states as a basis. For the initial states we use eigenstates of H_0, denoted by $|n_{k1}n_{k2}\cdots\rangle$; as base vectors for the states with a roton present we use products of an eigenstate of H_0 and a position eigenfunction (delta function) for the roton, denoted by $|\mathbf{x};n_{k1}n_{k2}\cdots\rangle$. Since we deal only with initial states with no rotons,

$$[\mathrm{Tr}\exp(-\beta H)]^{-1}=\prod_{k<k_c}[1-\exp(-\beta\hbar ck)].$$

For the other trace we obtain

$$\mathrm{Tr}[V\exp(-i\eta H)V^*\exp(-\beta H+i\eta H)]$$
$$=\sum_{n_{k1},n_{k2},\ldots}\int d\mathbf{x}d\mathbf{y}\langle n_{k1}n_{k2}\cdots|V|\mathbf{y};n_{k1}n_{k2}\cdots\rangle$$
$$\times\langle\mathbf{y};n_{k1}n_{k2}\cdots|\exp[-i\eta(H_0+H_1)]|\mathbf{x};n_{k1}n_{k2}\cdots\rangle$$
$$\times\langle\mathbf{x};n_{k1}n_{k2}\cdots|V^*|n_{k1}n_{k2}\cdots\rangle$$
$$\times\exp[(-\beta+i\eta)\sum_{k<k_c}n_k\hbar ck]. \quad (6')$$

[16] We are only computing the shape of $n(E)$ for energies near the roton energy. Fast neutrons tend also to produce multiple excitations, with large energy losses, but we are not studying that part of the spectrum.

From (3′) it follows that

$$\langle n_{\mathbf{k}1}n_{\mathbf{k}2}\cdots|V|\mathbf{y};n_{\mathbf{k}1}n_{\mathbf{k}2}\cdots\rangle=[NS(q)/V]^{\frac{1}{2}}\exp(i\mathbf{q}\cdot\mathbf{y}),$$
$$\langle\mathbf{x};n_{\mathbf{k}1}n_{\mathbf{k}2}\cdots|V^*|n_{\mathbf{k}1}n_{\mathbf{k}2}\cdots\rangle=[NS(q)/V]^{\frac{1}{2}}\exp(-i\mathbf{q}\cdot\mathbf{x}).$$

A great simplification is effected if we neglect the dependence of $E(\mathbf{k})$ on $\mathbf{k}$, i.e., let $E(\mathbf{k})=\Delta$. We now make this approximation and shall later consider the effects of restoring the dependence. Since H does not involve the momentum of the roton, if the wave function is initially a position eigenfunction of the roton it will remain a position eigenfunction. Consequently,

$$\langle\mathbf{y};n_{\mathbf{k}1}n_{\mathbf{k}2}\cdots|\exp[-i\eta(H_0+H_1)]|\mathbf{x};n_{\mathbf{k}1}n_{\mathbf{k}2}\cdots\rangle$$
$$=\delta(\mathbf{x}-\mathbf{y})\langle\mathbf{x};n_{\mathbf{k}1}n_{\mathbf{k}2}\cdots|\exp[-i\eta(H_0+H_1)]|\mathbf{x};n_{\mathbf{k}1}n_{\mathbf{k}2}\cdots\rangle. \quad (7')$$

The second factor on the right is simply a diagonal element of the Green's function (represented in occupation-number space) for a collection of oscillators forced by the function H_1. The matrix elements G_{mn} for a forced oscillator are easily worked out, by operator calculus[17] or other methods. If the Hamiltonian for a forced oscillator is

$$H=a^*a\epsilon+g(t)a+g^*(t)a^*, \quad (8')$$

and $|m\rangle$ and $|n\rangle$ are eigenstates of the unforced oscillator with energies $m\epsilon$ and $n\epsilon$, respectively, then

$$G_{mn}=\left\langle m\left|\exp\left[-i\int_{t'}^{t''}H(t)dt\right]\right|n\right\rangle$$
$$=\exp(i\epsilon mt'-i\epsilon nt'')(m!\,n!)^{-\frac{1}{2}}\sum_r\binom{m}{r}\binom{n}{r}r!$$
$$\times(-iB^*)^{m-r}(-iB)^{n-r}G_{00}, \quad (9')$$

where

$$B=\int_{t'}^{t''}g(t)\exp(-i\epsilon t)dt,$$

and

$$G_{00}=\exp\left(-\int\int_{t''>t>s>t'}dtdsg(t)g^*(s)e^{-i\epsilon(t-s)}\right).$$

The trace is now easily obtained. We find

$$\sum_n\exp[(-\beta+it''-it')\epsilon n]G_{nn}$$
$$=G_{00}\sum_{n\geq r\geq 0}e^{-\beta\epsilon n}\frac{n!}{r![(n-r)!]^2}(-BB^*)^{n-r}$$
$$=G_{00}(1-e^{-\beta\epsilon})^{-1}\exp[-BB^*/(e^{\beta\epsilon}-1)]. \quad (10')$$

From (5′) we have $g(t)=(\partial\Delta/\partial\rho)(\rho_0\hbar/2V_c)^{\frac{1}{2}}k^{\frac{1}{2}}e^{i\mathbf{k}\cdot\mathbf{r}}$. Replacing the sum over oscillators by $V(2\pi)^{-3}\int d\mathbf{k}$,

[17] See, for instance, R. P. Feynman, Phys. Rev. 84, 108 (1951), Eq. (38). The present case is a trivial generalization of the result given there.

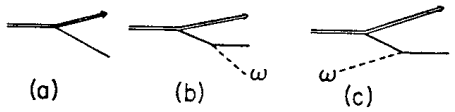


FIG. 7. Interaction of neutron (double line) with helium. In (a) the neutron produces a real roton. In (b) the neutron produces a virtual roton which decays into a real roton plus a phonon of frequency ω. In (c) the virtual roton absorbs a phonon and becomes a real roton.

we find [letting $\alpha=(\partial\Delta/\partial\rho)(\rho_0\hbar/c)^{\frac{1}{2}}$]

$$f(\eta)=(a^2\hbar k_i/m)NS(q)\exp(-i\Delta\eta)$$
$$\times\exp\left\{-\frac{\alpha^2}{(2\pi)^2(\hbar c)^4}\int_0^{\hbar ck_c}d\omega\left[-i\eta\omega^2\right.\right.$$
$$\left.\left.-\omega(e^{-i\omega\eta}-1)+\frac{4\omega\sin^2(\omega\eta/2)}{e^{\beta\omega}-1}\right]\right\}. \quad (11')$$

The various terms in the exponent of (11′) are easily understood by applying perturbation theory to (5′), treating α as small. The coefficient of $-i\eta$ is simply the energy Δ of a roton, plus a correction arising from the fact that the roton can emit and reabsorb, or absorb and re-emit, phonons. The rate of emission and reabsorption of phonons of momentum $\hbar\mathbf{k}$ is proportional to $n_{\mathbf{k}}+1$, while the rate of absorption and re-emission is proportional to $n_{\mathbf{k}}$, with an energy denominator of equal magnitude but opposite sign. Hence, the energy correction is independent of the number of phonons present, and does not depend on the temperature. The numerical value of the self-energy is $\delta E/\Delta=-0.04\,k_c^3$, with k_c measured in reciprocal angstroms. The cutoff k_c should correspond to a wavelength equal to the roton size, i.e., several interatomic spacings; we estimate $k_c\sim 0.5$ A^{-1}.

The remaining terms in the exponent represent the possibility of production of multiple excitations [Fig. 7(b)], or production of a roton which then absorbs a phonon [Fig. 7(c)]. In all diagrams the neutron (double line) interacts with the liquid only once, and produces a roton. The roton may then emit or absorb an arbitrary number of phonons, each roton-phonon interaction contributing a factor α to the amplitude and α^2 to the probability. The only processes with rates proportional to α^2 are those shown in Figs. 7(b) and 7(c). In 7(b) the neutron suffers an energy loss $\Delta+\omega$, while in 7(c) the energy loss is $\Delta-\omega$. The rate of 7(b) is proportional to $n(\omega)+1$, while 7(c) is proportional to $n(\omega)$. Hence at low temperatures the line shape is strongly asymmetric. If the exponent of (11′) is expanded into complex exponentials, the coefficient of $\exp(-i\omega\eta)$ is the rate of 7(b) as computed in perturbation theory, and the coefficient of $\exp(i\omega\eta)$ is the rate of 7(c). The remaining term, which is independent of η, represents a change in the rate of single roton production 7(a) arising from the

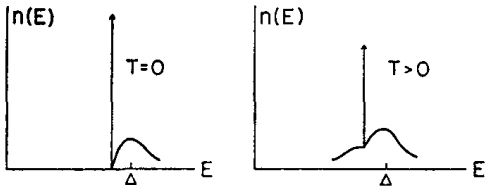


FIG. 8. Shape of a line in the neutron spectrum resulting from roton production, computed from (11′). In the actual case, the delta-functions (represented by arrows) are slightly smeared out, but the rest of the shape is as shown.

distortion of the single-roton wave function by virtual phonons.

Since α is small, the exponential in (11′) can be accurately replaced by the first two terms of a power series. Then $n(E)$ is just the coefficient of $\exp(-iE\eta)$, and we can plot the line form (Fig. 8). If only the first two terms of the power series were retained, $n(E)$ would cut off sharply at energies more than $\hbar ck_c$ from the line center; the higher terms in the power series smear out the cutoff. The one-sidedness of the curve for $T=0$ arises from the fact that the roton is the lowest excitation of momentum $\mathbf{q}$. Since no annihilation is possible at zero temperature, the neutron cannot lose any less energy than that needed to produce a roton.

At high temperatures ($\beta\hbar ck_c \ll 1$) the line becomes Gaussian with width

$$\sigma^2 = \langle (E-\bar{E})^2 \rangle_{\text{Av}} = \frac{1}{6\pi^2}\left(\frac{\partial\Delta}{\partial\rho}\right)^2 \frac{\rho_0 k_c^3}{\beta c^2}. \qquad (12')$$

It is readily shown that in thermal equilibrium the matter density fluctuates according to the Gaussian distribution. The width of the Gaussian at high temperatures is $\rho_0 k_c^3/6\pi^2\beta c^2$. Reasoning classically, one might say that although we do not know the value of the density at the place where the roton is created, nevertheless the density has an instantaneous value, which determines the amount of energy needed to create the roton. Accordingly, the line shape would be the same as the shape of the statistical distribution of the density fluctuations, as is indeed the case at high temperatures.

The classical argument fails at low temperatures, especially near the line center. For large values of η, $\sin^2(\omega\eta/2)$ may be replaced by its average value. The integral $\int d\omega\ \omega \exp(-i\omega\eta)$ approaches zero because of the oscillation of the exponential, and we find that $f(\eta) \sim \text{const} \times \exp[-i(\Delta+\delta E)\eta]$. Therefore $n(E)$ contains a delta function at the center of the line (the strength of the delta function approaches zero for large T). To understand this classically, we would have to say that there is a finite chance that the density is exactly equal to ρ_0 at the place where the roton is created; this, of course, is wrong. Since the density operator does not commute with the Hamiltonian, a density measurement would change the state of the system. Consequently, as in the double-slit experiment,

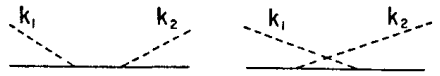


FIG. 9. Diagrams representing phonon-roton scattering.

the different possible outcomes can interfere with each other. Amplitudes, rather than probabilities must be added. Part of our uncertainty about the density comes from the fact that we do not know what state the liquid is in, because the temperature is finite; and part of the uncertainty is the quantum-mechanical uncertainty which still exists when the system is in a pure state. The former is correctly analyzable by classical reasoning, and the latter is not. Accordingly, one might say that a finite fraction of the density fluctuations is congenitally unobservable; this fraction gives rise to the delta function.

The uncertainty principle says that the width of a line arising from roton production is inversely proportional to the lifetime of the roton. Since there is nothing in the Hamiltonian (5′) which would allow the roton to disintegrate,[18] the "lifetime" is the time till the roton is scattered by a phonon. The two diagrams of Fig. 9 contribute to the scattering. If all rotons have the same energy, then energy conservation requires $k_1 = k_2$. Then the energy denominators for the two diagrams are equal in magnitude and opposite in sign, and the scattering rate is zero. This result also holds true in higher orders and is well known in meson theory. Hence the lifetime of the roton is infinite, and the line width zero.

We expect the delta function to spread out if we can analyze the Hamiltonian (5′), including the momentum dependence of $E(k)$. The width, presumably, will be the rate of scattering of rotons by phonons.[19] We thought it worthwhile to extend our analysis to include this case, since it is not entirely obvious that the roton has a finite lifetime, in the sense required by the uncertainty principle, simply because it can scatter. Furthermore, the occurrence of delta functions is not clearly precluded by the uncertainty principle; a line form consisting of a delta function super-imposed on the center of (say) a Gaussian would have a finite energy spread. Therefore we continue the analysis.

The Lagrangian form of quantum mechanics[20] is useful here. In reference 20 the problem of a particle interacting with an oscillator has been studied. Suppose the total Hamiltonian is

$$H = H_{\text{part}} + a^* a\epsilon + g(\mathbf{x},t)a + g^*(\mathbf{x},t)a^* = H_{\text{part}} + H',$$

[18] A roton cannot emit a real phonon, since the roton is the lowest state of given momentum.

[19] Another way to see that the delta function must spread out is to note that $n(\mathbf{q},E)$ is the Fourier transform in space and time of the "time-dependent pair distribution function" $p(\mathbf{r},t)$, which is the probability of finding an atom at $\mathbf{r}$ at time t, if there was an atom at the origin at $t=0$ [see L. Van Hove, Phys. Rev. **95**, 249 (1954)]. If $n(\mathbf{q},E)$ contained a delta function for some E, it would follow that $\int p(\mathbf{r},t) \exp(i\mathbf{q}\cdot\mathbf{r})d\mathbf{r}$ does not approach zero for large t. But $p(\mathbf{r},t)$ clearly becomes independent of $\mathbf{r}$ and t for large t, and the integral must approach zero.

[20] R. P. Feynman, Revs. Modern Phys. **20**, 367 (1948).

where $\mathbf{x}$ is the particle coordinate, and H_{part} is derivable from a Lagrangian L. Then the amplitude for the oscillator to go from state n at time t' to state m at time t'', while the particle goes from $\mathbf{x}$ to $\mathbf{y}$, is given by the sum over all paths $\mathbf{x}(t)$ of the functional

$$\exp\left(i\int_{t'}^{t''} L[\mathbf{x}(t)]dt\right)\left\langle m\left|\exp\left(-i\int_{t'}^{t''} H'(t)dt\right)\right|n\right\rangle,$$

where the sum is taken only over paths such that $\mathbf{x}(t')=\mathbf{x}$ and $\mathbf{x}(t'')=\mathbf{y}$. This sum is denoted by

$$\int_{\substack{\mathbf{x}(t')=\mathbf{x}\\ \mathbf{x}(t'')=\mathbf{y}}} \mathfrak{D}\mathbf{x}(t).$$

H' depends on the path $\mathbf{x}(t)$ through the forcing function g. For any particular path, however, the matrix element is given by (9′). Hence, if we save the sum over paths and the integration on $\mathbf{x}$ and $\mathbf{y}$ till the end, the oscillator sums in (6′) can be carried out as before, and we obtain[21]

$$f(\eta)=\left(\frac{a^2\hbar k_i}{m}\right)NS(q)\int d(\mathbf{y}-\mathbf{x})\exp[i\mathbf{q}\cdot(\mathbf{x}-\mathbf{y})]\times\int_{\substack{\mathbf{x}(0)=\mathbf{x}\\ \mathbf{x}(\eta)=\mathbf{y}}}\mathfrak{D}\mathbf{x}(t)\exp\left(i\int_0^{\eta}L[\mathbf{x}(t)]dt\right)\times\exp\{-\gamma A[\mathbf{x}(t)]\},\quad(13')$$

where

$$A[\mathbf{x}(t)]=\int d\mathbf{k}k\int\int_{0<s<t<\eta}dtds\left[e^{i\mathbf{k}\cdot(\mathbf{x}_t-\mathbf{x}_s)}e^{-i\omega(t-s)}\times\left(1+\frac{1}{e^{\beta\omega}-1}\right)+e^{-i\mathbf{k}\cdot(\mathbf{x}_t-\mathbf{x}_s)}\frac{e^{i\omega(t-s)}}{e^{\beta\omega}-1}\right],\quad(14')$$

$$\gamma=\left(\frac{\partial\Delta}{\partial\rho}\right)^2\left(\frac{\rho_0\hbar}{2c}\right)\left(\frac{1}{(2\pi)^3}\right),\quad\omega=\hbar ck.$$

The Hamiltonian $\Delta+(p-p_0)^2/2\mu$ comes from the Lagrangian $L=-\Delta+\frac{1}{2}\mu|d\mathbf{x}/dt|^2+|d\mathbf{x}/dt|p_0$. As $\mu\to\infty$, L becomes very large for all paths except the one with $d\mathbf{x}/dt=0$. Consequently the main contribution to (13′) comes from paths whose end points $\mathbf{x}$ and $\mathbf{y}$ are very close to each other; furthermore, the entire path must stay close to $\mathbf{x}$ and $\mathbf{y}$. If μ is actually infinite, $\mathbf{x}_t-\mathbf{x}_s$ may be set equal to zero in (14′), and then (13′) reduces to (11′), which we henceforth call $f_\infty(\eta)$. The actual value of μ is large; consequently, only for large η can the paths stray far enough from the origin to make $f(\eta)$ appreciably different from $f_\infty(\eta)$. Hence the line form is the same as that previously computed, except near the line center.

[21] We have replaced $\mathbf{q}$ by $-\mathbf{q}$. This clearly does not affect $f(\eta)$.

For any functional $M[\mathbf{x}(t)]$, we define

$$\langle M\rangle=\exp[iE(q)\eta]\int d(\mathbf{y}-\mathbf{x})\exp[i\mathbf{q}\cdot(\mathbf{x}-\mathbf{y})]\times\int_{\substack{\mathbf{x}(0)=\mathbf{x}\\ \mathbf{x}(\eta)=\mathbf{y}}}\mathfrak{D}\mathbf{x}(t)M[\mathbf{x}(t)]\exp\left(i\int_0^{\eta}L[\mathbf{x}(t)]dt\right).$$

The amplitude for a free particle to go from $\mathbf{x}$ to $\mathbf{y}$ in time η is

$$K_0(\mathbf{x},\mathbf{y},\eta)=\int_{\substack{\mathbf{x}(0)=\mathbf{x}\\ \mathbf{x}(\eta)=\mathbf{y}}}\mathfrak{D}\mathbf{x}(t)\exp\left(i\int_0^{\eta}L[\mathbf{x}(t)]dt\right)=(2\pi)^{-3}\int d\mathbf{k}\exp\{i[\mathbf{k}\cdot(\mathbf{y}-\mathbf{x})-E(k)\eta]\},\quad(15')$$

and therefore $\langle 1\rangle=1$. The operation $\langle\ \rangle$, may be regarded as a kind of average. We want to find

$$\exp[-iE(q)\eta]\langle\exp(-\gamma A)\rangle.$$

If we define $\langle\exp(-\gamma A)\rangle=\exp(-\varphi)$, then φ can be expanded as a power series in γ:

$$\varphi=\gamma\varphi_1+\gamma^2\varphi_2+\cdots,$$

where

$$\varphi_1=\langle A\rangle,\quad-\varphi_2=\tfrac{1}{2}(\langle A^2\rangle-\langle A\rangle^2),\quad\text{etc.}\quad(16')$$

If the integrations on $\mathbf{k}$, t, and s are postponed, then the kind of integral which must be done in evaluating $\langle A\rangle$ is

$$I=\int d(\mathbf{y}-\mathbf{x})e^{i\mathbf{q}\cdot(\mathbf{x}-\mathbf{y})}\int_{\substack{\mathbf{x}(0)=\mathbf{x}\\ \mathbf{x}(\eta)=\mathbf{y}}}\mathfrak{D}\mathbf{x}(t)e^{i\mathbf{k}\cdot(\mathbf{x}_t-\mathbf{x}_s)}e^{-i\omega(t-s)}\times\exp\left(i\int_0^{\eta}L[\mathbf{x}(\tau)]d\tau\right)$$

$$=e^{-i\omega(t-s)}\int\int\int d\mathbf{y}d\mathbf{x}_2d\mathbf{x}_1e^{-i\mathbf{q}\cdot\mathbf{y}}\times\int_{\substack{\mathbf{x}(t)=\mathbf{x}_2\\ \mathbf{x}(\eta)=\mathbf{y}}}\mathfrak{D}\mathbf{x}(\tau)\exp\left(i\int_t^{\eta}Ld\tau\right)e^{i\mathbf{k}\cdot\mathbf{x}_2}\times\int_{\substack{\mathbf{x}(s)=\mathbf{x}_1\\ \mathbf{x}(t)=\mathbf{x}_2}}\mathfrak{D}\mathbf{x}(\tau)\exp\left(i\int_s^{t}Ld\tau\right)e^{-i\mathbf{k}\cdot\mathbf{x}_1}\times\int_{\substack{\mathbf{x}(0)=\mathbf{x}\\ \mathbf{x}(s)=\mathbf{x}_1}}\mathfrak{D}\mathbf{x}(\tau)\exp\left(i\int_0^{s}Ld\tau\right)e^{i\mathbf{q}\cdot\mathbf{x}}$$

$$=\int\int\int d(\mathbf{y}-\mathbf{x}_2)d(\mathbf{x}_2-\mathbf{x}_1)d(\mathbf{x}_1-\mathbf{x})e^{-i\mathbf{q}\cdot(\mathbf{y}-\mathbf{x}_2)}\times K_0(\mathbf{y},\mathbf{x}_2,\eta-t)e^{-i\omega(t-s)}e^{-i(\mathbf{q}-\mathbf{k})\cdot(\mathbf{x}_2-\mathbf{x}_1)}\times K_0(\mathbf{x}_2,\mathbf{x}_1,t-s)e^{-i\mathbf{q}\cdot(\mathbf{x}_1-\mathbf{x})}K_0(\mathbf{x}_1,\mathbf{x},s)$$

$$=\exp\{-i[E(\mathbf{q})(\eta-t)+(E(\mathbf{q}-\mathbf{k})+\omega)(t-s)+E(\mathbf{q})s]\}.$$

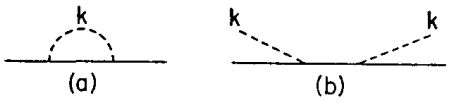


FIG. 10. Diagrams representing the two terms in $\langle A \rangle$.

This integral clearly describes the emission and re-absorption of a phonon of momentum $\hbar\mathbf{k}$ [Fig. 10(a)]. The second term in A, corresponding to absorption and re-emission [Fig. 10(b)], results in the integral

$$I'=\exp\{-i[E(\mathbf{q})(\eta-t)+(E(\mathbf{q}+\mathbf{k})-\omega)(t-s)+E(\mathbf{q})s]\}.$$

Similarly, $\langle A^2 \rangle$ gives rise to terms involving two phonons. To calculate $\langle A^n \rangle$, one must evaluate integrals of the form

$$J_\nu=\int_{0<t_1<t_2<\cdots<t_\nu<\eta}\cdots\int dt_1\cdots dt_\nu \exp(i\sum\alpha_i t_i).$$

Defining new variables $x_1=t_1$, $x_2=t_2-t_1$, $\cdots$, $x_\nu=t_\nu-t_{\nu-1}$, and introducing the function

$$f(y)=\begin{cases}1, & 0<y<\eta\\ 0, & y>\eta\end{cases}$$

$$=\frac{i}{2\pi}\int_{-\infty}^{\infty}\left(\frac{e^{-is\eta}-1}{s}\right)e^{isy}ds,$$

we find

$$J_\nu=\int_0^\infty\cdots\int_0^\infty f(\sum x_i)\exp(i\sum\beta_i x_i)dx_1\cdots dx_\nu \qquad (\beta_i=\sum_{j=i}^{\nu}\alpha_j)$$

$$=\frac{i}{2\pi}\int_{-\infty}^{\infty}ds\frac{e^{-is\eta}-1}{s}\int_0^\infty\cdots\int_0^\infty \times\exp[i\sum(\beta_i+s)x_i]dx_1\cdots dx_\nu$$

$$=\frac{1}{2\pi i^{\nu-1}}\int_{-\infty}^{\infty}ds\left[\left(\frac{e^{-is\eta}-1}{s}\right)\prod_{i=1}^{\nu}\left(\frac{1}{\beta_i+s+i\epsilon_i}\right)\right].$$

To make the next to last integral converge, we have added a small positive imaginary part $i\epsilon_i$ to each β_i. All the ϵ_i will be taken as different, so that all poles are simple, even if some of the β_i are equal. Closing the contour in the lower half-plane, we have finally

$$J_\nu=\frac{1}{i^{\nu-2}}\sum\left[\left(\frac{e^{(i\beta_i-\epsilon_i)\eta}-1}{\beta_i+i\epsilon_i}\right)\prod_{j\neq i}\times\left(\frac{1}{\beta_j-\beta_i+i(\epsilon_j-\epsilon_i)}\right)\right]. \qquad (17')$$

The integral resulting from I has the form J_2, with $\beta_1=0$, $\beta_2=E(\mathbf{q})-E(\mathbf{q}-\mathbf{k})-\omega$. In general, the β_i are the energy differences between the initial state and various intermediate states. The time integral of I' is similarly evaluated, and we find

$$-\gamma\langle A\rangle=-\gamma\int d\mathbf{k}\,k\left[\left(1+\frac{1}{e^{\beta\omega}-1}\right)\left(\frac{i\eta}{E(\mathbf{q})-E(\mathbf{q}-\mathbf{k})-\omega}-\frac{e^{i[E(\mathbf{q})-E(\mathbf{q}-\mathbf{k})-\omega]\eta}-1}{[E(\mathbf{q})-E(\mathbf{q}-\mathbf{k})-\omega]^2}\right)\right.$$
$$\left.+\frac{1}{e^{\beta\omega}-1}\left(\frac{i\eta}{E(\mathbf{q})-E(\mathbf{q}+\mathbf{k})+\omega}-\frac{e^{i[E(\mathbf{q})-E(\mathbf{q}+\mathbf{k})+\omega]\eta}-1}{[E(\mathbf{q})-E(\mathbf{q}+\mathbf{k})+\omega]^2}\right)\right]. \qquad (18')$$

In the limit of infinite roton mass, (18') is the same as the exponent in (11'). This is to be expected, since if there is only one possible path, then $\langle\exp(-\gamma A)\rangle=\exp(-\gamma\langle A\rangle)$. The various terms in (18') have the same significance as in (11'), and agree with the results of perturbation theory. None of the denominators in (18') can vanish, since the roton is defined as the lowest state of momentum $\mathbf{q}$. If we approximate $\langle\exp(-\gamma A)\rangle$ by $\exp(-\gamma\langle A\rangle)$ in our evaluation of $f(\eta)$, then the delta function in $n(E)$ still persists.

One naturally thinks of going further with the series (16') and replacing $\langle\exp(-\gamma A)\rangle$ by $\exp[-(\gamma\langle A\rangle+\gamma^2\varphi_2)]$. This procedure is not obviously valid, however, since we are interested in $f(\eta)$ for large η. Suppose, for instance, that φ_2 is linear in η for large η,

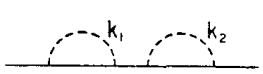


FIG. 11. A diagram which contributes a term to $\langle A^2 \rangle$ proportional to η^2. This term is canceled by the square of (10a).

but some subsequent term in the series (16') involves higher powers of η. Then neglect of the subsequent terms would make $f(\eta)$ entirely incorrect for large η. However, by inspection of the diagrams which contribute to φ_n, it is quite easy to see that φ_n is always linear in η for large η. For instance, the "bubble" in Fig. 10(a) contributes a term to $\langle A \rangle$ proportional to η because it can occur anywhere on the solid line. Among the processes contributing to $\langle A^2 \rangle$ is the one shown in Fig. 11. Since each of the bubbles can occur (almost) anywhere on the solid line, Fig. 11 contributes a term proportional to η^2; but since the two bubbles are independent if the separation between them is large, the square of (10a) cancels the term in η^2. There is a correction term in $\langle A^2 \rangle$ proportional to η because the two bubbles may overlap. Hence φ_2 is linear in η; similar arguments apply to φ_n. Another way to see that φ is asymptotically linear in η is to observe that

for large η_1 and η_2, $f(\eta)$ obeys the functional equation $f(\eta_1+\eta_2)\cong\text{const}\times f(\eta_1)f(\eta_2)$. We omit the proof.

The only part of $\gamma^2\varphi_2$ which interests us is the part proportional to η, since the rest is negligible compared with $\gamma\varphi_1$. Furthermore, pieces of $\gamma^2\varphi_2$ proportional to $i\eta$ represent small corrections to the self-energy and can be omitted. Hence we are interested only in the piece of φ_2 (if any) which is proportional to η with a real coefficient. Such a piece would cause attenuation of $f(\eta)$ for large η, and would imply that the delta function is gone.

The diagrams of Fig. 12 (and no others in this order) contribute the kind of terms we are looking for. For (12a) we have $\beta_1=0$, $\beta_2=\beta_4=E(\mathbf{q})+\omega_1-E(\mathbf{q}+\mathbf{k}_1)$, $\beta_3=E(\mathbf{q})+\omega_1-E(\mathbf{q}+\mathbf{k}_1-\mathbf{k}_2)-\omega_2$, and

$$-J_4(12a)=\frac{e^{-\epsilon_1\eta}-1}{i\epsilon_1}\left(\frac{1}{\beta_2^2(\beta_3+i\epsilon_3-i\epsilon_1)}\right)+\frac{e^{(i\beta_3-\epsilon_3)\eta}-1}{\beta_3+i\epsilon_3}\left(\frac{1}{-\beta_3+i\epsilon_1-i\epsilon_3}\right)\left(\frac{1}{(\beta_2-\beta_3)^2}\right)+\cdots,$$

where the dots indicate that uninteresting terms have been omitted. It is possible for β_3 to vanish; in fact the condition $\beta_3=0$ states that phonon-roton scattering is energetically possible. Since $1/(z+i\epsilon)=P(1/z)-i\pi\delta(z)$, the first term becomes[22]

$$\frac{i\eta}{\beta_2^2}\left\{P\left(\frac{1}{\beta_3}\right)+i\pi\delta(\beta_3)\right\}.$$

The principal value term is a correction to the self-energy and is omitted. Omitting uninteresting terms again, we find for the second term

$$\frac{\cos(\beta_3\eta)-1+i\sin(\beta_3\eta)}{\beta_3(\beta_2-\beta_3)^2}\left[-P\left(\frac{1}{\beta_3}\right)-i\pi\delta(\beta_3)\right]=-\frac{\pi\eta}{\beta_2^2}\delta(\beta_3)+\frac{\pi\eta}{\beta_2^2}\delta(\beta_3)+\cdots=0+\cdots.$$

The net contribution of (12a) is

$$J_4(12a)=\frac{\pi\eta}{\beta_2^2}\delta(\beta_3)+\cdots.$$

The contribution of (12b) has the same form, with $\beta_2=E(\mathbf{q})-E(\mathbf{q}-\mathbf{k}_1)-\omega_1$, $\beta_3=E(\mathbf{q})+\omega_2-E(\mathbf{q}-\mathbf{k}_1+\mathbf{k}_2)-\omega_1$. Similarly, (12c) contributes a term $(\pi\eta/\beta_2\beta_4)\delta(\beta_3)$, with $\beta_2=E(\mathbf{q})+\omega_1-E(\mathbf{q}+\mathbf{k}_1)$, $\beta_3=E(\mathbf{q})+\omega_1-E(\mathbf{q}+\mathbf{k}_1-\mathbf{k}_2)-\omega_2$, $\beta_4=E(\mathbf{q})-E(\mathbf{q}-\mathbf{k}_2)-\omega_2$; (12d) contributes

[22] We are free to let ϵ_1 and ϵ_2 approach zero in any order, so long as we are consistent. We let $\epsilon_3\to 0$ first.

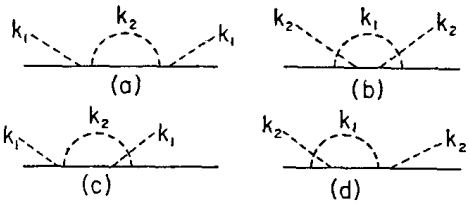


FIG. 12. Diagrams contributing attenuation to $f(\eta)$.

the same type of term, with

$$\beta_2=E(\mathbf{q})-E(\mathbf{q}-\mathbf{k}_1)-\omega_1,$$
$$\beta_3=E(\mathbf{q})+\omega_2-E(\mathbf{q}+\mathbf{k}_2-\mathbf{k}_1)-\omega_1,$$
$$\beta_4=E(\mathbf{q})+\omega_2-E(\mathbf{q}+\mathbf{k}_2).$$

For fixed $\mathbf{k}_1$ and $\mathbf{k}_2$ we add the contributions of diagrams (12a)–(12d) to the contributions of the corresponding diagrams with $\mathbf{k}_1$ and $\mathbf{k}_2$ exchanged. Noting that $\delta(\beta_3)=-\delta(-\beta_3)$, we find

$$-\gamma^2\varphi_2=\tfrac{1}{2}\gamma^2(\langle A^2\rangle-\langle A\rangle^2)$$
$$=-\eta\gamma^2\pi\int\int d\mathbf{k}_1 d\mathbf{k}_2 k_1 k_2\left(\frac{1}{e^{\beta\omega_1}-1}\right)\times\left(\frac{1}{e^{\beta\omega_2}-1}+1\right)\left[\frac{1}{E(\mathbf{q})+\omega_1-E(\mathbf{q}+\mathbf{k}_1)}+\frac{1}{E(\mathbf{q})-\omega_2-E(\mathbf{q}-\mathbf{k}_2)}\right]^2\times\delta[E(\mathbf{q}+\mathbf{k}_1-\mathbf{k}_2)+\omega_2-E(\mathbf{q})-\omega_1]$$
$$=-\eta\hbar R/2,$$

where R is just the rate of roton-phonon scattering (Fig. 8) as computed in perturbation theory. Finally we find

$$f(\eta)\simeq(a^2\hbar k_i/m)NS(q)\exp[-(i\Delta\eta+\gamma\langle A\rangle+\eta\hbar R/2)]\simeq f_\infty(\eta)\exp(-\eta\hbar R/2). \quad (19')$$

The only appreciable change in the line shape $n(E)$ is the replacement of the delta function by

$$\frac{1}{2\pi}\left(\frac{\hbar R}{(E-\Delta')^2+\hbar^2R^2/4}\right), \quad (20')$$

where Δ' is the corrected roton energy. The uncertainty principle is evidently satisfied.

As we remarked earlier, the real interaction between phonons and rotons is more complicated than the one we assumed. The line width is determined by the lifetime for roton-roton collisions, rather than roton-phonon collisions. Landau and Khalatnikov[14] fitted the viscosity data by assuming a roton-roton interaction of the form $V_0\delta(\mathbf{r}_1-\mathbf{r}_2)$ with $V_0=0.5\times10^{-38}$ erg-cm^3. The interaction term in (1′), which we call H', allows one roton to emit a phonon which is absorbed by another roton.

In computing the equivalent potential, it is permissible to regard the rotons as distinguishable. If the initial state is $\psi_i=\exp[i(\mathbf{k}_1\cdot\mathbf{r}_1+\mathbf{k}_2\cdot\mathbf{r}_2)]$ and the final state is $\psi_f=\exp\{i[(\mathbf{k}_1+\mathbf{k})\cdot\mathbf{r}_1+(\mathbf{k}_2-\mathbf{k})\cdot\mathbf{r}_2]\}$, then the amplitude to go from i to f is

$$\sum_I \frac{H'_{iI}H'_{If}}{E_I-E_i}=V_{\mathbf{k}}.$$

Roton 2 can emit a phonon $\mathbf{k}$ which is then absorbed by 1, or roton 1 can emit a phonon $-\mathbf{k}$ which is then absorbed by 2. Neglecting the dependence of roton energy on the momentum, we find $V_{\mathbf{k}}=(\partial\Delta/\partial\rho)^2(\rho_0/c^2)$. The equivalent potential is $\sum V_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{r}_{12}}=(\partial\Delta/\partial\rho)^2(\rho_0/c^2)\times\delta(\mathbf{r}_{12})$ (actually, the delta function is smeared out by the high-momentum cutoff). Using the value of Atkins and Edwards[15] for $\partial\Delta/\partial\rho$, we find $V_0=0.66\times10^{-38}$ erg-cm^3. The $\mathbf{p}\cdot\mathbf{v}$ coupling between rotons and phonons (see next paragraph) gives rise to a velocity-dependent roton-roton interaction which seems comparable in strength to the one we have computed. If the picture of a roton as a moving smoke ring is correct, then the interaction between rotons would depend strongly on their relative orientations. Accordingly, even though the delta-function interaction can be simply explained, we think the actual roton-roton interaction is more complicated.

The interaction between phonons and rotons has been discussed by Landau and Khalatnikov, and involves other terms besides $(\partial\Delta/\partial\rho)(\rho-\rho_0)$. A phonon induces a velocity field $\mathbf{v}(\mathbf{r})$ which can be represented in terms of the $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^*$. In the presence of such a field, the energy of a roton of momentum $\mathbf{p}$ is $E(p)+\mathbf{p}\cdot\mathbf{v}$. Furthermore, there are terms in $(\rho-\rho_0)^2$, such as $(\partial^2\Delta/\partial\rho^2)(\rho-\rho_0)^2$, which are responsible for most of the phonon-roton scattering. Nevertheless, a line arising from roton production will still have the shape pictured in Fig. 8, provided the delta function is replaced by a "witch" of the form (20′), where R is the rate of roton-roton scattering. As the temperature approaches zero, the rate R approaches zero because no other excitations are present to scatter the roton. Furthermore, at $T=0$, $n(E)$ is "one sided" because it is impossible to produce an excitation of momentum $\mathbf{q}$ with less energy than a roton. The curve hits the axis with finite slope because the rate of production of rotons, plus a phonon of frequency of ω, is proportional to ω for small ω. At finite temperatures, the background curve intersects the "witch" with finite slope on the right, and zero slope on the left,[23] because the rate of phonon production is proportional to $\omega[1+(e^{\beta\omega}-1)^{-1}]$ while the annihilation rate is proportional to $\omega(e^{\beta\omega}-1)^{-1}$. All these statements depend only on the fact that the coupling is a power series in the $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^*$, with each creation or destruction operator accompanied by a factor $k^{\frac{1}{2}}$.

[23] This is not exactly true. Processes involving several phonons can give rise to a small finite slope on the left.

V. Physics of Elementary Particles

Throughout his career, Feynman worked mainly in the area of physics denoted as "particles and fields," and while his interests also roamed into wider pastures, his home base remained the study of the strong, electromagnetic and weak interactions of the so-called elementary particles. Thus it was inevitable that he would become a major player in the development of high energy physics, trying to explain the important experimental discoveries in cosmic rays of the decade following World War II, and continuing with the "particle explosion" resulting from the use of the great accelerators and detectors as they came on-line in the subsequent decades.

V.A *Weak Interactions (V–A Theory)*

Since Becquerel's discovery of radioactivity in 1896, before the discovery of the electron, the proton, or the atomic nucleus, the weak interactions were observed and studied in the form of β-decay. Enrico Fermi formulated a theory of this process at the end of 1933, which described the phenomenon and, while not perfect, gave a good measure of agreement with experiment. Like QED, and the atomic and nuclear theories of the time, Fermi's four-fermion weak interaction theory obeyed the left–right symmetry known as parity. (The parity operation, switching from a left-handed to a right-handed spatial coordinate system, left the physics unchanged.) It was a shock, therefore, when a suspected violation of parity in weak interactions (such as β-decay and particle decays) was experimentally confirmed in 1957, leading in a short time interval to theories and experiments that became known as the "parity revolution." It was a real revolution, for after this no symmetry principle was taken as so sacrosanct that it was assumed to be valid without experimental confirmation.[1]

Feynman played a significant role in that revolution, publicly questioning the necessity for left–right symmetry at the 1956 Rochester Conference,[2] by informally suggesting an early version of the V–A interaction at the 1957 Rochester Conference, and by coauthoring paper [41] with Murray Gell-Mann.

Paper [41] was not without controversy as concerned its collaboration, priority, etc.[3] However, compared to its competition it uniquely emphasized certain features: the heuristic value of the two-component relativistic electron equation of second order in the time (Kramers equation),[4] current–current interaction possibly mediated by heavy intermediate bosons (see p. 197 of [41]), and the conserved vector current.[5]

[1]For a short discussion on the parity revolution, see *Pions to Quarks*, edited by L.M. Brown, M. Dresden, and L. Hoddeson (Cambridge: 1989), pp. 23–29.

[2]The Rochester Conferences, organized originally by Robert Marshak, formed an important annual series of international conferences on high energy physics. At the 1956 conference, Feynman famously repeated a question asked of him by Martin Block, about whether parity might be violated in the decay of the K-meson. For his suggestion of V–A at the 1957 conference, see note 3 of [41].

[3]For various relevant accounts of the parity revolution and V–A by some of the participants, see *Pions to Quarks* (note 2), pp. 434–496 and note 88, on p. 38. See also the sections on the V–A theory in a recent biography of Murray Gell-Mann: George Johnson, *Strange Beauty* (New York: 1999).

[4]See L.M. Brown, "Two-component fermion theory," *Phys. Rev.* **111** (1958): 957–965.

[5]V and A stand for vector and axial vector couplings, two of five interaction choices permitted by the special theory of relativity. V–A denotes a particular mixture of the two interactions, which have opposite parity, and it is their interference term in a weak interaction process that violates parity invariance.

Item [65] is a set of six pedagogical lectures on the weak interactions of hadrons (i.e. baryons and mesons) delivered by Feynman in 1964 at the International School of Physics "Ettore Majorana" in Erice, Sicily. The lectures discussed the consequences in β-decay and weak particle decay of the conserved vector current (CVC) and partially conserved vector current (PCAC). Feynman treated the decay of strange particles, based on Cabibbo's theory, and derived the Goldberger–Treiman relation between the strong meson–nucleon coupling constant and the weak coupling constant. He discussed the consequences of the Gell-Mann–Ne'eman symmetry SU(3) for the weak interactions of hadrons.

Selected Papers

[41] With M. Gell-Mann. Theory of the Fermi interaction. *Phys. Rev.* **109** (1958): 193–198.
[65] Consequences of SU(3) symmetry in weak interactions. In *Symmetries in Elementary Particle Physics* (Ettore Majorana). New York, Academic (1966), pp. 111–174.

PHYSICAL REVIEW VOLUME 109, NUMBER 1 JANUARY 1, 1958

Theory of the Fermi Interaction

R. P. FEYNMAN AND M. GELL-MANN
California Institute of Technology, Pasadena, California

(Received September 16, 1957)

The representation of Fermi particles by two-component Pauli spinors satisfying a second order differential equation and the suggestion that in β decay these spinors act without gradient couplings leads to an essentially unique weak four-fermion coupling. It is equivalent to equal amounts of vector and axial vector coupling with two-component neutrinos and conservation of leptons. (The relative sign is not determined theoretically.) It is taken to be "universal"; the lifetime of the μ agrees to within the experimental errors of 2%. The vector part of the coupling is, by analogy with electric charge, assumed to be not renormalized by virtual mesons. This requires, for example, that pions are also "charged" in the sense that there is a direct interaction in which, say, a π^0 goes to π^- and an electron goes to a neutrino. The weak decays of strange particles will result qualitatively if the universality is extended to include a coupling involving a Λ or Σ fermion. Parity is then not conserved even for those decays like $K\to 2\pi$ or 3π which involve no neutrinos. The theory is at variance with the measured angular correlation of electron and neutrino in He^6, and with the fact that fewer than 10^{-4} pion decay into electron and neutrino.

THE failure of the law of reflection symmetry for weak decays has prompted Salam, Landau, and Lee and Yang[1] to propose that the neutrino be described by a two-component wave function. As a consequence neutrinos emitted in β decay are fully polarized along their direction of motion. The simplicity of this idea makes it very appealing, and considerable experimental evidence is in its favor. There still remains the question of the determination of the coefficients of the scalar, vector, etc., couplings.

There is another way to introduce a violation of parity into weak decays which also has a certain amount of theoretical *raison d'être*. It has to do with the number of components used to describe the electron in the Dirac equation,

$$(i\nabla-\mathbf{A})\psi=m\psi. \tag{1}$$

Why must the wave function have four components? It is usually explained by pointing out that to describe the electron spin we must have two, and we must also represent the negative-energy states or positrons, requiring two more. Yet this argument is unsatisfactory. For a particle of spin zero we use a wave function of only one component. The sign of the energy is determined by how the wave function varies in space and time. The Klein-Gordon equation is second order and we need both the function and its time derivative to predict the future. So instead of two components for spin zero we use one, but it satisfies a second order equation. Initial states require specification of that one and its time derivative. Thus for the case of spin $\frac{1}{2}$ we would expect to be able to use a simple two-component spinor for the wave function, but have it satisfy a second order differential equation. For example, the wave function for a free particle would look like $U\exp[-i(Et-\mathbf{P}\cdot\mathbf{x})]$, where U has just the two components of a Pauli spinor and whether the particle refers to electron or positron depends on the sign of E in the four-vector $p_\mu=(E,\mathbf{P})$.

In fact it is easy to do this. If we substitute

$$\psi=\frac{1}{m}(i\nabla-\mathbf{A}+m)\chi \tag{2}$$

in the Dirac equation, we find that χ satisfies

$$(i\nabla-\mathbf{A})^2\chi=[(i\nabla_\mu-A_\mu)\cdot(i\nabla_\mu-A_\mu)-\tfrac{1}{2}\sigma_{\mu\nu}F_{\mu\nu}]\chi=m^2\chi, \tag{3}$$

where $F_{\mu\nu}=\partial A_\nu/\partial x_\mu-\partial A_\mu/\partial x_\nu$ and $\sigma_{\mu\nu}=\frac{1}{2}i(\gamma_\mu\gamma_\nu-\gamma_\nu\gamma_\mu)$. Now we have a second order equation, but χ still has four components and we have twice as many solutions as we want. But the operator $\gamma_5=\gamma_x\gamma_y\gamma_z\gamma_t$ commutes with $\sigma_{\mu\nu}$; therefore there are solutions of (3) for which $i\gamma_5\chi=\chi$ and solutions for $i\gamma_5\chi=-\chi$. We may select, say, the first set. We always take

$$i\gamma_5\chi=\chi. \tag{4}$$

Then we can put the solutions of (3) into one-to-one correspondence with the Dirac equation (1). For each ψ there is a unique χ; in fact we find

$$\chi=\tfrac{1}{2}(1+i\gamma_5)\psi \tag{5}$$

by multiplying (2) by $1+i\gamma_5$ and using (4). The function χ has really only two independent components. The conventional ψ requires knowledge of both χ and its time derivative [see Eq. (2)]. Further, the six $\sigma_{\mu\nu}$ in (3) can be reduced to just the three σ_{xy}, σ_{yz}, σ_{zx}. Since $\sigma_{zt}=i\gamma_z\gamma_t=i\sigma_{xy}\cdot i\gamma_5$, Eq. (4) shows that σ_{zt} may be replaced by $i\sigma_{xy}$ when operating on χ as it does in (3)

Let us use the representation

$$\gamma_t=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad \boldsymbol{\gamma}=\begin{pmatrix}0&\boldsymbol{\sigma}\\-\boldsymbol{\sigma}&0\end{pmatrix},\quad i\gamma_5=-\begin{pmatrix}0&1\\1&0\end{pmatrix},$$

where $\sigma_{x,y,z}$ are the Pauli matrices. If

$$\psi=\begin{pmatrix}a\\b\end{pmatrix},$$

[1] A. Salam, Nuovo cimento 5, 299 (1957); L. Landau, Nuclear Phys. 3, 127 (1957); T. D. Lee and C. N. Yang, Phys. Rev. 105, 1671 (1957).

where a, b are two-component spinors, we find from (5) that

$$\chi=\begin{pmatrix}\varphi\\-\varphi\end{pmatrix},$$

where $\varphi=\frac{1}{2}(a-b)$. Our Eq. (3) for the two-component spinor φ is

$$[(i\nabla_\mu-A_\mu)^2+\boldsymbol{\sigma}\cdot(\mathbf{B}+i\mathbf{E})]\varphi=m^2\varphi, \tag{6}$$

where $B_x=F_{yz}$, $E_x=F_{tx}$, etc., which is the equation we are looking for.

Rules of calculation for electrodynamics which involve only the algebra of the Pauli matrices can be worked out on the basis of (6). They, of course, give results exactly the same as those calculated with Dirac matrices. The details will perhaps be published later.

One of the authors has always had a predilection for this equation.[2] If one tries to represent relativistic quantum mechanics by the method of path integrals, the Klein-Gordon equation is easily handled, but the Dirac equation is very hard to represent directly. Instead, one is led first to (3), or (6), and from there one must work back to (1).

For this reason let us imagine that (6) had been discovered first, and (1) only deduced from it later. It would make no difference for any problem in electrodynamics, where electrons are neither created nor destroyed (except with positrons). But what would we do if we were trying to describe β decay, in which an electron is created? Would we use a field operator ψ directly in the Hamiltonian to represent the annihilation of an electron, or would we use φ? Now everything we can do one way, we can represent the other way. Thus if ψ were used it could be replaced by

$$\frac{1}{m}(p-A+m)\begin{pmatrix}\varphi\\-\varphi\end{pmatrix}, \tag{a}$$

while an expression in which φ was used could be rewritten by substituting

$$\tfrac{1}{2}(1+i\gamma_5)\psi. \tag{b}$$

If φ were really fundamental, however, we might be prejudiced against (a) on the grounds that gradients are involved. That is, an expression for β coupling which does not involve gradients from the point of view of ψ, does from the point of view of φ. So we are led to suggest *φ as the field annihilation operator to be used in β decay without gradients*. If φ is written as in (b), we see this does not conserve parity, but now we know that that is consistent with experiment.

For this reason one of us suggested the rule[3] that the electron in β decay is coupled directly through φ, or, what amounts to the same thing, in the usual four-particle coupling

$$\textstyle\sum_i C_i(\bar\psi_n O_i\psi_p)(\bar\psi_\nu O_i\psi_e), \tag{7}$$

we always replace ψ_e by $\frac{1}{2}(1+i\gamma_5)\psi_e$.

One direct consequence is that the electron emitted in β decay will always be left-hand polarized (and the positron right) with polarization approaching 100% as $v\to c$, irrespective of the kind of coupling. That is a direct consequence of the projection operator

$$a=\tfrac{1}{2}(1+i\gamma_5).$$

A priori we could equally well have made the other choice and used

$$\bar a=\tfrac{1}{2}(1-i\gamma_5);$$

electrons emitted would then be polarized to the right. We appeal to experiment[4] to determine the sign. Notice that $a^2=a$, $\bar a a=0$.

But now we go further, and suppose that the same rule applies to the wave functions of all the particles entering the interaction. We take for the β-decay interaction the form

$$\textstyle\sum C_i(\overline{a\psi}_n O_i a\psi_p)(\overline{a\psi}_\nu O_i a\psi_e),$$

and we should like to discuss the consequences of this hypothesis.

The coupling is now essentially completely determined. Since $\overline{a\psi}=\bar\psi\bar a$, we have in each term expressions like $\bar a O_i a$. Now for S, T, and P we have O_i commuting with γ_5 so that $\bar a O_i a=O_i\bar a a=0$. For A and V we have $aO_ia=O_ia^2=O_ia$ and the coupling survives. Furthermore, for axial vector $O_i=i\gamma_\mu\gamma_5$, and since $i\gamma_5 a=a$, we find $O_ia=\gamma_\mu a$; thus A leads to the same coupling as V:

$$(8)\tfrac{1}{4}G(\bar\psi_n\gamma_\mu a\psi_p)(\bar\psi_\nu\gamma_\mu a\psi_e), \tag{8}$$

the most general β-decay interaction possible with our hypothesis.[5]

This coupling is not yet completely unique, because our hypothesis could be varied in one respect. Instead of dealing with the neutron and proton, we could have made use of the antineutron and antiproton, considering them as the "true particles." Then it would be the wave function $\psi_{\bar n}$ of the antineutron that enters with the factor a. We would be led to

$$(8)\tfrac{1}{4}G(\bar\psi_{\bar p}\gamma_\mu a\psi_{\bar n})(\bar\psi_\nu\gamma_\mu a\psi_e). \tag{9}$$

This amounts to the same thing as

$$(8)\tfrac{1}{4}G(\bar\psi_n\gamma_\mu\bar a\psi_p)(\bar\psi_\nu\gamma_\mu a\psi_e), \tag{9'}$$

and from the *a priori* theoretical standpoint is just as good a choice as (8).

We have assumed that the neutron and proton are

[2] R. P. Feynman, Revs. Modern Phys. **20**, 367 (1948); Phys. Rev. **84**, 108 (1951).

[3] R. P. Feynman, *Proceedings of the Seventh Annual Rochester Conference on High-Energy Nuclear Physics, 1957* (Interscience Publishers, Inc., New York, 1957).

[4] See, for example, Boehm, Novey, Barnes, and Stech, Phys. Rev. **108**, 1497 (1957).

[5] A universal V, A interaction has also been proposed by E. C. G. Sudarshan and R. E. Marshak (to be published).

either both "particles" or both "antiparticles." We have defined the electron to be a "particle" and the neutrino must then be a particle too.

We shall further assume the interaction "universal," so for example it is

$$(8)^{\frac{1}{2}}G(\bar{\psi}_\mu\gamma_\mu a\psi_\nu)(\bar{\psi}_\nu\gamma_\mu a\psi_e) \tag{10}$$

for μ decay, as currently supposed; the μ^- is then a particle. Here the other choice, that the μ^- is an antiparticle, leads to $(8)^{\frac{1}{2}}G(\bar{\psi}_\nu\gamma_\mu a\psi_{\bar{\mu}})(\bar{\psi}_\nu\gamma_\mu a\psi_e)$, which is excluded by experiment since it leads to a spectrum falling off at high energy (Michel's $\rho=0$).

Since the neutrino function always appears in the form $a\psi_\nu$ only neutrinos with left-hand spin can exist. That is, the two-component neutrino theory with conservation of leptons is valid. Our neutrinos spin oppositely to those of Lee and Yang.[6] For example, a β particle is a lepton and spins to the left; emitted with it is an antineutrino which is an antilepton and spins to the right. In a transition with $\Delta J=0$ they tend to go parallel to cancel angular momentum. This is the angular correlation typical of vector coupling.

We have conservation of leptons and double β decay is excluded.

There is a symmetry in that the incoming particles can be exchanged without affecting the coupling. Thus if we define the symbol

$$(\bar{A}B)(\bar{C}D)=(\bar{\psi}_A\gamma_\mu a\psi_B)(\bar{\psi}_C\gamma_\mu a\psi_D),$$

we have $(\bar{A}B)(\bar{C}D)=(\bar{C}B)(\bar{A}D)$. (We have used anticommuting ψ's; for C-number ψ's the interchange gives a minus sign.[7])

The capture of muons by nucleons results from a coupling $(\bar{n}p)(\bar{\nu}\mu)$. It is already known that this capture is fitted very well if the coupling constant and coupling are the same as in β decay.[8]

If we postulate that the universality extends also to the strange particles, we may have couplings such as $(\bar{\Lambda}^0 p)(\bar{\nu}\mu)$, $(\bar{\Lambda}^0 p)(\bar{\nu}e)$, and $(\bar{\Lambda}^0 p)(\bar{p}n)$. The $(\bar{\Lambda}^0 p)$ might be replaced by $(\bar{\Sigma}^- n)$, etc. At any rate the existence of such couplings would account qualitatively for the existence of all the weak decays. Consider, for example, the decay of the K^+. It can go virtually into an anti-Λ^0 and a proton by the fairly strong coupling of strange particle production. This by the weak decay $(\bar{\Lambda}^0 p)(\bar{p}n)$ becomes a virtual antineutron and proton. These become, on annihilating, two or three pions. The parity is not conserved because of the a in front of the nucleons in the virtual transition. The theory in which only the neutrino carries the a cannot explain the parity failure for decays not involving neutrinos (the τ-θ puzzle). Here we turn the argument around; both the lack of parity conservation for the K and the fact that neutrinos are always fully polarized are consequences of the same universal weak coupling.

For β decay the expression (8) will be recognized as that for the two-component neutrino theory with couplings V and A with equal coefficients and opposite signs [expression (9) or (9′) makes the coupling $V+A$]. The coupling constant of the Fermi (V) part is equal to G. This constant has been determined[9] from the decay of O^{14} to be $(1.41\pm0.01)\times10^{-49}$ erg/cm^3. In units where $\hbar=c=1$, and M is the mass of the proton, this is

$$G=(1.01\pm0.01)\times10^{-5}/M^2. \tag{11}$$

At the present time several β-decay experiments seem to be in disagreement with one another. Limiting ourselves to those that are well established, we find that the most serious disagreement with our theory is the recoil experiment in He^6 of Rustad and Ruby[10] indicating that the T interaction is more likely than the A. Further check on this is obviously very desirable. Any experiment indicating that the electron is not 100% left polarized as $v\rightarrow c$ for any transition allowed or forbidden would mean that (8) and (9) are incorrect. An interesting experiment is the angular distribution of electrons from polarized neutrons for here there is an interference between the V and A contributions such that if the coupling is $V-A$ there is no asymmetry, while if it is $V+A$ there is a maximal asymmetry. This would permit us to choose between the alternatives (8) and (9). The present experimental results[11] agree with neither alternative.

We now look at the muon decay. The fact that the two neutrinos spin oppositely and the ρ parameter is $\frac{3}{4}$ permitted us to decide that the μ^- is a lepton if the electron is, and determines the order of $(\bar{\mu},\nu)$ which we write in (10). But now we can predict the direction of the electron in the $\pi^-\rightarrow\mu^-+\bar{\nu}\rightarrow e^-+\nu+\bar{\nu}$ sequence. Since the muon comes out with an antineutrino which spins to the right, the muon must also be spinning to the right (all senses of spin are taken looking down the direction of motion of the particle in question). When the muon disintegrates with a high-energy electron the two neutrinos are emitted in the opposite direction. They have spins opposed. The electron emitted must spin to the left, but must carry off the angular momentum of the muon, so it must proceed in the direction opposite to that of the muon. This direction agrees with experiment. The proposal of Lee and Yang predicted

[6] This is only because they used S and T couplings in β decay; had they used V and A, their theory would be similar to ours, with left-handed neutrinos.

[7] We can express $(\bar{A}B)(\bar{C}D)$ directly in terms of the two-component spinors φ: $(\bar{A}B)(\bar{C}D)=4(\varphi_A^*\varphi_B)(\varphi_C^*\varphi_D)-4(\varphi_A^*\sigma\varphi_B)\cdot(\varphi_C^*\sigma\varphi_D)$. If we put $\varphi_A=\begin{pmatrix}A_1\\A_2\end{pmatrix}$, etc., where A_1 and A_2 are complex numbers, we obtain $8(A_1^*C_2^*-A_2^*C_1^*)(B_1D_2-B_2D_1)$ and the symmetry is evident.

[8] See, for example, J. L. Lopes, Phys. Rev. (to be published); L. Michel, *Progress in Cosmic-Ray Physics*, edited by J. G. Wilson (Interscience Publishers, Inc., New York, 1952), Vol. 1, p. 125.

[9] Bromley, Almquist, Gove, Litherland, Paul, and Ferguson, Phys. Rev. **105**, 957 (1957).

[10] B. M. Rustad and S. L. Ruby, Phys. Rev. **97**, 991 (1955).

[11] Burgy, Epstein, Krohn, Novey, Raboy, Ringo, and Telegdi, Phys. Rev. **107**, 1731 (1957).

the electron spin here to be opposite to that in the case of β decay. Our β-decay coupling is V, A instead of S, T and this reverses the sign. That the electron have the same spin polarization in all decays (β, muon, or strange particles) is a consequence of putting $a\psi_e$ in the coupling for this particle. It would be interesting to test this for the muon decay.

Finally we can calculate the lifetime of the muon, which comes out

$$\tau = 192\pi^3/G^2\mu^5 = (2.26\pm 0.04)\times 10^{-6} \text{ sec}$$

using the value (11) of G. This agrees with the experimental lifetime[12] $(2.22\pm 0.02)\times 10^{-6}$ sec.

It might be asked why this agreement should be so good. Because nucleons can emit virtual pions there might be expected to be a renormalization of the effective coupling constant. On the other hand, if there is some truth in the idea of an interaction with a universal constant strength it may be that *the other interactions are so arranged so as not to destroy this constant.* We have an example in electrodynamics. Here the coupling constant e to the electromagnetic field is the same for all particles coupled. Yet the virtual mesons do not disturb the value of this coupling constant. Of course the distribution of charge is altered, so the coupling for high-energy fields is apparently reduced (as evidenced by the scattering of fast electrons by protons), but the coupling in the low-energy limit, which we call the total charge, is not changed.

Using this analogy to electrodynamics, we can see immediately how the Fermi part, at least, can be made to have no renormalization. For the sake of this discussion imagine that the interaction is due to some intermediate (electrically charged) vector meson of very high mass M_0. If this meson is coupled to the "current" $(\bar{\psi}_p\gamma_\mu a\psi_n)$ and $(\bar{\psi}_\mu\gamma_\mu a\psi_\nu)$ by a coupling $(4\pi f^2)^{\frac{1}{2}}$, then the interaction of the two "currents" would result from the exchange of this "meson" if $4\pi f^2 M_0^{-2} = (8)^{\frac{1}{2}}G$. Now we must arrange that the total current

$$J_\mu = (\bar{\psi}_p\gamma_\mu a\psi_n) + (\bar{\psi}_\nu\gamma_\mu a\psi_e) + (\bar{\psi}_\nu\gamma_\mu a\psi_\mu) + \cdots \qquad (12)$$

be not renormalized. There are no known large interaction terms to renormalize the $(\bar{\nu}e)$ or $(\bar{\nu}\mu)$, so let us concentrate on the nucleon term. This current can be split into two: $J_\mu = \frac{1}{2}(J_\mu{}^V + J_\mu{}^A)$, where $J_\mu{}^V = \bar{\psi}_p\gamma_\mu\psi_n$ and $J_\mu{}^A = \bar{\psi}_p i\gamma_\mu\gamma_5\psi_n$. The term $J_\mu{}^V = \bar{\psi}\gamma_\mu\tau_+\psi$, in isotopic spin notation, is just like the electric current. The electric current is

$$J_\mu{}^{\text{el}} = \bar{\psi}\gamma_\mu(\tfrac{1}{2}+\tau_z)\psi.$$

The term $\frac{1}{2}\bar{\psi}\gamma_\mu\psi$ is conserved, but the term $\bar{\psi}\gamma_\mu\tau_z\psi$ is not, unless we add the current of pions, $i[\varphi^*T_z\nabla_\mu\varphi - (\nabla_\mu\varphi^*)T_z\varphi]$, because the pions are charged. Likewise $\bar{\psi}\gamma_\mu\tau_+\psi$ is not conserved but the sum

$$J_\mu{}^V = \bar{\psi}\gamma_\mu\tau_+\psi + i[\varphi^*T_+\nabla_\mu\varphi - (\nabla_\mu\varphi)^*T_+\varphi] \qquad (13)$$

is conserved, and, like electricity, leads to a quantity whose value (for low-energy transitions) is unchanged by the interaction of pions and nucleons. If we include interactions with hyperons and K particles, further terms must be added to obtain the conserved quantity.

We therefore suppose that this conserved quantity be substituted for the vector part of the first term in (12). Then the Fermi coupling constant will be strictly universal, except for small electromagnetic corrections. That is, the constant G from the μ decay, which is accurately $V-A$, should be also the exact coupling constant for at least the vector part of the β decay. (Since the energies involved are so low, the spread in space of $J_\mu{}^V$ due to the meson couplings is not important, only the total "charge.") It is just this part which is determined by the experiment with O^{14}, and that is why the agreement should be so close.

The existence of the extra term in (13) means that other weak processes must be predicted. In this case there is, for example, a coupling

$$(8)^{\frac{1}{2}}Gi(\varphi^*\nabla_\mu T_+\varphi - (\nabla_\mu\varphi)^*T_+\varphi)(\bar{\psi}_\nu\gamma_\mu a\psi_e),$$

by which a π^- can go to a π^0 with emission of $\bar{\nu}$ and e. The amplitude is

$$4G(p_\mu{}^- + p_\mu{}^0)(\bar{\psi}_\nu\gamma_\mu a\psi_e),$$

where p^-, p^0 are the four-momenta of π^- and π^0. Because of the low energies involved, the probability of the disintegration is too low to be observable. To be sure, the process $\pi^-\to\pi^0+e+\bar{\nu}$ could be understood to be qualitatively necessary just from the existence of β decay. For the π^- may become virtually an antiproton and neutron, the neutron decay virtually to a proton, e, and $\bar{\nu}$ by β decay and the protons annihilate forming the π^0. But the point is that by our principle of a universal coupling whose vector part requires no renormalization we can calculate the rate directly without being involved in closed loops, strong couplings, and divergent intervals.

For any transition in which strangeness doesn't change, the current $J_\mu{}^V$ is the total current density of isotopic spin T_+. Thus the vector part gives transitions $\Delta T = 0$ with square matrix element $T(T+1) - T_zT_z'$ if we can neglect the energy release relative to the rest mass of the particle decaying. For the nucleon and $K^-\to K^0+e+\bar{\nu}$ the square of the matrix element is 1, for the pion and $\Sigma^-\to\Sigma^0+e+\bar{\nu}$ it is 2. The axial coupling in the low-energy limit is zero between states of zero angular momentum like the π meson or O^{14}, so for both of these we can compute the lifetime knowing only the vector part. Thus the $\pi^-\to\pi^0+e+\bar{\nu}$ decay should have the same ft value as O^{14}. Unfortunately because of the very small energies involved (because isotopic spin is such a good quantum number) none of these decays of mesons or hyperons are fast enough to observe in competition to other decay processes in which T or strangeness changes.

[12] W. E. Bell and E. P. Hincks, Phys. Rev. 84, 1243 (1951).

This principle, that the vector part is not renormalized, may be useful in deducing some relations among the decays of the strange particles.

Now with present knowledge it is not so easy to say whether or not a pseudovector current like $\bar{\psi} i\gamma_5\gamma_\mu\tau_+\psi$ can be arranged to be not renormalized. The present experiments[13] in β decay indicate that the ratio of the coupling constant squared for Gamow-Teller and Fermi is about 1.3 ± 0.1. This departure from 1 might be a renormalization effect.[14] On the other hand, an interesting theoretical possibility is that it is exactly unity and that the various interactions in nature are so arranged that it need not be renormalized (just as for V). It might be profitable to try to work out a way of doing this. Experimentally it is not excluded. One would have to say that the $ft_{\frac{1}{2}}$ value of 1220 ± 150 measured[15] for the neutron was really 1520, and that some uncertain matrix elements in the β decay of the mirror nuclei were incorrectly estimated.

The decay of the π^- into a μ^- and $\bar{\nu}$ might be understood as a result of a virtual process in which the π becomes a nucleon loop which decays into the $\mu+\bar{\nu}$. In any event one would expect a decay into $e+\bar{\nu}$ also. The ratio of the rates of the two processes can be calculated without knowledge of the character of the closed loops. It is $(m_e/m_\mu)^2(1-m_\mu^2/m_\pi^2)^{-2}=13.6\times10^{-5}$. Experimentally[16] no $\pi\to e+\nu$ have been found, indicating that the ratio is less than 10^{-5}. This is a very serious discrepancy. The authors have no idea on how it can be resolved.

We have adopted the point of view that the weak interactions all arise from the interaction of a current J_μ with itself, possibly via an intermediate charged vector meson of high mass. This has the consequence that any term in the current must interact with all the rest of the terms and with itself. To account for β decay and μ decay we have to introduce the terms in (12) into the current; the phenomenon of μ capture must then also occur. In addition, however, the pairs $e\nu$, $\mu\nu$, and pn must interact with themselves. In the case of the $(\bar{e}\nu)(\bar{\nu}e)$ coupling, experimental detection of electron-neutrino scattering might some day be possible if electron recoils are looked for in materials exposed to pile neutrinos; the cross section[17] with our universal coupling is of the order of 10^{-45} cm^2.

To account for all observed strange particle decays it is sufficient to add to the current a term like $(\bar{p}\Lambda^0)$, $(\bar{p}\Sigma^0)$, or $(\overline{\Sigma^-}n)$, in which strangeness is increased by one as charge is increased by one. For instance, $(\bar{p}\Lambda^0)$ gives us the couplings $(\bar{p}\Lambda^0)(\bar{e}\nu)$, $(\bar{p}\Lambda^0)(\bar{\mu}\nu)$, and $(\bar{p}\Lambda^0)(\bar{n}p)$. A direct consequence of the coupling $(\bar{p}\Lambda^0)(\bar{e}\nu)$ would be the reaction

$$\Lambda^0\to p+e+\bar{\nu} \tag{14}$$

at a rate 5.3×10^7 sec^{-1}, assuming no renormalization of the constants.[18] Since the observed lifetime of the Λ^0 (for disintegration into other products, like $p+\pi^-$, $n+\pi^0$) is about 3×10^{-10} sec, we should observe process (14) in about 1.6% of the disintegrations. This is not excluded by experiments. If a term like $(\overline{\Sigma^-}n)$ appears, the decay $\Sigma^-\to n+e^-+\nu$ is possible at a predicted rate 3.5×10^8 sec^{-1} and should occur (for $\tau_{\Sigma^-}=1.6\times10^{-10}$ sec) in about 5.6% of the disintegrations of the Σ^-. Decays with μ replacing the electron are still less frequent. That such disintegrations actually occur at the above rates is not excluded by present experiments. It would be very interesting to look for them and to measure their rates.

These rates were calculated from the formula Rate$=(2G^2W^5c/30\pi^3)$ derived with neglect of the electron mass. Here $W=(M_\Lambda^2-M_p^2)/2M_\Lambda$ is the maximum electron energy possible and c is a correction factor for recoil. If $x=W/M_\Lambda$ it is

$$c=-\tfrac{15}{16}x^{-5}(1-2x)^2\ln(1-2x)-\tfrac{5}{8}x^{-4}(1-x)(3-6x-2x^2),$$

and equals 1 for small x, about 1.25 for the Σ decay, and 2.5 for $M_p=0$.

It should be noted that decays like $\Sigma^+\to n+e^++\nu$ are forbidden if we add to the current only terms for which $\Delta S=+1$ when $\Delta Q=+1$. In order to cause such a decay, the current would have to contain a term with $\Delta S=-1$ when $\Delta Q=+1$, for example $(\overline{\Sigma^+}n)$. Such a term would then be coupled not only to $(\bar{\nu}e)$, but also to all the others, including one like $(\bar{p}\Lambda^0)$. But a coupling of the form $(\overline{\Sigma^+}n)(\overline{\Lambda^0}p)$ leads to strange particle decays with $\Delta S=\pm2$, violating the proposed rule $\Delta S=\pm1$. It is important to know whether this rule really holds; there is evidence for it in the apparent absence of the decay $\Xi^-\to\pi^-+n$, but so few Ξ particles have been seen that this is not really conclusive. We are not sure, therefore, whether terms like $(\overline{\Sigma^+}n)$ are excluded from the current.

We deliberately ignore the possibility of a neutral current, containing terms like $(\bar{e}e)$, $(\bar{\mu}e)$, $(\bar{n}n)$, etc., and possibly coupled to a neutral intermediate field. No weak coupling is known that requires the existence of such an interaction. Moreover, some of these couplings, like $(\bar{e}e)(\bar{\mu}e)$, leading to the decay of a muon into three electrons, are excluded by experiment.

It is amusing that this interaction satisfies simultaneously almost all the principles that have been

[13] A. Winther and O. Kofoed-Hansen, Kgl. Danske Vidensakb. Selskab, Mat.-fys. Medd. (to be published).

[14] This slight inequality of Fermi and Gamow-Teller coupling constants is not enough to account for the experimental results of reference 11 on the electron asymmetry in polarized neutron decay.

[15] Spivac, Sosnovsky, Prokofiev, and Sokolov, *Proceedings of the International Conference on the Peaceful Uses of Atomic Energy, Geneva, 1955* (United Nations, New York, 1956), A/Conf. 8/p/650.

[16] C. Lattes and H. L. Anderson, Nuovo cimento (to be published).

[17] For neutrinos of energy ω (in units of the electron mass m) the total cross section is $\sigma_0\omega^2/(1+2\omega)$, and the spectrum of recoil energies ϵ of the electron is uniform $d\epsilon$. For antineutrinos it is $\sigma_0(\omega/6)[1-(1+2\omega)^{-3}]$ with a recoil spectrum varying as $(1+\omega-\epsilon)^2$. Here $\sigma_0=2G^2m^2/\pi=8.3\times10^{-45}$ cm^2.

[18] R. E. Behrends and C. Fronsdal, Phys. Rev. 106, 345 (1957).

proposed on simple theoretical grounds to limit the possible β couplings. It is universal, it is symmetric, it produces two-component neutrinos, it conserves leptons, it preserves invariance under CP and T, and it is the simplest possibility from a certain point of view (that of two-component wave functions emphasized in this paper).

These theoretical arguments seem to the authors to be strong enough to suggest that the disagreement with the He^6 recoil experiment and with some other less accurate experiments indicates that these experiments are wrong. The $\pi \rightarrow e + \bar{\nu}$ problem may have a more subtle solution.

After all, the theory also has a number of successes. It yields the rate of μ decay to 2% and the asymmetry in direction in the $\pi \rightarrow \mu \rightarrow e$ chain. For β decay, it agrees with the recoil experiments[19] in A^{35} indicating a vector coupling, the absence of Fierz terms distorting the allowed spectra, and the more recent electron spin polarization[4] measurements in β decay.

[19] Herrmansfeldt, Maxson, Stähelin, and Allen, Phys. Rev. **107**, 641 (1957).

Besides the various experiments which this theory suggests be done or rechecked, there are a number of directions indicated for theoretical study. First it is suggested that all the various theories, such as meson theory, be recast in the form with the two-component wave functions to see if new possibilities of coupling, etc., are suggested. Second, it may be fruitful to analyze further the idea that the vector part of the weak coupling is not renormalized; to see if a set of couplings could be arranged so that the axial part is also not renormalized; and to study the meaning of the transformation groups which are involved. Finally, attempts to understand the strange particle decays should be made assuming that they are related to this universal interaction of definite form.

ACKNOWLEDGMENTS

The authors have profited by conversations with F. Boehm, A. H. Wapstra, and B. Stech. One of us (M. G. M.) would like to thank R. E. Marshak and E. C. G. Sudarshan for valuable discussions.

CONSEQUENCES OF SU_3 SYMMETRY IN WEAK INTERACTIONS

R.P. Feynman,
California Institute of Technology.

Introduction

1st LECTURE

These lectures will cover the relationship of SU_3 and the weak interactions. The lectures will be geared for experimental people so that they may get an idea of how our theoretical predictions arise. At first, I will speak a little bit about how calculations are made for the weak decays, then we shall consider SU_3, and finally the effects of unitary symmetry upon the weak interactions. The subject matter is split in this way so that one may get a clearer idea of the origins of the various problems in the theory of weak interactions. Not all of our difficulties arise from SU_3 nor do all of the successes, and it is important to realize this.

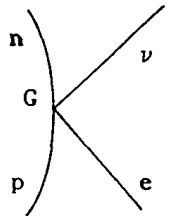
The theory of weak decays is very unsatisfactory except that it agrees with experiment. To understand that remark let us consider the muon. A muon is a particle which has exactly the same properties as the electron except that its mass is 207 times the mass of the electron. This statement completely describes our experiments with the muon, but such a comment is also unsatisfactory for a true theorist. Experimentalists find a beautiful and simple thing which is easy for the theorists to describe. Nevertheless, we must be unhappy about this situation because we have no idea of why this particle exists. Similarly, the theory of weak decay, up to the point where we encounter strangeness changing interactions, is accurate but unsatisfactory. There are various mysterious properties which I shall mention as I go on, but I should like to remind you that the most mysterious aspect of the weak decays is that they exist at all. It seems so much simpler just to forget them. There is no clue from electromagnetism, from gravity, or from nuclear forces that the weak interactions must exist. They seem to have no connection with the rest of the world.

Form of the four-fermion weak interaction

The theory of the weak decays was originated by Fermi. Fermi tried to describe mathematically an idea of Pauli's concerning the neutrino. For example, in K capture an orbital electron in the K shell is eaten by a proton in the nucleus with the result that a neutron plus a neutrino is created. Fermi attempted to write down an abstract description of this process. He assumed that there was a Hamiltonian which contained a term of the form

$$Q^*_\nu \, Q^*_n \, Q_e \, Q_p \, .$$

In this form the Q^*_n is an operator which creates a neutron, Q_p is an operator which destroys a proton, Q^*_ν creates a neutrino, and Q_e destroys an electron. Such a term is present in a Hamiltonian of the world with a certain strength characterized by a coupling constant G, and gives rise to a certain amplitude for the K-capture reaction, but no mechanism is described by it. Another way of describing the reaction is to draw a diagram with a four-point coupling



and to associate with that diagram the strength constant G, together with certain well-defined rules for calculating the amplitude. The decay amplitude is proportional to G and the amplitude for finding the proton and the electron together. Given the amplitude for the process, the rate at which the reaction proceeds is given by a Golden Rule of the form

$$\text{Rate} = 2\pi |\text{Amp}|^2 \ \text{(Density of final states)}.$$

Actually things are a bit more complicated. A fermion, described by the Dirac equation, is represented by four amplitudes, two spin possibilities times two charge possibilities. The product of four such Q's will produce 4^4 possible terms, each of which might have its own characteristic coupling

constant. Actually such a horrible mess is simplified by requiring relativistic invariance of the amplitude. This reduces the number of possible coefficients to ten. Mathematically, if we represent the four-component spinor wave function of a particle by its name, the possible forms for the coupling can be represented as

$$(\bar{\nu}1e)\ (\bar{n}1p),\ (\bar{\nu}\gamma_\alpha e)\ (\bar{n}\gamma_\alpha p),\ (\bar{\nu}1e)\ (\bar{n}\gamma_5 p)\ ,$$

and so on. Fermi, however, was not satisfied with writing down all the possibilities. He made a guess that the interaction was vector

$$(\bar{\nu}\gamma_\alpha e)\ (\bar{n}\gamma_\alpha p)\ .$$

Having limited the form of the amplitude by this guess, he was then able to calculate the properties of the beta decay.

Parity non-conservation

The most prominent feature is the shape of the spectrum, which depends almost wholly on the density of final states. Early experiments failed to confirm this spectrum, but those experiments were incorrect because of rescattering in the foils. The subsequent history of beta-decay investigations is beclouded by two decades of inaccurate observations and poor theoretical suggestions, which I shall not discuss. The final resolution of the puzzle is simply to multiply each wave function of a particle by the left-handed helicity operator,

$$a = (1 + i\gamma_5)/2\ .$$

This prescription reduces to one the number of independent coupling constants in beta decay.

The helicity operator has certain properties, namely

$$aa = a,\ \bar{a}\bar{a} = \bar{a},\ a\bar{a} = 0\ .$$

Using this helicity operator in the interaction also results in a violation of parity conservation in the weak interaction, which was found in 1957. The final form for the beta-decay coupling in the Lagrangian may then be written as

$$\sqrt{8}\ G(\bar{a}\bar{\nu}\gamma_\alpha ae)\ (\bar{a}\bar{n}\gamma_\alpha ap)\ .$$

With such a coupling, the particles are emitted polarized along the direction of their motion. They are spinning to the left with a probability of $(1+v)/2$, and to the right with a probability of $(1-v)/2$, where v is the particle velocity in units of c. The opposite polarization holds for the antiparticles. For beta decay the above Lagrangian may be written in the form

$$\frac{1}{\sqrt{2}}\ G\left[\bar{\nu}\left(\gamma_\alpha + i\gamma_\alpha\gamma_5\right) e\right] \cdot \left[\bar{n}\left(\gamma_\alpha + i\gamma_\alpha\gamma_5\right)p\right]$$

and it is common today to identify the various terms in this formula with current-density operators. People like to define the vector current of the leptons to be $J_{V\alpha}^{\bar{\nu}e} = (\bar{\nu}\gamma_\alpha e)$, and the axial vector current of the leptons as $J_{A\alpha}^{\bar{\nu}e} = (\bar{\nu}\gamma_\alpha i\gamma_5 e)$, together with similar definitions of the nuclear vector and axial vector currents. With this symbolism the beta-decay interaction takes the form

$$\frac{G}{\sqrt{2}}\left[J_{V\alpha}^{\bar{\nu}e} + J_{A\alpha}^{\bar{\nu}e}\right]\left[J_{V\alpha}^{\bar{n}p} + J_{A\alpha}^{\bar{n}p}\right]\ .$$

To discuss the decay of the muon we need only replace the proton spinor by a neutrino spinor and the neutron spinor by a muon spinor in the above coupling. I shall not go through the calculation of the spectrum and polarization properties of the decay on the basis of this proposed theory, but shall only comment on the results. This spectrum agrees very well with the experiments which I have seen. In fact the agreement between experiment and theory today is so detailed that one has to take into account the radiative corrections to the spectrum. The absolute value of the coupling constant G_μ, can be determined from the rate of the decay of the muon, and the result is $G_\mu M_p^2 = 1.01 \cdot 10^{-5}$. (I use natural units in which $\hbar = c = 1$.)

Strong interaction modification of weak decay matrix elements, and the conserved vector current theory

Let us return to the beta decay of the neutron. In this case we have two kinds of currents, vector and axial vector, which couple to the leptons. In beta decay the nucleons are non-relativistic and we can simplify

the calculation of matrix elements. The four-dimensional vector current for a non-relativistic particle is dominated by its time component, which is just the operator 1, while the axial vector current is dominated by its space-like terms, which are just the spin operators. The vector term does not change the spin of the nucleon and is called the Fermi coupling, while the axial vector term, called Gamow-Teller coupling, flips the spin of the nucleon. Because each type of coupling leads to distinctive selection rules it was known quite early that both types of coupling are present. In the muon decay the ratio of the axial-vector term to the vector term is unity, but from experiments on the disintegration of the neutron and O^{14} it was shown that the ratio in nuclear-beta decay differs substantially from unity. If we calculate the properties of neutron decay using the coupling

$$\frac{1}{\sqrt{2}}\left[\bar{\nu}\left(\gamma_\alpha + i\gamma_\alpha\gamma_5\right) e\right] \cdot \left[\bar{n}\left(G_V\gamma_\alpha - iG_A\gamma_\alpha\gamma_5\right) p\right] ,$$

we come to the conclusion that for polarized neutrons if $G_A = - G_V$, the electron is emitted isotropically, but the neutrino is much more likely to come out along the direction of the neutron spin than opposite. The recoil proton is also anisotropic, but the electron is not. Actually the electron is slightly unsymmetric in its emission direction. From measurements on the decay of polarized neutrons one can then find that the ratio of the coupling constants is negative and approximately 1.2. What a destruction of the beautifully simple theory! $(1 + i\gamma_5)/2$ is very pretty because its square equals itself, and it projects out a certain helicity component, but $(1 + 1.2\ i\gamma_5)/2$ is dirty. However, this situation is really not unexpected, because the nucleons are fairly complicated particles due to strong interactions about which we know little. Indeed, in the sense of perturbation theory, it would be expected that a fundamental simple-type of coupling would no longer be simple. In other words, it is still possible that in the ddep heart of matter the coupling formula for the nucleons involves the helicity projection operator, but that the spinors which entered do not represent the real proton and neutron but rather some kind of

idealized p and n. In calculating matrix elements for the real nucleons, corrections would then have to be made for pions and other strongly interacting particles which would renormalize the relative coefficient of the vector and axial vector current.

Strong interactions in fact would be expected to renormalize not only the axial current but also the vector current. After some mental effort one can see, however, that it is quite possible that the vector current need not be renormalized at all, as experimental clues hinted, when it was found that the muon coupling constant was the same as the Fermi coupling constant to within a few per cent. The extraction of the Fermi constant from the rate of O^{14} beta decay involves an additional assumption about nuclear forces unless the vector current for the weak interactions is not renormalized. In the decay the parent nucleus contains eight protons and six neutrons, while the daughter nucleus contains seven protons and seven neutrons. Now, there is a state in N^{14} which is the same as the ground state in O^{14} in the sense of isotopic spin. The nuclear forces are independent of whether the nucleons are protons or neutrons, and therefore in every system of seven protons and seven neutrons there is a state which has essentially the same character as any state that exists for eight protons and six neutrons. (The inverse is not true because of the exclusion principle.) The sister state of the ground state of N^{14}, which has isotopic spin 1, lies lower in energy than that state and so the nitrogen nucleus does decay to its sister state. Since the kinematic features of the wave functions of the initial and final states are the same, the matrix element of the integral of the beta-decay charge density involves simply an isotopic spin Clebsch-Gordan coefficient, which is just $\sqrt{2}$. The calculation is not exact because the violation of isotopic spin invariance by electromagnetism destroys the perfect overlap of the wave functions, but this effect is at most a half per cent.

This matrix element is almost the sole example of an accurate calculation in nuclear physics, because one does not have to know any nuclear physics to calculate it. This trick is possible only at very low momentum transfer where the integral of the fourth component of the charge density becomes the

operator for the total charge. Since the isotopic spin is conserved by the strong interactions we have therefore no renormalization of such matrix elements.

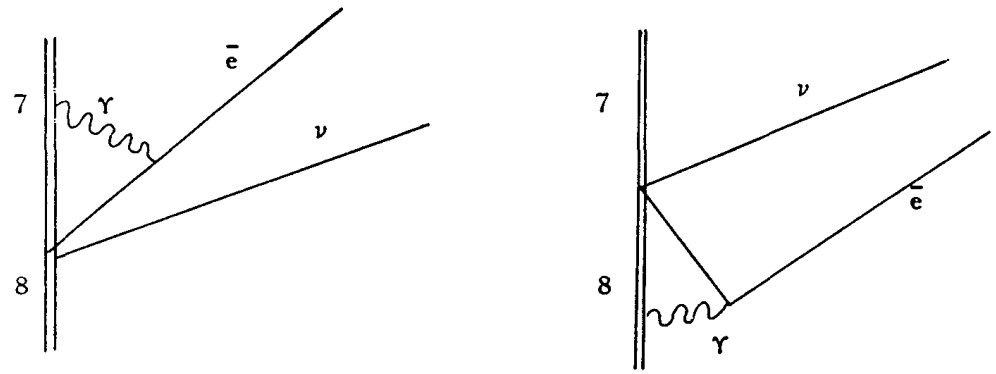
The assumption that the vector part of the beta-decay current is a component of the isotopic-spin current, implies also that two pions must have a well-defined weak coupling to two leptons. It is quite easy to see how strong the coupling is by using our electricity analogue. Since the isotopic spin of the pion is the same as that for O^{14}, namely one, we again get a factor of $\sqrt{2}$ for the matrix element. Similarly, there will be matrix elements of the currents between two kaons, two Σ hyperons, and so on. We propose that this, in fact, is the way the world works and, consequently, that the vector part of the beta-decay coupling which conserves hypercharge is not renormalized.

Electromagnetic corrections to weak interaction matrix elements

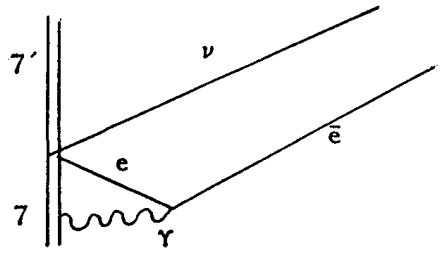
When the coupling constants for the Fermi beta decay and that for the muon beta decay are determined in the manner indicated, they turn out not to be equal as we had expected, but to differ by a few per cent. To be more accurate, the vector coupling constant determined from the rate of decay of O^{14} is 0.985 times the muon beta-decay constant, whereas, that determined from the decay of Al^{26} is 0.975 G_μ. In doing calculations we have a problem with certain relativistic electromagnetic corrections. We are trying to investigate a difference of a few per cent and the order of magnitude of electromagnetic effects is the same. Because the agreement between the two coupling constants is so close, one is reluctant to guess that the idea of equality is wrong ... . One would prefer to speculate that the discrepancy may be due to something else, such as a misunderstanding in the calculation of the electromagnetic corrections. It is therefore worth while to discuss the question of how accurately we know such corrections. In computing them for the decay of the muon we have no problem; the electromagnetic structure of the muon is well known and it has no other anomalous moments. The calculation goes through straightforwardly. However, in calculating proton corrections to the disintegration of the neutron we get into difficulties because the integrals diverge. The divergence is due to the anomalous magnetic moment of the nucleons. Now, when we have to compute such integrals, frankly we do not get completely satisfactory answers; depending on where one puts the cut-off one might at first be able to account for any

discrepancy. However, in spite of this apparent uncertainty, it turns out that the answer is not very sensitive to the cut-off. In fact, the analysis is not uncertain by more than about ± ½ per cent. Even though the electromagnetic effects are larger than this, the biggest part of them can be understood without serious ambiguity.

The uncertainty in the radiative correction is expected to be quite small because it is really a correction to a correction. Let me explain the origin of the major part of the correction by referring to the oxygen decay. In this case we would have the following types of diagrams



In these diagrams we see that the virtual photon interacts both with an object of charge 7 and one of charge 8. The usual procedure would be to use a Coulomb wave function for the outgoing positrons (that is, for a field of a nucleus with charge 7). This results in a well-known correction to the f value. But the second diagram is not included correctly in this procedure. The approximation actually includes the following diagram



in which, outside the nucleus, an electron-positron pair is created by the virtual photon which interacts with a charge of seven units. These two diagrams A and B have to be added, and are comparable in strength for a relativistic positron. In fact, f accounts for the sum of A and C. The error is the difference of diagram B and C, which is a diagram like C but with charge 1 on the nucleus. When this is calculated, disregarding recoil of the nucleus, one obtains $(e^2/2\pi)$ ln (X/E) where E is the positron energy and X is ∞. Actually, the result should be modified for recoil and magnetic-moment corrections etc., the result being to replace X by some unknown of the order M_p. If I stop the integral at about $2M_p$, I will then have to add an unknown amount to account for the high energy piece of the integral. It is this unknown amount that is estimated to be of the order of ½ per cent. From my experience with such calculations I do not believe that this correction could be significantly larger than ½ per cent. The point is that the major part of the radiative correction is well known (the log is large because the lower limit E is so much smaller than M_p) and that the remaining part is very unlikely to account for the discrepancy between the nuclear and muon weak decay constants, even though the calculation involves some uncertainty.

Can the radiative correction calculation ever be made more precisely? I believe that it can and would like to propose a programme for theorists. What is really needed is an evaluation of the accuracy of the approximation of keeping only a one-particle intermediate state in the calculation of matrix elements for the product of two currents. Many theorists profess to believe that such an approximation is quite good, but they really lack any reason to support their stand. It is not possible to compute the degree of validity of this approximation, but if we analyse suitable experiments, we ought to be able to obtain a very good answer to our problem. One should conduct a careful investigation of high-energy Compton scattering off a proton, and of the hyperfine structure of hydrogen. The amplitudes involved in these phenomena are matrix elements of the square of the electromagnetic current. Now Platzman and Iddings have computed the hydrogen hyperfine splitting, keeping only the proton intermediate state and using the Hofstader form factors at each vertex,

and their result disagrees with experiment. However, their work is valuable because it provides an evaluation of the accuracy of this type of approximation, and it is on this basis that I believe our estimate of the radiative correction to nuclear beta decay is good to ½ per cent. It would be useful to have another check on this approximation by studying Compton scattering, and it is likely that such a study will provide additional substantiation for our estimate of the ½ per cent accuracy.

Strangeness changing weak decays — 2nd LECTURE

The beta decay of nucleons and the decay of the muon are only two examples of the weak decays. There are many more and it is now our task to outline the possible terms which are required in the weak interaction coupling in order that all types of decays may be accounted for, at least qualitatively. To do this let us use a more convenient symbolism. Let us abbreviate the V-A current term $[\bar{A}\gamma_\alpha(1 + i\gamma_5)B]$ by $(\bar{A}B)$. In this language, to describe nuclear beta decay, we need a coupling term of the form $(\bar{\nu}e)(\bar{n}p)$, together with its Hermitian conjugate. In our discussion of neutron decay we have seen that such a form is correct only qualitatively. Indications are that the vector current is not renormalized and, consequently, that additional terms for other strongly interacting particles such as pions, K's, etc. must be present.

We have speculated that the vector beta-decay current which is strangeness preserving is a suitable component of the isotopic spin current, which is conserved. For that reason the theory we have discussed is called the conserved vector current (CVC) theory. (It is well to remind ourselves that our main requirement was that the vector coupling constant should not be modified by the strong interactions. An easy way to guarantee this is to propose that the vector current is conserved, but that may not be the only way.) The axial vector coupling constant differs from its ideal value [1.26 instead of 1*)]. Some people consider this a small deviation so that they are tempted

*) Bernardini has told me that Miss Wu now finds the ratio of the axial vector to the vector nuclear beta-decay coupling constants to be 1.16.

to supply some reason for such a relatively small renormalization. Many attempts have been made to construct a conserved axial current, but all have failed because of mass terms. However, it is by no means self-evident that we cannot set up the axial current so that its coupling is essentially unchanged by renormalization, even though such a current will not be conserved.

The decay of the muon shows that the weak coupling must contain a term of the form $(\bar{\nu} e)(\bar{\mu} \nu_\mu)$, and today it is known definitely that ν_μ, which we might call the neutretto, is distinct from ν, the neutrino. Since bound muons disappear faster than free ones, we also know that there must be a term like $(\bar{\nu}_\mu \mu)(\bar{n} p)$. By measuring the degree to which the muon lifetime is shortened in various nuclei, we can get a pretty fair idea of the strength of the muonic lepton coupling to nucleons. The experimental evidence allows one to assume that the strength of this coupling is the same as the electron beta-decay coupling strength, which is also the same as G to a few per cent.

These three terms are sufficient to give rise to many other observed decays. For example, the modes $\pi \to \mu + \nu_\mu$ and $\pi \to e + \nu$ would be expected qualitatively because they can proceed through an intermediate $\bar{N}N$ state which is coupled to leptons*). In fact, all weak decays which conserve hypercharge are predicted by the three types of weak coupling already considered. But to account for the observed strangeness changing decays we need to postulate at least three more types of weak coupling. Consider the following three decays of the K meson: $K \to \mu + \nu_\mu$, $K \to \pi + e + \nu$, and $K \to \pi + \pi$. We see that we need a strangeness changing term, which, without prejudice as to its true nature, we shall abbreviate by $(\bar{\Lambda} p)$, coupled to both types of the leptons, and to strongly interacting particles with zero strangeness: $(\bar{\nu}_\mu \mu)(\bar{\Lambda} p)$, $(\bar{\nu} e)(\bar{\Lambda} p)$, and $(\bar{p} n)(\bar{\Lambda} p)$. These three terms are sufficient because all the strangeness changing weak decays, about which there is no dispute experimentally, obey the selection rule that the change in strangeness equals the change in the charge $\Delta S/\Delta Q = +1$. Now one would guess that the strength of the last three types of

*) For example, π^+ goes virtually to p and antineutron via strong interactions. The proton then goes to N, e^+ and ν, the neutron and antineutron annihilating.

weak coupling might be the same as the first three which conserve hypercharge, but it turns out that this is not the case. The strangeness changing decay rates are weaker by an order of magnitude from what would be expected if the strength G were universal.

The current-current theory

These six terms are somewhat messy and the question is whether things can be organized in a more pleasing way. One idea that pops into view is to combine the four currents into one grand weak interaction current

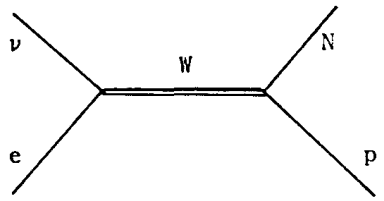
$$J_\alpha = (\bar{\nu}e) + (\bar{\nu}_\mu \mu) + a(\bar{p}n) + b(\bar{p}\Lambda) ,$$

with the coefficients, a and b, to be determined by some symmetry principles, and to suggest that the weak coupling is simply a current-current interaction $1/\sqrt{2}\ G \bar{J}_\alpha J_\alpha$. In the cross-products one finds the six types of terms that are required experimentally. Such a proposal automatically eliminates neutral lepton currents, for which there is no evidence experimentally. One may ask why we write charged currents. The answer is that if we rewrite for example $(\bar{\nu}e)\ (\bar{\mu}\nu_\mu)$ as $-(\bar{\nu}\nu_\mu)\ (\bar{\mu}e)$, and then pursue the idea of a current-current coupling, the decay $K^0 \to \mu + \bar{e}$ would be predicted. Since such a decay is definitely not seen, we feel that the charged current hypothesis is much to be preferred.

The current-current interaction not only leads to the six desired cross terms from the four basic types of current, but also predicts four new types of parity non-conserving interactions, which are the diagonal terms in the product. Three of them will be extremely difficult to observe; for example, the diagonal term $(\bar{\nu}e)\ (\bar{e}\nu)$ leads to a direct parity non-conserving scattering cross-section between neutrinos and electrons, but since the cross-section is so tiny and neutrinos are so hard to detect, the existence of such a term has not yet been verified experimentally. However, experiments in nuclear physics can be designed which are extremely sensitive to the existence of a parity-violating contribution to the nuclear forces. Recently, Felix Boehm at Caltech has measured the circular polarization of the gamma rays produced in a partially forbidded M1 transition in a heavy ellipsoidal nucleus. This polarization can be non-zero only through an admixture of the parity forbidden

El matrix element. Boehm found good evidence for a small parity violation in the nuclear forces. Such an effect is predicted by the $(\bar{n}p)(\bar{p}n)$ diagonal term in the current-current theory. Because of strong renormalizations which are not calculable, it is impossible to state whether or not the size and sign of the experimental effect agree with the prediction of the current-current theory of weak interactions. The order of magnitude of the effect nevertheless, is correct, and hence this does constitute a qualitative verification of the hypothesis.

If the weak interaction does have the current-current form an appealing theoretical possibility is that a new vector meson exists, which mediates the interaction in the same way that the photon mediates the interaction between two charge currents. Such a new field will give rise through the following type of diagram



to a weak interaction between two currents of the form

$$4\pi\ e_W^2\ \bar{J}_\alpha \left(\delta_{\alpha\beta} - q_\alpha q_\beta / m_W^2\right) J_\beta / (q^2 - m_W^2)\ .$$

In this formula, m_W is the mass of the intermediate boson, and e_W is its coupling to the weak interaction current. For $q << m_W$, the interaction reduces to our current-current form once we identify $4\pi\ e_W^2/m_W^2$ with $g/\sqrt{2}$!

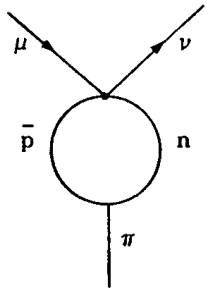
The charged vector meson theory, like the four-fermion point interaction theory, is not renormalizable, but this never bothers me. All of our theories are wrong at high energies, and renormalizability only refers to whether this universal disease can be swept under the rug for the analysis of low-energy phenomena. I personally have felt no particular favour for renormalizable theories, and will just as well accept one that is unrenormalizable as one that is renormalizable. It has never been proved that renormaliz-

able theories are superior because they are consistent; indeed, it seems to me likely that renormalizable theories suffer from ghost difficulties at high energies. I make a point of this because in my later analysis I shall continue to discuss non-renormalizable theories, thus opposing the conventional practice, and shall make no apology for doing so.

From the size of the weak-boson coupling constant e_W, one can estimate the production rate of these particles. Once threshold energy has been passed sufficiently, they should be produced copiously enough to be seen readily if they exist at all.

Pion decay

Let us now turn to the question of the extension of the idea of non-renormalizability from the vector current to the axial current. At the same time we will study the treatment of another prominent non-strangeness changing decay, the decay of the pion. To clarify the ideas involved, we stick to a model of the universe in which nucleons and pions are the only strongly interacting particles. As already explained, one expects the pion to decay into leptons just because of the existence of the following type of diagram:



In the coupling of the pion to the nucleon loop one usually sees the Dirac matrix $g\gamma_5$. However, we shall come back later and discuss an alternate form for the vertex, which I prefer. The nucleon loop is coupled to leptons by both a γ_μ matrix and a $\gamma_\mu\gamma_5$ matrix. For this problem only the $\gamma_\mu\gamma_5$ term contributes, since the pion is a pseudoscalar particle. This loop cannot be calculated because it is divergent. Even if it were not, one would not believe the answer because pion corrections to it would be very important.

However, the form it gives for the matrix element

$$\sqrt{8\pi}\; q_\alpha F_\pi (\bar{\psi}_\mu \gamma_\alpha a \psi_\nu) ,$$

where q is the momentum of the pion, is the only possible invariant form for the amplitude. The number F_π is the result of all possible diagrams. We do not know it nor do we know how to calculate it.

Given this form for the amplitude we can calculate the decay rate using the Golden Rule given in the first lecture. It turns out to be

$$\frac{1}{\tau_\pi} = \left(F_\pi m_\pi\right)^2 \frac{m_\pi}{2} \frac{m_\mu^2}{m_\pi^2} \left[\frac{m_\pi^2 - m_\mu^2}{m_\pi^2}\right]^2 .$$

Using the known decay rate of the charged pion, one finds that

$$F_\pi m_\pi = 5.95 \times 10^{-8} .$$

Incdientally, the same mathematics hold for the decay of kaons into leptons, in which case $F_k m_k = 5.55 \times 10^{-8}$.

To describe the decay of $\pi \to e + \nu$, the amplitude would have the same invariant structure. You will note that if the coupling constants are of the same order for both $(\bar{e}\nu)$ and $(\bar{\mu}\nu_\mu)$ the rate for decay into $(\bar{e}\nu)$ is severely inhibited. The inhibition is due to the fact that the charged lepton must come out with its spin aligned along its direction of motion. This is very difficult for electrons since their velocity is quite high and, in the weak interactions, particles prefer to be polarized in the opposite sense. The probability of right-hand polarization goes as I - V/2. The muons, on the other hand, are not highly relativistic and, consequently, not so greatly inhibited. Theoretically one finds the ratio:

$$\frac{\pi \to e + \nu}{\pi \to \mu + \nu_\mu} = 1.36 \times 10^{-4} ,$$

where this theoretical estimate includes some radiative corrections.

The theoretical estimate for the ratio of the decay rates assumes that the coupling of $(\bar{e}\nu)$ and $(\bar{\mu}\nu_\mu)$ is the same for weak interactions. Experiment agrees with that result to within two per cent; that means that within one per cent the electron and muon couple in the same way in the weak interactions. The pion decay experiment thus provides the best check of the hypothesis that the strength of the muon coupling is the same as the electron coupling in the weak interactions. This is a much better result than is obtained from studying μ capture in complicated nuclei. Note that this check is entirely independent of the value of F_π which is not calculable.

Is there some theoretical argument for determining the pion decay amplitude F_π? It is a delightful problem because it is apparently hopeless. However, there appeared a paper by Goldberger and Treiman, who found a formula for F_π. The argument given by Goldberger and Treiman is inadequate since they ignored terms of the same order as the ones that were kept. Similar formulae had also been discovered before by several other people who had not pursued them because they did not agree well with experiment. Goldberger and Treiman's contribution was to put renormalized constants into the formula, thus obtaining much better agreement with experiment.

In this lecture I should like to describe my first approach to a derivation of the Goldberger-Treiman relation. In the next lecture I shall discuss Gell-Mann's refinement of this approach.

To understand the ideas involved in this derivation, let us recall our attempt to construct the vector beta-decay current in such a way that it was not modified by renormalization. In our model universe of pions and nucleons one can show that if the coefficient A of the pion current, in the combination $(\bar{\psi}_p \gamma_\alpha \bar{\psi}_n) + A(\pi^- \partial_\alpha \pi^0 - \pi^0 \partial_\alpha \pi^-)$ is suitably chosen, the renormalized vector beta-decay coupling constant is the same as the bare vector-coupling constant. An obvious question is whether or not one can write the axial vector current in such a way that there is no renormalization. The answer to that, generally speaking, is that we cannot. Furthermore, there is some renormalization effect. Nevertheless, because the renormalization is small, one is

tempted to construct the axial vector current in such a way that only a small renormalization would result. To study this question let us write a possible Lagrangian for our model universe

$$\mathcal{L} = \tfrac{1}{2}\left[(\nabla\varphi)^2 - m_0^2\varphi^2\right] + \bar{\psi}\,\not{\partial}\,\psi - \bar{\psi}\,M_0\,\psi$$
$$+ a_0\,\bar{\psi}\,\tau(\gamma_5\not{\partial}\,\Phi)\,\psi + e_W\left[\bar{\psi}\gamma_\mu\left(\frac{1+i\gamma_5}{2}\right)\tau_4\,\psi\right]W_\mu\ .$$

For the purpose of this lecture we neglect complications due to isotopic spin. The essential point is to get an idea, with the aid of a single model, which might be true in a more realistic description of nature. The first three terms in the above Lagrangian are those characterizing free meson and nucleon fields. The a_0 term describes the coupling of the nucleon field to the pseudoscalar mesons. The final term describes the coupling of the nucleon fields to intermediate vector bosons. It may be noted that in the pseudoscalar meson-nucleon coupling I have used a gradient form, which is called pseudovector coupling. For single pion interactions at low pion momenta, this form is equivalent to the conventional pseudoscalar coupling $\bar{\psi}\gamma_5\,\psi\varphi$, providing that one identifies the coefficient $2m_N a_0$ with the conventional coupling constant g_0. For two pion interactions there is a great distinction between pseudovector and pseudoscalar coupling. The absence of low-energy s wave scattering is compatible with the prediction of pseudovector coupling and very difficult to explain with pseudoscalar coupling. For this and other reasons I prefer the pseudovector form, and I do not care in the least that the pseudovector form belongs to the class of interactions called unrenormalizable. It is very difficult to check which of these couplings is more correct because calculations cannot be made for strong interactions, except possibly by noting the following fact. If, in any amplitude, we can extrapolate the pion four-momentum off the mass shell down to the point where it is zero (zero momentum and zero energy), then that amplitude should vanish in the case of pseudovector coupling. This is an interesting principle which should aid in developing good trial formulae for pion interactions, but it has not been used very much up to now.

In electricity we have the coupling of charged particles to the electromagnetic field A, in such a way that if one adds to A a pure gradient

$$A_\alpha \rightarrow A_\alpha + \partial_\alpha \lambda$$

it makes no difference. Let us pursue the idea of constructing the coupling for the weak interaction such that if we add a pure gradient $\nabla\lambda$, to the vector boson field W_α, then there is also no effect. For the vector part of the coupling there is no change if we have the W meson coupled to a current which is conserved, as we have already assumed. For the axial vector term the addition of the gradient leads to an extra term in the Lagrangian of the form

$$e_W[\psi_i \nabla\lambda \gamma_5 \tau_+ \psi] \ .$$

Can we cancel this? It turns out that we can, partially, if we use the pseudovector coupling. Suppose that when we change the W field by adding the gradient, we also change the pion field by

$$\varphi \rightarrow \varphi - \frac{e_W}{a_0} \lambda \ .$$

In that case the $a_0[\bar{\psi}\tau_+\gamma_5 \not{\partial}\varphi\psi]$ term will be modified in a way that exactly compensates the change in the weak interaction Lagrangian. The compensation is not exact because of the mass term in the pion field, and it is impossible to produce an exact compensation. However, having lived under the influence of Gell-Mann, who likes to suggest that if all the masses were zero there would be much greater symmetry in the Lagrangian, I shall disregard the mass term. Thus, neglecting the mass squared term in the free-pion field Lagrangian, the change is $-e_W/a_0 \partial_\alpha \varphi \partial_\alpha \lambda$ to the first order and this is compensated if we add a term $(e_W/a_0)\ \partial_\alpha \Phi W_\alpha$ to the Lagrangian. It is in this way that we obtain a definite prescription for the direct coupling of the pion to the weak-vector boson, or equivalently to lepton currents. Combining terms, it is seen that the axial vector current which is coupled to the W field is $e_W(\frac{1}{2}\psi\gamma_\alpha i\gamma_5\psi + 1/a_0\ \partial_\alpha \Phi)$.

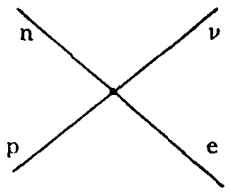
The result of these considerations is to predict that there is a direct bare coupling of pions to leptons with a strength

$$F_\pi \text{ (theoretical)} = \frac{G}{\sqrt{8\pi}\, a_0} = \frac{m_p G}{\sqrt{2\pi}\, g_0}$$

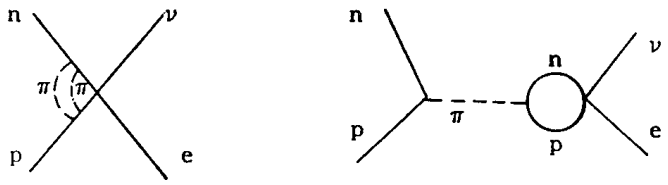
This would be the famous Goldberger-Treiman relation if G and g_0 were renormalized. G and g_0, which are coupling constants before renormalization, are unknown. Thus the problem is: what can be done with this formula? If, without justification, at this point one inserts for G the axial vector coupling constant for nuclear beta decay and for g_0 the experimental pion-nucleon coupling constant, then one finds for F_π (theoretical) the value 5.5×10^{-8}. This value is not very different from the experimental result (5.95×10^{-8}). The difference may have something to do with inaccurate analysis of the effects of renormalization.

We now have to discuss why it is legitimate to replace these bare-coupling constants in the Goldberger-Treiman relation by their renormalized experimental values. To do this we shall have to inquire as to how higher order diagrams with many pions lead to a renormalization of the coupling constants. There is a clever way to do this, by which one can prove the theorem in a few lines. However, since the ideas of renormalization are not self-evident to everyone here, we shall study a more detailed treatment.

We begin by defining some terms. First, the unrenormalized axial vector beta-decay coupling to nucleons, which is $G_V \bar{n} \gamma_\mu \gamma_5 p$, (where we have assumed that $G_{A(\text{bare})} = -G_{V(\text{bare})} = -G_V$), and which is represented by

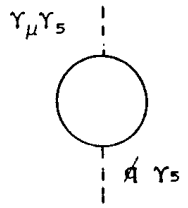


is modified by pion corrections, for example

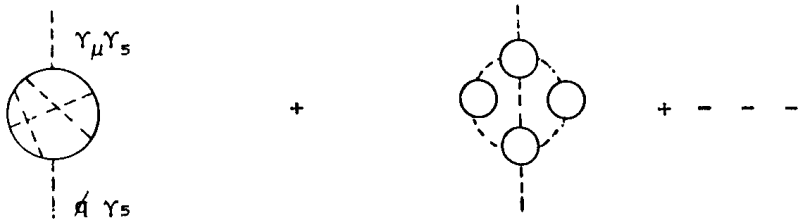


so that the resultant effective beta-decay coupling is $G_A \bar{n} \gamma_\mu \gamma_5 p$. The effective coupling constant is G_A only in those cases where the leptons effectively carry off zero momentum. When the momentum transfer becomes large, the strength of the coupling will be modified. We may indicate this by a function $G_A(q^2)$; $G_A(0) \triangleq G_A$. Theoreticians call the ratio of the bare-coupling constant to the experimental one a renormalization constant Z_A. That is $G_A/G_{A(\text{bare})} = 1/Z_A$.

In our subsequent considerations we shall encounter the following nucleon loop

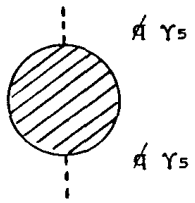


together with all the pion correction diagrams, e.g.



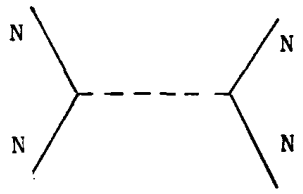
The sum of all these loop diagrams must be a vector. Since the only vector around is q_μ, the momentum transfer at each of the two vertices, then the sum must be represented mathematically by $q_\mu K(q^2)$. $K(q^2)$ does not blow up as $q^2 \to 0$.

One would expect that $K(q^2)$ is slowly varying over momentum transfers comparable to the pion mass, the main variation only occurring for q of order of the nucleon mass. Although it is true that there are intermediate states of mass as low as $3m_\pi$, these do not come in strongly in the pseudovector coupling model. For, if each of the three pion momenta are of order m_π, the coupling is weak (as $f^2 = 0.08$). Strong effective couplings come only for higher values of q, and hence not close to the pole with three intermediate pions. Finally, if we consider the sum of all diagrams of the form

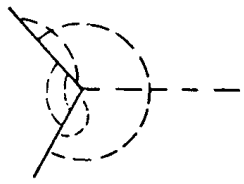


the answer is $q^2K(q^2)$. The blob stands for the sum of all intermediate states except a single pion.

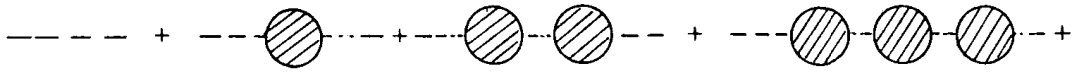
Let us now go on to consider the renormalization of the pion exchange diagram. The simplest diagram giving rise to a nuclear force is



At each vertex the pion couples through $a_0 q\!\!\!/ \gamma_5$, where q is the momentum transfer. There will be corrections to this exchange diagram of two types. First, each vertex will be modified by virtual pions, of which a typical diagram is



Secondly, the propagation of the pion will not be simply ,
but will be



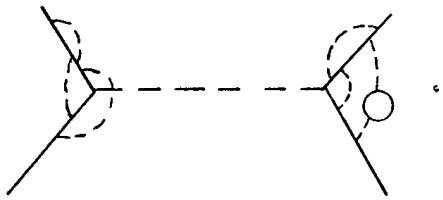
The result of including all these corrections will be the experimental pion-exchange contribution to the scattering amplitude, which is

$$a_{exp}^2 \not{A}\gamma_5 \frac{1}{q^2 - m_\pi^2} \not{A}\gamma_5 .$$

Let us go through this more slowly for the case of pion-nucleon scattering. The single bare-pion exchange diagram gives rise to an amplitude

$$a_o^2 \not{A}\gamma_5 (1/q^2 - m_o^2) \not{A}\gamma_5 .$$

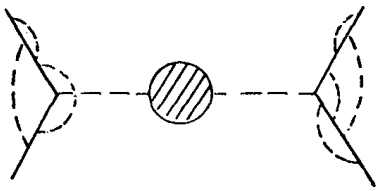
Suppose we consider still, a single bare-pion exchange but correct the vertices to all orders,



The mathematical contribution to the amplitude is

$$\left[\frac{a_o}{Z_A(q^2)}\right]^2 \not{A}\gamma_5 (1/q^2 - m_o^2) \not{A}\gamma_5 ,$$

since, in the discussion of the axial vector beta decay we defined $[Z_A(q^2)]^{-1}$ to be the sum of all vertex corrections for axial vector coupling. Next, if we consider the sum of all diagrams with a single generalized nucleon loop breaking the pion exchange:



we get

$$\left[\frac{a_0}{Z_A(q^2)}\right]^2 \not{A}\gamma_5 \frac{1}{q^2 - \pi_0^2} \; a_0^2 \, q^2 K(q^2) \frac{1}{q^2 - m_0^2} \not{A}\gamma_5 \; .$$

If we add another generalized nucleon loop to the pion propagator, we multiply the preceding amplitude by the factor

$$a_0^2 \, q^2 K(q^2) \frac{1}{q^2 - m_0^2} \; .$$

Evidently, if we add up such diagrams with all numbers of nucleon loops in the pion propagator, then we are just adding up a geometrical series and the final answer is

$$\left[\frac{a_0}{Z_A(q^2)}\right]^2 \not{A}\gamma_5 \frac{1}{q^2 - m_0^2 - a_0^2 \, q^2 K(q^2)} \not{A}\gamma_5 \; .$$

We must now compare this result with the experimental pion-exchange contribution by noting that the amplitude must have a pole at $q^2 = m_\pi^2$, where m_π is the experimental pion mass. We see that

$$m_\pi^2 = m_0^2 + a_0^2 m_\pi^2 \, K(m_\pi^2) \quad \text{or} \quad \frac{m_\pi^2}{m_0^2} = \frac{1}{1 - a_0^2 \, K(m_\pi^2)} \; .$$

Also by expanding the corrected propagator near $q^2 = m_\pi^2$, we find that

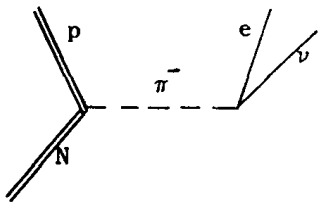
$$[q^2 - m_0^2 - a_0^2 \, K(q^2)] \approx [1 - a_0^2 \, K(m_\pi^2) - a_0^2 \, m_\pi^2 \, K'(m_\pi^2)] \cdot (q^2 - m_\pi^2)$$

so that the experimental pion-nucleon coupling constant is related to the bare pion-nucleon coupling by

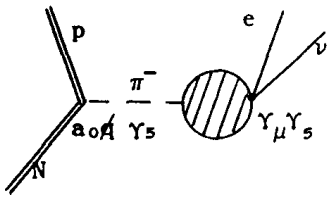
$$a^2_{exp} = \left[\frac{a_0}{Z_A(m_\pi^2)}\right]^2 \frac{1}{[1 - a_0^2\, K(m_\pi^2) - a_0^2\, m_\pi^2\, K'(m_\pi^2)]} \quad .$$

The theoreticians like to call the renormalization factor, $1/[\ ]$, for the pion Z_π.

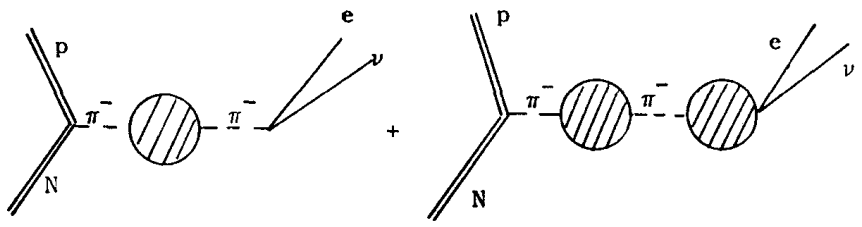
So much for the discussion of the axial vector vertex and the pion propagator renormalization constants. We now turn to an evaluation of the pion decay rate in terms of the nuclear beta-decay axial vector coupling constant. To get all the renormalization constants straight, we will consider the pion decay under the assumption that the pion is produced off a nucleon. Thus, we shall consider the matrix element for a neutron to go into a proton, plus a lepton pair at a momentum transfer near the pion pole. Assuming that there is a direct pion-lepton weak coupling, the simplest diagram contributing to the nuclear beta decay which gives rise to a pole at the pion mass is



The direct nucleon-lepton term gives no pole so we leave it out. Another contribution of the same type is



To these two diagrams we must add those diagrams with two bare-pion propagators



and three, and four and so on. Again the terms can be summed because they form geometric series. The result of all diagrams of this type is

$$\frac{a_0}{Z_A(q^2)} \not{q}\gamma_5 \frac{1}{q^2 - m_0^2 - a_0^2\, q^2 K(q^2)} \left[-\frac{G_{A(\text{bare})}}{a_0}\, q_\alpha + a_0 q_\alpha K(q^2) G_{A(\text{bare})} \right] (\bar{\psi}_e \gamma_\alpha a \psi_\nu) .$$

In the square bracket the coefficient $-G_{A(\text{bare})}/a_0$ is the bare coupling of the pion to the lepton field, which we found earlier. The term $a_0 q_\alpha K(q^2) G_{A(\text{bare})}$ represents the coupling of a bare pion to the generalized nucleon loop times the axial vector coupling of the nucleon loop to the lepton current.

Now this sum of diagrams must represent, near the pion pole, the dominant part of the nuclear beta-decay amplitude, which we can write down using the experimental coupling of pions to nucleons, and the experimental amplitude for a pion to decay into leptons. That is,

$$a_{\text{exp}} \frac{1}{q^2 - m_\pi^2} \sqrt{8\pi}\, F_\pi q_\alpha \bar{\psi}_e \gamma_\alpha a \psi_\nu .$$

Equating the two expressions near the pion pole, one finds

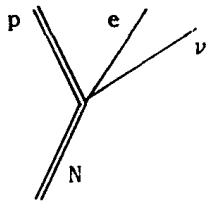
$$F_\pi = -\frac{G_{A(\text{bare})}}{\sqrt{8\pi}\, a_{\text{exp}}\, Z_A(m_\pi^2)} \frac{1 - a_0^2 K(m_\pi^2)}{1 - a_0^2 \left[K(m_\pi^2) + m_\pi^2 K'(m_\pi^2) \right]} .$$

Recalling that the experimental nuclear beta-decay constant is $G_A = G_{A(bare)}\sqrt{Z_A(0)}$ we see that we get the Goldberger-Treiman relation in terms of experimental coupling strength, except for a factor which is a ratio of renormalization functions

$$\frac{Z_A(0)}{Z_A(m_\pi^2)} \cdot \frac{1 - a_0^2\, K(m_\pi^2)}{1 - a_0^2\, [K(m_\pi^2) + m_\pi^2\, K'(m_\pi^2)]} \cdot$$

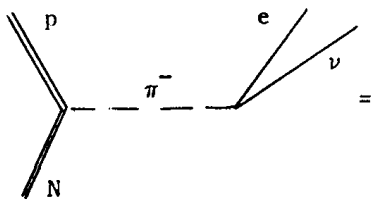
Since the functions $Z_A(q^2)$ and $K(q^2)$ presumably vary appreciably only for q^2 of order m_N^2, the above ratio should be quite close to one. One expects, therefore, that the Goldberger-Treiman relationship should be quite good. Experimentally this relationship holds to eight per cent.

I shall now describe a more sophisticated way to get the same result. Our theory states that, in the limit that the pion mass goes to zero, if the lepton current is replaced by a pure gradient, the axial vector amplitude should then vanish. Let us now consider the beta decay of the nucleon. We would ordinarily write down the amplitude as the sum of two terms. One of them is the direct axial-vector coupling:



$$= G_A(\bar{p}\gamma_\alpha\gamma_5 n)\,(\bar{e}\gamma_\alpha a_\nu) .$$

The other is the coupling through the pion:



$$= a_{exp}(\bar{p}\gamma_5 \not{q} n)\,\frac{1}{q^2 - m_\pi^2}\sqrt{8\pi}\; F_\pi q_\alpha \cdot (\bar{e}\gamma_\alpha a_\nu) .$$

In the limit $m_\pi \to 0$, the sum of these two must vanish when the lepton current $(\bar{e}\gamma_\alpha a\nu)$, is replaced by q_α, i.e.

$$G_A + \sqrt{8\pi}\ F_\pi\, a_{exp} = 0 \ .$$

This is the Goldberger-Treiman relation! The reason we apparently do not have a small correction is that we have skipped over a small point. If the pion mass were zero then the axial-vector coupling constant G_A, and the pion-nucleon coupling constant a_{exp}, would be different from their true values. It is this small change which is represented by the ratio of the renormalization constants that we encountered in the more detailed argument.

We can argue that the coupling strengths would not change materially in the gradient coupling theory if the pion mass went to zero. This is because pions with small momentum are effectively decoupled. Pions with momenta of the order m_π are coupled with a strength of only 0.08. Only when the pion momentum gets as large as a nucleon mass does the coupling tend to 15. When the pions have a high momentum, the fact that they have a mass is unimportant. Thus, the renormalization factors for the coupling due to pions will be the same, to an excellent approximation, when the pions have zero mass as when they have the physical mass. It is for this reason that I believe that the Goldberger-Treiman relation should be accurate.

In working with the field theory model described in the previous lecture we may notice a certain property of the axial vector current. The divergence of this current is proportional to the bare-pion field operator. In taking matrix elements, the axial vector current has to be renormalized in order to get G_A and the pion-field operator suffers a renormalization. These renormalization factors are not too different from unity. When this observation is analysed it results in the Goldberger-Treiman relation.

Recall that the matrix element of the axial current between two nucleons consisted of two terms. One was a $\gamma_\mu\gamma_5$ type coupling and another term, which arose via a virtual pion, was proportional to $q_\mu\gamma_5$. If we now compute the divergence of the axial vector current with $m_\pi^2 \neq 0$, we find that

the contribution of the second term, the induced pseudoscalar term, goes to zero as $q^2 \to 0$. Consider now the matrix element of the pion-field operator between two nucleons. Since the source of the pion field is $ig/2M\ (\bar{n}\not{q}\gamma_5 p)$ and the source is equal to $(\Box^2 - m_o^2)\Phi$, we see that the matrix element of the pion-field operator is equal to

$$\left(\frac{i g_o}{2M}\right)\frac{(\bar{n} q \gamma_5 p)}{q^2 - m_o^2}\ .$$

Equating this expression with the matrix element of the divergence of the axial vector current we get again the Goldberger-Treiman relation.

Gell-Mann abstracted this result by assuming only that the divergence of the axial vector was proportional to the pion field $\nabla_\alpha J_\alpha^A = K\Phi$, without getting the constant of proportionality, K, from my specific theory. He then calculated the divergence of the nuclear-axial vector matrix element in terms of K, and the decay rate of the pion in terms of K. If one considers the ratio of these two amplitudes the factors of K drop out and one gets the Goldberger-Treiman relation.

More recently Gell-Mann has refined the argument by dropping all references to field operators, but I shall not give that argument.

Introduction to SU_3 — 3rd LECTURE

It is time, in our study of the weak interactions, to consider those decays which violate the conservation of hypercharge. Some understanding of these has been obtained with the help of SU_3 symmetry considerations and we therefore turn now to a short discussion of SU_3.

Let me begin by reviewing isotopic spin or SU_2. Suppose we have two particles with identical dynamical properties as far as a large group of interactions are concerned, e.g. the strong interactions, but which differ when another type of interaction, e.g. electromagnetic, is considered. Let them

be a neutral particle $A \equiv A^0$, and a negatively charged particle $B \equiv B^-$, and let us assign to each of them an eigenvalue for the z component of isotopic spin: $+\frac{1}{2}$ for A, and $-\frac{1}{2}$ for B. The particles are assumed to have corresponding antiparticles $\bar{A}$ and $\bar{B}$, the quantum numbers of which are the opposite to those of the corresponding particle. By considering the states of two objects, one can generate states with isotopic spin 1 and 0, from these isodoublets. For example, if we consider new basis states A and B which are related to A′ and B′ by unitary transformations, then the state of an antiparticle and a particle, which is unchanged by the transformation, is $(\bar{A}A + \bar{B}B)/\sqrt{2}$. That state is therefore an isosinglet; i.e. it has I = 0. Under the unitary transformations, the three other orthonormal states $\bar{A}B$, $(\bar{A}A - \bar{B}B)/2$ and $\bar{B}A$, are transformed into linear combinations of each other, so that they form an isotopic triplet, i.e. I = 1. Let us call these three states by a new set of names so that we recall immediately that they form a triplet; in analogy with the Σ hyperons we use the symbol σ:

$$\sigma^- = -\bar{A}B, \quad \sigma^0 = -(\bar{A}A - \bar{B}B)/\sqrt{2}, \quad \sigma^+ = \bar{B}A$$

and in analogy with the Λ^0 hyperon we designate the isotopic singlet state by λ^0:

$$\lambda^0 = -(\bar{A}A + \bar{B}B)/\sqrt{2} .$$

Two basic objects allow us to construct all states which differ in their isotopic spin properties. However, in order to construct states which differ in another quantum number, say hypercharge, we have to introduce a third basic object. Thus, we add to our set of A and B a third particle $C \equiv C^-$, which has the same dynamical characteristics for a restricted set of interactions and which is negatively charged. We define its isotopic spin so that C forms an isotopic singlet. Under unitary transformations generated by isotopic- spin operators, C is not mixed with A and B. We define also a quantum number, e.g. hypercharge, which is -1 for C and 0 for A and B. With three basic objects it is interesting to see what happens when, in analogy to

the procedure for isotopic transformations, we study the effects of unitary transformations on states composed of several of the three particles A, B, C, and the corresponding transformations on their antiparticles $\bar{A}$, $\bar{B}$ and $\bar{C}$.

If we concentrate on the states formed from a particle and an antiparticle, of which there are nine, we find that the state

$$\alpha = (\bar{A}A + \bar{B}B + \bar{C}C)/\sqrt{3}$$

is unchanged by such transformations and thus transforms as a singlet, whereas the other eight orthonormal states are transformed among themselves by a general change of the three basis states. The other eight, therefore, form an irreducible octet, which we display below:

Chart 1

$$-\underset{\xi^-}{(\bar{A}C)} \qquad \underset{\xi^0}{(\bar{B}C)}$$

$$-\underset{\sigma^-}{(\bar{A}B)} \qquad \underset{\sigma^0}{1/\sqrt{2}\,(\bar{B}B - \bar{A}A)} \qquad \underset{\sigma^+}{(\bar{B}A)}$$

$$-\underset{\lambda^0}{1/\sqrt{6}\,[\bar{A}A + \bar{B}B - 2(\bar{C}C)]}$$

$$\underset{n}{(\bar{C}B)} \qquad \underset{p}{(\bar{C}A)}$$

The sign factors assigned to these states are chosen by a convention which is an extension of the Condon and Shortley specification. We have arranged the states so that a state gets transformed only into states in the same row, when the unitary transformations mix only A and B, that is, when we restrict ourselves to isospin transformations. The states are labled ξ^-, n, σ^+ etc. in analogue to the real particles Ξ^-, N, Σ^+ etc. as they have the same quantum numbers. If the SU_3 theory were perfect, then the particle states (or the corresponding particle states of other octets such as K^-, K^0, π^+, etc.) could be substituted for these small letters ξ^-, n, σ^+ in any expression written

below without changing its transformation properties. Likewise antibaryons could be substituted with the corresponding quantum numbers, i.e. $-\bar{p}$ for ξ^-; $\bar{n}$ for ξ^0, $\bar{\lambda}$ for λ, $-\bar{\sigma}^+$ for σ^-, $\overline{\sigma^0}$ for σ^0, $-\overline{\sigma^-}$ for σ^+, $-\overline{\xi^-}$ for p, $\overline{\xi^0}$ for n. The above octet chart is written down immediately just by requiring that the states have the right quantum numbers. The coefficient $-2/\sqrt{6}$ for $\bar{C}C$ in λ is determined so that the λ state is orthogonal to the singlet.

Sum rules for mass splittings

We can generalize the scheme slightly at this point and thus obtain the first order mass sum rule. For this we assign masses to the C and $\bar{C}$ particles which are different from K/2, where we take K/2 to be the mass of A, $\bar{A}$, B and $\bar{B}$. The generalization consists in not identifying $\bar{A}$, $\bar{B}$ and $\bar{C}$ with the antiparticles of A, B and C, but just requiring that $\bar{A}$, $\bar{B}$ and $\bar{C}$ transform in the same way as the antiparticles. Thus, the mass of $\bar{C}$ can be different from that of C. We designate the mass of C by $K/2 + a - b$, and that of $\bar{C}$ by $K/2 + a + b$.

The other virtue of not necessarily identifying $\bar{A}$, $\bar{B}$ and $\bar{C}$ with the antiparticles of A, B and C is that then the ξ^- is not the antiparticle of p and so on. We will assume that the mass of the composite states is just equal to the sum of the masses of the component states. Carrying out this simple calculation we find the following table of masses:

$$
\begin{aligned}
p &= K + b + a \\
\Xi &= K - b + a \\
\Sigma &= K \\
\Lambda &= K + 2/3 \cdot 2a \ .
\end{aligned}
$$

Eliminating a and b leads to a sum rule for the masses:

$$3M_\Lambda + M_\Sigma = 2M_N + 2M_\Xi \ .$$

Note that the b term breaks the mass symmetry for different strangeness, so that if we were constructing an octet of mesons using this scheme we would identify $\bar{A}$, $\bar{B}$ and $\bar{C}$ with the antiparticles of A, B and C, and thus set $b = 0$.

Incidentally, it is worth noting that because the masses are no longer equal, the mass operator connects the singlet state with that state of the octet which has I = 0 and Y = 0. It is thus not diagonal. If we diagonalize it we find new eigenvalues and new eigenvectors which are useful in the interpretation of ω and Φ, two of the vector mesons.

Octet operators

For the analysis of beta decay I would like to find the total isotopic-spin current in terms of my basic set of objects. In fact, with three basic objects I will get eight components of a generalized current instead of just the three that I get by considering isotopic-spin transformations.

The chart of the eight particles given on page 143 also permits us to discover a set of eight operators which transform like an octet. It is only necessary to read the A, $\bar{C}$ etc. as annhihilation of A, creation of C (or creation of $-\bar{A}$, destruction of $\bar{C}$) etc. Thus, the p-like operator (the operator which transforms like p) is $\bar{C}A$ or annihilate A, create C plus annihilate $\bar{C}$ create $-\bar{A}$. That is, the result of the p-like operator on any state is found by taking each term and rewriting it with each A replaced by C, and adding what one gets with each $\bar{C}$ replaced by $-\bar{A}$ (terms with neither A nor $\bar{C}$ in them are to be dropped). For example, the p-like operator on $n = (\bar{C}B)$ converts it to $-\bar{A}B$ or σ^-. This we write as $+(\sigma^- n)$. Again the p-like operator on p itself $(\bar{C}A)$ converts it to $(\bar{C}C - \bar{A}A)$ which is $\sqrt{1/2}\ \sigma^0 + \sqrt{3/2}\ \lambda$, which we write as $+\sqrt{1/2}\ (\overline{\sigma^0} p) + \sqrt{3/2}\ (\overline{\lambda^0} p)$. Proceeding in this way we find the effect of each operator on each member of the octet, with results given in the following chart:

Chart 2

$$p \text{ operator} = \frac{1}{\sqrt{12}}\left[-(\bar{\xi}^{-}\sigma^{0}) - (\bar{\sigma^{0}}p) - \sqrt{3}\ (\bar{\xi}^{-}\lambda) - \sqrt{3}\ (\bar{\lambda}p) - \sqrt{2}\ (\bar{\xi}^{0}\sigma^{+}) - \sqrt{2}\ (\bar{\sigma}^{-}n)\right]$$

$$n \text{ operator} = \frac{1}{\sqrt{12}}\left[-\sqrt{2}\ (\bar{\xi}^{-}\sigma^{-}) + \sqrt{2}\ (\bar{\sigma^{+}}p) + (\bar{\sigma^{0}}n) - (\bar{\xi}^{0}\sigma^{0}) + \sqrt{3}\ (\bar{\xi}^{0}\lambda) - \sqrt{3}\ (\bar{\lambda}n)\right]$$

$$\sigma^{+} \text{ operator} = \frac{1}{\sqrt{8}}\left[-(\bar{n}p) - (\bar{\xi}^{-}\xi^{0}) - \sqrt{2}\ (\bar{\sigma^{0}}\sigma^{+}) - \sqrt{2}\ (\bar{\sigma}^{-}\sigma^{0})\right]$$

$$\sigma^{0} \text{ operator} = \frac{1}{\sqrt{12}}\left[(\bar{\xi}^{0}\xi^{0}) - (\bar{\xi}^{-}\xi^{-}) + (\bar{p}p) - (\bar{n}n) + 2\ (\bar{\sigma^{+}}\sigma^{+}) - 2(\bar{\sigma}^{-}\sigma^{-})\right]$$

$$\sigma^{-} \text{ operator} = \frac{1}{\sqrt{6}}\left[(\bar{p}n) + (\bar{\xi}^{0}\xi^{-}) + \sqrt{2}\ (\bar{\sigma^{0}}\sigma^{-}) + \sqrt{2}\ (\bar{\sigma^{+}}\sigma^{0})\right]$$

$$\lambda \text{ operator} = \frac{1}{2}\left[-(\bar{\xi}^{-}\xi^{-}) - (\bar{\xi}^{0}\xi^{0}) + (\bar{n}n) + (\bar{p}p)\right]$$

$$\xi^{0} \text{ operator} = \frac{1}{\sqrt{12}}\left[(\bar{n}\ \sigma^{0}) - (\bar{\sigma^{0}}\xi^{0}) - \sqrt{2}\ (\bar{\sigma}^{-}\xi^{-}) + \sqrt{2}\ (\bar{p}\ \sigma^{+}) + \sqrt{3}\ (\bar{\lambda}\xi^{0}) - \sqrt{3}\ (\bar{n}\lambda)\right]$$

$$\xi^{-} \text{ operator} = \frac{1}{\sqrt{12}}\left[(\bar{\sigma^{0}}\xi^{-}) + (\bar{p}\sigma^{0}) + \sqrt{2}\ (\bar{\sigma^{+}}\xi^{0}) + \sqrt{2}\ (\bar{n}\sigma^{-}) + \sqrt{3}\ (\bar{\lambda}\xi^{-}) + \sqrt{3}\ (\bar{p}\lambda)\right]$$

Note the operators σ_{+}, σ_{0} and σ_{-} are the isotopic spin operators: $-I_{-}$, I_{Z} and I_{+}, respectively.

You will note from these tables that the σ^{+} operator transforms like the lowering operator for isotopic spin and is thus the operator we would use for the non-strangeness changing beta-decay current, at least for the vector part. Similarly, to describe strangeness changing weak decays, the corresponding operator that we would use would be the operator corresponding to p, which, as you will note, automatically implies that $\Delta S = \Delta Q$.

Thus, SU_3 determines the relative coefficients in the beta decay for us. It turns out, however, that there is another octet of operators, which we shall display later, that may be used for the beta-decay currents. The use of these will introduce another parameter into the analysis of the weak decays, but will not change the isotopic and hypercharge selection rules that are apparent from the octet we have just derived.

The use of such currents to describe beta decay is just a guess and we shall try to check it out in later lectures. This, in fact, is Cabibbo's theory of weak decays.

We have called this set of eight operators an octet. In fact, it is an octet in the sense of SU_3 because if we make transformations among the three basic elements used to construct our scheme, the set of eight operators will transform into each other in exactly the same way that the set of eight composite particles transform into each other.

Reduction of the direct product of two octets

Combining two particles with octet transformation properties produces 64 possible states which, under the various operators listed above in Chart 2, are transformed into one another. From these 64, multiplets may be formed whose states transform only among themselves. These multiplets contain 1, 8, 8, 10, $\overline{10}$ and 27 members. Let the state of two particles be written as $p\lambda$, say, meaning the first is p, the second is λ in its transformation properties. (For example, the first may be an anti Ξ^-; the second, η^0.) The state λp means the first is λ, the second is p (in our example, the first is anti Λ, the second is K^+). We shall write, for short

$$(pn) = (pn - np)$$
$$[pn] = (pn + np) .$$

Since $(\bar{n}n) + (\bar{p}p) + (\bar{\xi}^-\xi^-) + (\bar{\xi}^0\xi^0) + (\bar{\sigma}^+\sigma^+)$ + etc. is evidently invariant, replacing each antiparticle by a particle of the same transformation properties, we get that $\xi^0 n - \xi^- p + n\xi^0 - \sigma^-\sigma^+$ etc. is invariant. Hence, our normalized singlet is

SINGLET

$$\{[\xi^- p] - [\xi^0 n] + [\sigma^+ \sigma^-] - \sigma^0\sigma^0 - \lambda\lambda\}/\sqrt{8}\,.$$

The forms for the p-like operator etc. in Chart 2 permit us, in the same way, to find an octet called the antisymmetric octet.

We denote the states by their quantum numbers $\{Y,I,I_z\}$ within the families.

ANTISYMMETRIC OCTET (8_A)

$$\{1,\tfrac{1}{2},\tfrac{1}{2}\} = \{(p\sigma^0) + \sqrt{2}\,(\sigma^+ n) + \sqrt{3}\,(p\lambda)\}/\sqrt{12}$$

$$\{1,\tfrac{1}{2},-\tfrac{1}{2}\} = \{(\sigma^0 n) + \sqrt{2}\,(p\sigma^-) + \sqrt{3}\,(n\lambda)\}/\sqrt{12}$$

$$\{0,1,1\} = \{(p\xi^0) + \sqrt{2}\,(\sigma^+\sigma^0)\}/\sqrt{6}$$

$$\{0,1,0\} = \{(p\xi^-) + (n\xi^0) + 2\,(\sigma^+\sigma^-)\}/\sqrt{12}$$

$$\{0,1,-1\} = \{(n\xi^-) + \sqrt{2}\,(\sigma^0\sigma^-)\}/\sqrt{6}$$

$$\{0,0,0\} = \{(n\xi^0) + (p\xi^-)\}/2$$

$$\{-1,\tfrac{1}{2},\tfrac{1}{2}\} = \{(\xi^0\sigma^0) + \sqrt{2}\,(\sigma^+\xi^-) + \sqrt{3}\,(\lambda\xi^0)\}/\sqrt{12}$$

$$\{-1,\tfrac{1}{2},-\tfrac{1}{2}\} = \{(\sigma^0\xi^-) + \sqrt{2}\,(\xi^0\sigma^-) + \sqrt{3}\,(\lambda\xi^-)\}/\sqrt{12}$$

The other octet is composed of states which are symmetric under the exchange of the two particles.

SYMMETRIC OCTET (8_S)

$$\{1,\tfrac{1}{2},\tfrac{1}{2}\} = \{\sqrt{3}\,[p\sigma^0] - [p\lambda] - \sqrt{6}\,[n\sigma^+]\}/\sqrt{20}$$

$$\{1,\tfrac{1}{2},-\tfrac{1}{2}\} = \{-\sqrt{3}\,[n\sigma^0] - [n\lambda] + \sqrt{6}\,[p\sigma^-]\}/\sqrt{20}$$

$$\{0,1,1\} = \{-\sqrt{3}\,[p\xi^0] + \sqrt{2}\,[\lambda\sigma^+]\}/\sqrt{10}$$

$$\{0,1,0\} = \{2\,[\lambda\sigma^0] - \sqrt{3}\,[p\xi^-] - \sqrt{3}\,[n\xi^0]\}/\sqrt{20}$$

$$\{0,1,-1\} = \{\sqrt{2}[\lambda\sigma^-] - \sqrt{3}\,[n\xi^-]\}/\sqrt{10}$$

$$\{0,0,0\} = \{[\sigma^0\sigma^0] - [\lambda\lambda] - 2\,[\sigma^+\sigma^-] + [p\xi^-] - [n\xi^0]\}/\sqrt{20}$$

$$\{-1,\tfrac{1}{2},\tfrac{1}{2}\} = \{-[\lambda\xi^0] - \sqrt{3}\,[\sigma^0\xi^0] + \sqrt{6}\,[\sigma^+\xi^-]\}/\sqrt{20}$$

$$\{-1,\tfrac{1}{2},-\tfrac{1}{2}\} = \{-[\lambda\xi^-] + \sqrt{3}\,[\sigma^0\xi^-] - \sqrt{6}\,[\sigma^-\xi^0]\}/\sqrt{20}$$

The remaining twenty states which are antisymmetric under exchange of the two particles fall into two multiplets of ten states. These deciments contain an isotopic singlet, doublet, triplet and quartet, and the hypercharge of the isotopic multiplets differs by one unit in a progressive way. The two decimets are related by a reflection transformation called R, under which: $Y \to -Y$, $I \to I$, $I_z \to -I_z$ and $p \to \xi^-$, $n \to \xi^0$, $\sigma^- \to \sigma^+$, $\lambda \to \lambda$.

DECIMET (10)

$$\{1,\tfrac{3}{2},\tfrac{3}{2}\} = (\sigma^+ p)/\sqrt{2}$$

$$\{1,\tfrac{3}{2},\tfrac{1}{2}\} = \{\sqrt{2}\,(\sigma^0 p) + (\sigma^+ n)\}/\sqrt{6}$$

$$\{1,\tfrac{3}{2},-\tfrac{1}{2}\} = \{\sqrt{2}\,(\sigma^0 n) + (\sigma^- p)\}/\sqrt{6}$$

$$\{1,\tfrac{3}{2},-\tfrac{3}{2}\} = (\sigma^- n)/\sqrt{2}$$

$$\{0,1,1\} = \{(\sigma^+\sigma^0) + \sqrt{2}\ (\xi^0 p) + \sqrt{3}\ (\sigma^+\lambda)\}/\sqrt{12}$$

$$\{0,1,0\} = \{(\sigma^+\sigma^-) + (\xi^- p) + (\xi^0 n) + \sqrt{3}\ (\sigma^0\lambda)\}/\sqrt{12}$$

$$\{0,1,-1\} = \{(\sigma^0\sigma^-) + \sqrt{2}\ (\xi^- n) + \sqrt{3}\ (\sigma^-\lambda)\}/\sqrt{12}$$

$$\{-1,\tfrac{1}{2},\tfrac{1}{2}\} = \{(\xi^0\sigma^0) + \sqrt{2}\ (\sigma^+\xi^-) + \sqrt{3}\ (\xi^0\lambda)\}/\sqrt{12}$$

$$\{-1,\tfrac{1}{2},-\tfrac{1}{2}\} = \{(\sigma^0\xi^-) + \sqrt{2}\ (\xi^0\sigma^-) + \sqrt{3}\ (\xi^-\lambda)\}/\sqrt{12}$$

$$\{-2,0,0\} = (\xi^0\xi^-)/\sqrt{2}$$

DECIMET ($\overline{10}$)

$$\{2,0,0\} = (np)/\sqrt{2}$$

$$\{1,\tfrac{1}{2},\tfrac{1}{2}\} = \{(\sigma^0 p) + \sqrt{2}\ (n\sigma^+) + \sqrt{3}\ (p\lambda)\}/\sqrt{12}$$

$$\{1,\tfrac{1}{2},-\tfrac{1}{2}\} = \{(n\sigma^0) + \sqrt{2}\ (\sigma^- p) + \sqrt{3}\ (n\lambda)\}/\sqrt{12}$$

$$\{0,1,1\} = \{(\sigma^0\sigma^+) + \sqrt{2}\ (p\xi^0) + \sqrt{3}\ (\sigma^+\lambda)\}/\sqrt{12}$$

$$\{0,1,0\} = \{(\sigma^-\sigma^+) + (p\xi^-) + (n\xi^0) + \sqrt{3}\ (\sigma^0\lambda)\}/\sqrt{12}$$

$$\{0,1,-1\} = \{(\sigma\ \sigma^0) + \sqrt{2}\ (n\xi^-) + \sqrt{3}\ (\sigma^-\lambda)\}/\sqrt{12}$$

$$\{-1,\tfrac{3}{2},\tfrac{1}{2}\} = (\sigma^+\xi^0)/\sqrt{2}$$

$$\{-1,\tfrac{3}{2},\tfrac{1}{2}\} = \{\sqrt{2}\ (\sigma^0\xi^0) + (\sigma^+\xi^-)\}/\sqrt{6}$$

$$\{-1,\tfrac{3}{2},-\tfrac{1}{2}\} = \{\sqrt{2}\ (\sigma^0\xi^-) + (\sigma^-\xi^0)\}/\sqrt{6}$$

$$\{-1,\tfrac{3}{2},-\tfrac{3}{2}\} = (\sigma^-\xi^-)/\sqrt{2}$$

The remaining twenty-seven states form a multiplet which is symmetric under the interchange of the particles. It is composed of a triplet with Y = 2, a doublet and a quarter with Y = 1, a singlet, a triplet and a quintet with Y = 0, a doublet and a quarter with Y = -1, and a triplet with Y = -2.

(27)

$\{2,1,1\} = pp$

$\{2,1,0\} = [np]/\sqrt{2}$

$\{2,1,-1\} = nn$

$\{1,3/2,3/2\} = [p\sigma^+]/\sqrt{2}$

$\{1,3/2,1/2\} = \{\sqrt{2}\,[p\sigma^0] + [n\sigma^+]\}/\sqrt{6}$

$\{1,3/2,-1/2\} = \{\sqrt{2}\,[n\sigma^0] + [p\sigma^-]\}/\sqrt{6}$

$\{1,3/2,-3/2\} = [n\sigma^-]/\sqrt{2}$

$\{1,1/2,1/2\} = \{[p\sigma^0] + 3\sqrt{3}\,[p\lambda] - \sqrt{2}\,[n\sigma^+]\}/\sqrt{60}$

$\{1,1/2,-1/2\} = \{-[n\sigma^0] + 3\sqrt{3}\,[n\lambda] + \sqrt{2}\,[p\sigma^-]\}/\sqrt{60}$

$\{0,2,2\} = \sigma^+\sigma^+$

$\{0,2,1\} = [\sigma^+\sigma^0]/\sqrt{2}$

$\{0,2,0\} = \{[\sigma^+\sigma^-] + [\sigma^0\sigma^0]\}/\sqrt{6}$

$\{0,2,-1\} = [\sigma^0\sigma^-]/\sqrt{2}$

$\{0,2,-2\} = \sigma^-\sigma^-$

$$\{0,1,1\} = \{\sqrt{2}\,[p\xi^0] + \sqrt{3}\,[\sigma^+\lambda]\}/\sqrt{10}$$

$$\{0,1,0\} = \{[p\xi^-] + [n\xi^0] + \sqrt{3}\,[\sigma^0\lambda]\}/\sqrt{10}$$

$$\{0,1,-1\} = \{\sqrt{2}\,[n\xi^-] + \sqrt{3}\,[\sigma^-\lambda]\}/\sqrt{10}$$

$$\{0,0,0\} = \{3\,[p\xi^-] - 3\,[n\xi^0] - [\sigma^+\sigma^-] + \sigma^0\sigma^0 + 9\,\lambda\lambda\}/\sqrt{120}$$

$$\{-1,3/2,3/2\} = [\sigma^+\xi^0]/\sqrt{2}$$

$$\{-1,3/2,1/2\} = \{\sqrt{2}\,[\sigma^0\xi^0] + [\sigma^+\xi^-]\}/\sqrt{6}$$

$$\{-1,3/2,-1/2\} = \{\sqrt{2}\,[\sigma^0\xi^-] + [\sigma^-\xi^0]\}/\sqrt{6}$$

$$\{-1,3/2,-3/2\} = [\sigma^-\xi^-]/\sqrt{2}$$

$$\{-1,1/2,1/2\} = \{\sqrt{2}\,[\sigma^+\xi^-] - [\sigma^0\xi^0] + 3\sqrt{3}\,[\lambda\xi^0]\}/\sqrt{60}$$

$$\{-1,1/2,-1/2\} = \{-\sqrt{2}\,[\sigma^-\xi^0] + [\sigma^0\xi^-] + 3\sqrt{3}\,[\lambda\xi^-]\}/\sqrt{60}$$

$$\{-2,1,1\} = \xi^0\xi^0$$

$$\{-2,1,0\} = [\xi^0\xi^-]/\sqrt{2}$$

$$\{-2,1,-1\} = \xi^-\xi^-$$

One way to deduce these multiplets is to take an extreme case, say pp, which has isotopic spin 1 and hypercharge +2, and operate on it with all the raising and lowering operators, which in fact are just the octet of operators that we found at first (Chart 2). Doing this, we would get the multiplet with 27 objects. Next, we can take a state with extreme quantum numbers which is orthogonal to the member of the 27 with the same quantum numbers. We then follow the same procedure of generating the complete multiplet by use of the raising and lowering operators. By iteration of this procedure we will generate all the multiplets.

Couplings of the baryons and mesons — 4th LECTURE

Last time we described how one could construct states of two particles, each one of which was a member of an octet, so that the resultant states would be grouped into multiplets. We found a singlet state, two '8''s, a '10', a '$\overline{10}$' and a '27'. The two '8''s can be distinguished by means of a certain type of inversion transformation which we define in such a way that $p \leftrightarrow \xi^-$, $n \leftrightarrow \xi^0$, $\sigma^- \leftrightarrow \sigma^+$, $\sigma^0 \leftrightarrow \sigma^0$ and $\lambda \leftrightarrow \lambda$. Under this transformation, the first '8' we constructed was antisymmetric and the second '8' was symmetric.

One of the uses of the charts of the previous lecture is to construct couplings of the baryons and mesons which will be invariant under unitary spin transformations. Although this is not directly related to the weak interactions, I would like to pause a moment and discuss that application of the charts. From the construction of a unitary singlet from two '8''s, we know that the combination $-\pi^-\sigma^+ - \pi^+\sigma^- + \pi^0\pi^0 - K^+\xi^- - K^-p + K^0\xi^0 + \overline{K^0}n + \eta\lambda$ is an invariant. We stated before that you can construct certain combinations of two particles that transform like an octet. For example, if we substitute for σ^+, σ^-, σ^0 the antisymmetric σ^+, σ^-, σ^0 operators derived in the last lecture, then we get all the coupling coefficients of the pseudoscalar mesons to the nucleons. Let me write out part of this. To keep the normalization the same, the coupling constant will have to be $\sqrt{6}\,a$, where $a = g/2M$.

$$\sqrt{6}a\Big\{\pi^+\Big[\sqrt{\tfrac{1}{6}}(\bar{n}\gamma_5\not{A}p) + \sqrt{\tfrac{1}{6}}(\bar{\Xi}^-\gamma_5\not{A}\Xi^0) + \sqrt{\tfrac{1}{3}}(\bar{\Sigma}^0\gamma_5\not{A}\Sigma^+) + \sqrt{\tfrac{1}{3}}\,\bar{\Sigma}^-\gamma_5\not{A}\Sigma^0)\Big] + K^-\Big[\sqrt{\tfrac{1}{6}}(\bar{\Xi}^0\gamma_5\not{A}\Sigma^+) + \sqrt{\tfrac{1}{6}}(\bar{\Sigma}^-\gamma_5\not{A}n) + \tfrac{1}{2}(\bar{\Xi}^-\gamma_5\not{A}\Lambda) + \tfrac{1}{2}(\bar{\Lambda}\gamma_5\not{A}p) + \frac{1}{2\sqrt{3}}(\bar{\Sigma}^0\gamma_5\not{A}p) + \frac{1}{2\sqrt{3}}(\bar{\Xi}^-\gamma_5\not{A}\Sigma^0)\Big] + \ldots\Big\} \quad .$$

However, the coupling is not completely determined because one is also allowed to use the symmetric set of eight pairs of baryons. Since there are only two '8''s, there are thus only two constants which characterize the meson-baryon coupling. It is standard to call the coupling constant of the normalized antisymmetric octet of baryons to the mesons $\sqrt{6}$ a F, and to call the coupling constant of the normalized symmetric octet of baryon pairs $\sqrt{10/3}$ a D. The F and D are two parameters which must be discovered by comparison with experiment. There are rough indications from the study of Λ hyperfragments that the ratio D/F is positive and is of the order two or three.

SU_3 in the weak interactions

Returning to the weak interactions we find that we are faced with a slightly more difficult problem. We cannot make use of SU_3 symmetry to construct invariant weak couplings, because the weak Lagrangian must destroy conservation laws. More ingenuity is therefore required to construct the couplings governing the weak decays. As before, one can only try to guess the answer and see if experiment agrees with that guess. The most satisfactory guess that has been made to date is that of Cabibbo, which I shall now introduce in a way that seems most reasonable to me.

Let us consider first, the non-strangeness charging piece of the weak interaction current, because it is this piece that we think we know best of all. We expect from the conserved vector-current theorem that this part of the current should be the isotopic-spin current. The isotopic spin current is just the σ^{+}_{anti}, which is the antisymmetric-octet operator. Considering still the vector current, what would a logical extension of this idea be in order to cover strangeness changing decays? There is some evidence in leptonic decays that this part of the current must violate isotopic spin conservation in such a way that only a violation of a half unit is possible, $\Delta I = \frac{1}{2}$. For example, in the decays $K \to \pi +$ leptons, the prediction of the $\Delta I = \frac{1}{2}$ rule is

$$\frac{K_2^0 \to \pi + \text{leptons}}{K^+ \to \pi + \text{leptons}} = (2/1) \text{ prediction} \quad \square \quad \frac{11{\cdot}1 \pm 1{\cdot}2}{6{\cdot}2 \pm 0{\cdot}9} \text{ experiment.}$$

There is other evidence for this rule but it is not very definitive. Nevertheless, the selection rule is the standard guess in the construction of the current. On the other hand, there is much better evidence for the selection rule $\Delta S = \Delta Q$. Assuming that the strangeness changing piece is also a member of an octet there is only one choice for the current, namely, the operator that transforms like the proton. From our charts you will note that the antisymmetric proton operator obeys both the $\Delta I = \frac{1}{2}$ and the $\Delta S = \Delta Q$ selection rules. We are led, thus, to propose that the vector current is $\sigma^{+}_{anti} + sp_{anti}$.

Experimentally the leptonic decays in which strangeness is violated are 20 times weaker than the non-strangeness changing decays. Therefore, the coefficient s, which measures the proportion of the strangeness-changing current must be definitely smaller than 1. The $K \to \pi e \nu$ is also somewhat smaller than might be expected. We will use this rate later to determine quite accurately the value of s.

Universality and the strangeness changing decays

How do we reconcile weakness of the strangeness-changing decays with the concept of universality. In our study of μ capture, nuclear beta-decay and muon decay, we found that np and the leptons were all coupled with the same strength within a few per cent. This observation, in fact, led to the idea that there was a universal coupling characterizing the weak interactions. However, today we know for example, that p couples much more weakly. Can we construct a theory of the coupling which retains the idea of universality? The answer is yes, and the way to do it was shown by Cabibbo. It is simply that we assume that the weak current is not the component σ^{+} nor the component p, but rather a skew component, but still normalized, i.e. we assume that the vector current is

$$\cos\Theta\ \sigma^{+V}_{anti} + \sin\Theta\ p^{V}_{anti}\ .$$

I have added the superscript V to designate the vector piece of the weak-interaction current. The idea then, is that the skew component is coupled to the leptons with the universal-coupling constant G_μ.

We turn now to the axial current. If there were no complications due to the strong interaction, the obvious choice for the axial current would be to use the same isotopic-coupling coefficients, and merely insert a γ_5. In other words, if we could deal with bare particles, the axial vector part of the current would be

$$\cos \Theta \; \sigma^{+A}_{anti} + \sin \Theta \; p^{A}_{anti} \; ,$$

where the A superscript is to indicate the Dirac matrix $\gamma_\mu\gamma_5$. But, of course, life is more complicated. If SU_3 were perfect, any component of the antisymmetric octet of operators would be conserved. Thus, the vector current would not be changed by renormalization. In fact, Gatto has shown that even to the first order in the mass shifts, the vector current is not renormalized. The axial vector current, however, is not conserved, even in the limit of exact SU_3 symmetry, so one must expect modifications due to renormalization. This means that the strength of the axial vector current will be changed and also that, starting from the antisymmetric octet for the axial vector current, we can get both the antisymmetric and the symmetric octets in the renormalized axial current. Thus, the renormalized axial vector current will take the form:

$$\sqrt{6}\; F_{BBW}\left[\cos\Theta\;\sigma^{+A}_{anti} + \sin\Theta\; p^{A}_{anti}\right] + \sqrt{10/3}\; D_{BBW}\left[\cos\Theta\;\sigma^{+A}_{symm} + \sin\Theta\; p^{A}_{symm}\right] .$$

In the limit of SU_3 symmetry the angle Θ is independent of the particles which are coupled to the weak current, because renormalization cannot change the axis of the octet that is coupled with the weak decays. The renormalization coefficients, F and D are strongly dependent on the type of octets, i.e. pseudoscalar mesons, baryons and vector mesons.

Comparison of the Cabibbo theory with experiment

Let us turn now to a discussion of how well this proposal of Cabibbo agrees with experiment. The most important parameter of the theory is the angle Θ. One of the best ways to determine it is to consider the decay

$$K^+ \rightarrow \pi^0 + e^+ + \nu \ .$$

Two pseudoscalar mesons in a J = 1 state have a parity of (-), and therefore only the vector-weak current is involved in this decay. We have two vectors P_K and P_π, to make the current. The most general form of the matrix element is

$$(P_K + P_\pi) + \xi \ (P_K - P_\pi) \ .$$

In the limit of unitary symmetry the ξ term would be absent. Experimentally, a study of the spectrum shows it to be very small and therefore we will drop it. Putting in the coefficient $1/\sqrt{2}$ from the $K^-\pi^0$ term in the p-like current and the coupling strength $\sqrt{2}$ G, we get for the matrix element:

$$G \sin\Theta \ (P_K + P_\pi)_\alpha \ (\bar{\nu}\gamma_\alpha a e) \ .$$

Making the calculation of the rate and comparing with my data (which may not be the latest, most accurate) I find that $|\sin\Theta| = 0.23 \pm 0.015$ (Cabibbo, using different data in his paper, found $|\sin\Theta| = 0.260 \pm 0.015$).

Another way to determine the angle which Cabibbo also proposed, was to look at the ratio of K and π decays into leptons. Both of these decays proceed only through the axial vector current because the K and π are pseudoscalar mesons. The geometric structure of the matrix element is, as we have already discussed, $\sqrt{8\pi}\ F_\pi q_\alpha^\pi (\bar{\mu}\gamma_\alpha A\nu)$ and $\sqrt{8\pi}\ F_K q_\alpha^K (\bar{\mu}\gamma_\alpha a\nu)$.

There is only one octet that can be made from a single set of mesons. Hence, the Cabibbo theory would state that $F^K/F^\pi = \tan\Theta$.. Experimentally

$$F_\pi = \frac{5.95 \times 10^{-8}}{m_\pi} \quad \text{and} \quad F_K = \frac{5.55 \times 10^{-8}}{m_K} \ ,$$

so that tan Θ is very nearly equal to m_π/m_K. Numerically $|\sin\Theta|$ from the experiments is 0.26 ± 0.02, which agrees with the first determination. This provides the first check on the Cabibbo hypothesis. (Incidentally, in my later calculations I shall use $|\sin\Theta| = 0.245$, because then $\sin^2\Theta = 0.06$, a simple number for computational purposes.)

Another consequence of the Cabibbo theory is that the nuclear beta-decay vector coupling constant G_V, should differ from the muon beta-decay coupling G_μ. In fact, $G_V = G_\mu \cos\Theta$ in the Cabibbo theory. Since $\cos\Theta$ is about 0.97, we see that this theory very possibly resolves this difficulty in the theory of weak interactions!

Determination of the axial vector current F/D ratio

The next question is how well the Cabibbo theory fits with the leptonic decays of the hyperons. To study this let us make a table of the leptonic decay matrix elements, and use our charts of the F and D octets to construct the proportions of vector and axial vector currents in each decay.

Non-strangeness changing decays. These are to be multiplied by $\cos\Theta$.

Reactions	Vector	Axial
$n \to p + e + \nu$	1	F + D
$\Sigma^0 \to \Sigma^+ + e + \bar{\nu}$	$-\sqrt{2}$	$-\sqrt{2}$ F
$\Sigma^- \to \Sigma^0 + e + \bar{\nu}$	$\sqrt{2}$	$\sqrt{2}$ F
$\Sigma^- \to \Lambda + e + \bar{\nu}$	0	$\sqrt{2/3}$ D
$\Sigma^+ \to \Lambda + e + \bar{\nu}$	0	$\sqrt{2/3}$ D
$\Xi^- \to \Xi^0 + e + \bar{\nu}$	1	(F - D)

Strangeness changing decays. These are to be multiplied by sin Θ

Reactions	Vector	Axial
$\Sigma^0 \to p + e + \bar{\nu}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}(F - D)$
$\Lambda \to p + e + \bar{\nu}$	$\sqrt{3/2}$	$\sqrt{3/2}\ (F + \frac{1}{3} D)$
$\Xi^- \to \Sigma^0 + e + \bar{\nu}$	$\sqrt{1/2}$	$\frac{1}{\sqrt{2}}(F + D)$
$\Xi^- \to \Lambda + e + \bar{\nu}$	$\sqrt{3/2}$	$\sqrt{3/2}\ (F - \frac{1}{3} D)$
$\Xi^0 \to \Sigma^+ + e + \bar{\nu}$	1	$(F + D)$
$\Sigma^- \to n + e + \bar{\nu}$	1	$(F - D)$

The best relation we can use to partially determine F and D is the fact that for the decay of the neutron the axial vector coupling is 1.20 ± 0.04, relative to the vector coupling. Thus,

$$F + D = 1.20 \pm 0.04 \ .$$

Strangeness-changing decay rates can now be used to determine other relations between F and D. There is a standard rate for these decays predicted under the assumption that the weak current is just the matrix element of $\gamma_\alpha(1 + i\gamma_5)$. This rate is called the universal Fermi interaction (U.F.I.) rate. Neglecting small relativistic corrections, for any of these decays the rate should be $(V^2 + 3\ A^2/4)$ times the U.F.I. rate. The experimental branching ratios, together with the U.F.I. predictions, are given in the following table.

Decay	U.F.I.	Experimental branching ratios
$\Lambda \to p + e + \bar{\nu}$	1.5×10^{-2}	$0.81 \pm 0.10 \times 10^{-3}$
$\Sigma^- \to N + e + \bar{\nu}$	5.8×10^{-2}	$1.37 \pm 0.34 \times 10^{-2}$
$\Sigma^- \to \Lambda + e + \bar{\nu}$	1.0×10^{-4}	$0.07 \pm 0.03 \times 10^{-4}$
$\Sigma \to \Lambda + \bar{e} + \nu$	0.6×10^{-4}	$0.07 \pm 0.04 \times 10^{-4}$ (4 events).

From the rate of decay of the Λ, we find

$$F + \tfrac{1}{3}D = 0.68 \pm 0.07 \ .$$

There is evidence from the polarization of this leptonic decay that $F + \frac{1}{3} D > 0$. Using this relation, together with the first, we deduce

$$F = 0.40 \mp 0.10, \ D = 0.78 \pm 0.12 \ .$$

Cabibbo's original values were F = 0.30, D = 0.95. The small difference results from slightly different data and from the fact that Cabibbo used 1.26 instead of 1.20 for $-G_A/G_V$

With F and D thus determined, do the other leptonic decays agree with Cabibbo's theory? In many cases only a few events of a given decay have been observed so that the statistical errors and, even more important, the systematic errors in the determination of the rates, are very large. Of the remaining decays, for which experimental information is available, perhaps the reaction $\Sigma^- \to n + e + \bar{\nu}$ is known to the greatest accuracy. Converting the experimental rate for this reaction, one finds $|F - D| = 0.40 \pm 0.16$. You will note that this is certainly consistent with the value $F - D = -0.38 \pm 0.22$ which was predicted above. Another prediction is that the reaction $\Xi^- \to \Lambda + e + \bar{\nu}$ should have a branching ratio of $(0.50 \pm 0.05) \times 10^{-3}$. Experimentally, the branching ratio is two or three times 10^{-3}, but this is based on only a handful of events. The predictions on the $\Sigma \to \Lambda$ decays are also consistent with the crude experimental values.

The D to F ratio for the axial vector current is 2·2 ± 0·8. This is the same ratio as is claimed for the coupling of the pseudoscalar mesons to the baryons, which is thought to be about 2, 3 or 4. This value came originally from very crude theoretical analysis of the strong interactions and some study of the ΛN interaction in hyperfragments. I may remark here that an extension of the Goldberger-Treiman argument to the strangeness-changing current will result in the prediction that the D to F ratio for the strong interactions should be the same as it is for the axial vector weak baryon-baryon current.

Generalization of the Goldberger-Treiman Relation 5th LECTURE

One can explore the question of a supposed equality between the F/D ratios in the strong interactions and in the axial vector weak current by studying a generalization of the Goldberger-Treiman relation. There are several ways to obtain this generalization. One of them would proceed in analogy with my first derivation of the G-T relation for pions. To do this I would extend my model so that it had eight baryons and eight pseudoscalar mesons coupled to baryon pairs using the pseudovector $\not{A}\gamma_5$ coupling. I would then couple both the baryon pairs to the lepton current in such a way that if I neglected the meson masses a divergence added to the lepton current would have no effect. In complete analogy to the pionic case, I would have in lowest order the relation

$$F_{K(\text{unrenormalized})} = \frac{G^{\text{Bare}}_{A(n\Lambda)}}{\sqrt{8\pi}\, a_o(n\Lambda k)} \quad .$$

Renormalization corrections would have to be made of course. It is clear that the structure of the renormalization in the K case would be the same as it is in the π case. That is,

$$F_K = \frac{G_{A(n\Lambda)}}{\sqrt{8\pi}\, a_{(n\Lambda K)}} \frac{Z_{A(n\Lambda K)}(0)}{Z_{A(n\Lambda K)}(m_K^2)} \frac{1 - a_o^2(n\Lambda K)\, K_K(m_K^2)}{\left[1 - a_o^2(n\Lambda K)\, K_K(m_K^2) + m_K^2\, K_K'(m_K^2)\right]}$$

$$= F_{K(\text{unrenormalized})} \times \frac{1 - a_o^2(n\Lambda K)\, K_K(m_K^2)}{\left[1 - a_o^2(n\Lambda K)\,(K_K(m_K^2) + m_K^2\, K_K'(m_K^2))\right]^{1/2}} \quad .$$

One can then attempt to argue that the ratio of the renormalization factors is close to one. This will be far less convincing, however, in the case of the K because m_K^2 is far greater than m_π^2, and m_K^2 is much nearer to the next threshold for intermediate states than m_π^2 was. In the case of the pion,

the distance from $q^2 = 0$ to the pole is m_π^2, whereas the next smallest energy denominator lies at a distance $9\, m_\pi^2$ from $q^2 = 0$. On the other hand, for the K, $q^2 = 0$ is $(500\ \text{MeV})^2$ away from the K pole and only about $(770\ \text{MeV})^2$ away from the $K\pi\pi$ intermediate state. Thus, in the π case we had a factor of 9 in our favour, but in the K case this factor is only about 2. On the other hand, with pseudovector coupling, as we explained before, low momentum pions (and kaons) are only weakly coupled since $f^2 = 0.08$ is small, so these "nearest pole terms" are not expected to be large. I suspect that the functions $Z(q^2)$ and $K(q^2)$ only begin their serious variation near q^2 equal to the nucleon mass squared.

Granting the validity of these arguments that lead to the Goldberger-Treiman relation, the following point can be made. In the limit of SU_3 symmetry $F_\pi \tan\Theta$ should equal F_K, at least if unrenormalized. But the renormalization factors of F exhibited in the second form of the above equation should also be nearly the same. This explains why we can get $\tan\Theta$ from comparison of $\pi\mu\nu$ and $K\mu\nu$ decay. This being the case, the first form of the equation shows that the experimental meson-baryon couplings should be proportional to the experimental axial vector beta-decay couplings. Another way of stating this (since it is assumed that both types of couplings transform as an octet) is that both octets must have the same character. Thus, assuming the Goldberger-Treiman relation, one finds that F/D for the weak axial vector current is the same as F/D for the meson-baryon coupling. Although there are indications that this may be true, there is as yet no sharp experimental test available to check this prediction.

Leptic Decays

Let me summarize what we know about the leptic decays. If there is no strangeness change, we have essentially a complete theory of all the interactions with two coupling constants, G_V and G_A. F_π is supplied via the Goldberger-Treiman relations. When we come to the strangeness-changing decays, we do not know whether we have a good theory or not until we gather more

experimental data. We do have a theory, the Cabibbo theory, which is consistent with the data now available, but the sharpness of the tests is not very great. Cabibbo's theory involves two more parameters: $\sin\Theta$ and the F/D ratio. If some day we can justify more adequately the Goldberger-Treiman relation, then the F/D ratio would come out of a strong interaction theory. The Cabibbo theory would then involve only the parameter Θ. All observed leptic decays then have theoretically predictable properties, with the exception of $K \to \pi + \pi + e + \nu$.

Non-leptic decays

The non-leptic decays also arise from the assumption of a current-current interaction. If we take the theoretical view that the current is the sum of pieces of the form

$$J = (\bar{\mu}\,\nu_{\mu}) + (\bar{e}\,\nu) + \text{"}\sigma^{+}\text{"} \cos\Theta + \text{"p"} \sin\Theta \ ,$$

then the quadratic terms in the expansion of $\bar{J} \cdot J$

$$\sigma^{-}\sigma^{+} \cos^2\Theta + \xi^{-}\, p \sin^2\Theta$$

give rise to non-strangeness changing, non-leptic, weak interactions. It is these terms which give rise to a parity non-conserving nuclear force, for which evidence has been acquired at Cal Tech by Boehm. Other than by experiments of this type, such terms are difficult to get at experimentally.

The cross-products in the expansion of $\bar{J} \cdot J$ are $[\sigma^{-}p + \xi^{-}\sigma^{+}] \sin\Theta \cos\Theta$. The term $\sigma^{-}p$ accounts for those decays for which $\Delta S = +1$ and the term $\xi^{-}\sigma^{+}$, the hermitian conjugate term, accounts for those decays for which $\Delta S = -1$. Since one is the adjoint of the other, we need consider only one of them.

We have a much more difficult task in making predictions for the non-leptic decays because the renormalizations are much more elaborate. In other words, we have a four-fermion coupling term which is much screwed about

by renormalizations of the strong interactions. In the theory of the leptic decays much was predicted on the basis of the hypothesis that the weak current had certain transformation properties. For the non-leptonic decays, on the other hand, the form $\bar{J} \cdot J$ does not have well-defined transformation properties, with regard to unitary spin or even with regard to isotopic spin. Because the situation is not very clear, let us concentrate at first on the transformation properties of the currents.

The strangeness changing decays result from the coupling of an isotopic-spinor current with an isotopic-vector current. With respect to isospin, therefore, the interaction $\bar{J} \cdot J$ has a piece which transforms like isotopic spin $\frac{3}{2}$, as well as one of isotopic spin $\frac{1}{2}$. In the $\Delta S = |1|$ decays, therefore, our current-current interaction leads to the selection rules $\Delta I = \frac{1}{2}$ or $\Delta I = \frac{3}{2}$. The first thing that we must check is whether there is any evidence for $\Delta I = \frac{5}{2}$ in these decays. If one looks at the data to check this, one is surprised and finds a terribly interesting fact. Almost all the data seem to be consistent with only $\Delta I = \frac{1}{2}$ terms in the amplitude. Very small $\Delta I = \frac{3}{2}$ amplitudes are required to fit the data, and no $\Delta I = \frac{5}{2}$ is required at all. The fact that $\Delta I = \frac{3}{2}$ was so small was totally unexpected and has no explanation theoretically on the basis of the current-current hypothesis.

Tests of the $\Delta I = \frac{1}{2}$ rule

In the absence of being able to give a satisfactory theoretical justification for this rule, I will discuss the data which substantiate it. Consider first the disintegration of the Λ. There are two non-leptic channels open

$$\Lambda \to p + \pi^-, \quad \Lambda \to n + \pi^0 \ .$$

On the basis of the $\Delta I = \frac{1}{2}$ rule, since the isotopic spin of the Λ is zero, the charged pion decay amplitude should be $\sqrt{2}$ times the neutral-pion decay amplitude. Hence, the rates should be in the ratio 2 to 1. If there is a decay amplitude into the isospin $\frac{3}{2}$ state of the pion-nucleon system, measured by a coefficient a, then the ratio of the two rates should be $|(\sqrt{2} + a)/(1 - \sqrt{2}a)|^2$.

The experimental branching fraction for the decay into $n+\pi^0$ is 0.35 ± 0.02 of all non-leptic decays. The data, therefore, indicate complete agreement with the $\Delta I = \frac{1}{2}$ rule and show that $|a| < 0.05$. This excellent evidence for the $\Delta I = \frac{1}{2}$ rule is substantiated by the fact that the polarization properties for the neutral and charged modes are the same. The ratio of the s to p-wave amplitudes is measured by the anisotropy of the direction of the pion with respect to the spin of the Λ. The coefficient of this anisotropy term is called α, and if you check the measured values of α for the two modes, you will find that they agree within the experimental errors.

The next baryon to consider is the Σ. There are three non-leptic decays

$$1)\ \Sigma^+ \rightarrow p + \pi^0$$

$$2)\ \Sigma^+ \rightarrow n + \pi^+$$

$$3)\ \Sigma^- \rightarrow n + \pi^- \ .$$

If we label the amplitudes for each process by A_1, A_2 and A_3 respectively, it turns out that the only prediction of the $\Delta I = \frac{1}{2}$ rule is

$$\sqrt{2}\, A_1 = A_2 + A_3 \ .$$

Now each amplitude A_1 is a two-component vector, one component for the s wave and one component for the p wave. The rates are measured by the squares of the lengths of these vectors and are almost all the same. This means that the three vector amplitudes must form an isosceles right triangle with $\sqrt{2}$ A being the base of the triangle. Since the anisotropy coefficients for the second and third decays are almost zero, both amplitudes lie along the coordinate axes in an s, p-plane. Therefore, the first amplitude should lie at 45° to these axes. The anisotropy coefficient α, should therefore be 1 for the reaction. Experimentally, $\alpha = 0.73$. This fact is slightly annoying, but it is not sufficiently bad that we should consider it as evidence against the $\Delta I = \frac{1}{2}$ rule. This measurement is not a very good test of the $\Delta I = \frac{1}{2}$

rule because the experimental errors could be large. The experimental errors will have to be improved a great deal before one can consider the Σ decays to be evidence against the $\Delta I = \frac{1}{2}$ rule.

Another test of the $\Delta I = \frac{1}{2}$ rule comes from the study of $K \to \pi + \pi$. The decay rate for $K^+ \to \pi^+ + \pi^0$ is suppressed relative to $K_1^0 \to \pi + \pi$ by a factor 500. Two pions in a symmetric spatial state must have either I = 0 or I = 2. The wave function for $\pi^+ + \pi^0$, with the spin-parity assignment 0^+, must therefore be I = 2, whereas a neutral 0^+ state of two pions can have both I = 2 and I = 0. Since the K meson has $I = \frac{1}{2}$, the $\Delta I = \frac{1}{2}$ rule predicts that the decay $K^+ \to \pi^+ + \pi^0$ is severely suppressed. Also, the $\Delta I = \frac{1}{2}$ rule predicts that the branching ratio

$$\frac{K_1^0 \to \pi^+ + \pi^-}{K_1^0 \to \pi^0 + \pi^0} = 2/1 \ .$$

Experimentally, the branching fraction for neutral K_1^0 decays is either 0.26 ± 0.02 or 0.335 ± 0.014 depending on which experiment you choose to believe. From the observed rate of $K^+ \to \pi^+ + \pi^0$, we know the amplitude for decay of the K into the I = 2 state. If we take into account the possible interference of the I = 2 with the I = 0 amplitude in K decays, we find that the ratio $K_1^0 \to \pi^0 + \pi^0 / K_1^0 \to \pi + \pi$ need not be exactly 1/3 but may be somewhere between 0.28 and 0.38 (depending on the phase of the interference). Interference effects are thus quite large.

Finally, consider the decays of the K into three pions. Experimentally it is known that the Dalitz plot is almost flat, so that all three pions appear to be in S states. Thus, the spatial wave function is totally symmetric. By use of the generalized Pauli principle, the isotopic spin-wave function must also be totally symmetric. Therefore only I = 1 and I = 3 is possible. There is only one totally symmetric isospin wave function compatible with the $\Delta I = \frac{1}{2}$ rule, and that function has I = 1. The isotopic factors in the rates are given in the following table:

	Weight
$K_2^0 \to \pi^0 + \pi^0 + \pi^0$	3
$K_2^0 \to \pi^+ + \pi^- + \pi^0$	2
$K^+ \to \pi^+ + \pi^0 + \pi^0$	1
$K^+ \to \pi^+ + \pi^- + \pi^+$	4

The ratio of the first two, $3/2$, and the last two, $1/4$ are correct even if $\Delta I = 3/2$ is admitted (for only $\Delta I = 5/2$ can reach $I = 3$). But the relation of the K_2^0 rates to the K^+ rates depends explicitly on the $\Delta I = 1/2$ assumption. Since the available kinetic energy in these three pion decays is small, it is important to include also the change in phase space due to the fact that the charged pion is heavier than the neutral one. The data then, are in accord with predictions, but the tests are not yet very stringent. For example, the ratio of the last two rates given here as $1/4$ is changed to 0.325 by phase-space factors. Experimentally, the ratio is 0.30 ± 0.04. Phase space predicts the ratio of the first two should be 1.89 (instead of $3/2$); experimentally it is 1.6 ± 0.6.

Up to the present point there could have been a mixture of $\Delta I = 1/2$ and $\Delta I = 3/2$. If there were, then the K_2^0 rates would not be determined relative to the K^+ rates. If the $\Delta I = 1/2$ rule is assumed then all four rates are related in the proportion given by the weights in the above table. Another way of putting this is that from the known rate $K^+ \to \pi + \pi + \pi$, one can predict the rate for $K_2^0 \to \pi^+ + \pi^- + \pi^0$. This turns out to be $3.1 \pm 0.2 \times 10^6$/sec. There are two experiments giving $2.9 \pm 0.1 \times 10^6$/sec and $2.3 \pm 0.5 \times 10^6$/sec; both are thus consistent with the $\Delta I = 1/2$ hypothesis.

In summary, the evidence for the $\Delta I = 1/2$ rule is quite good. The $\Delta I = 3/2$ amplitudes are small, but not necessarily zero. There is no evidence for the $\Delta I = 5/2$ amplitude. The intriguing questions are:

a) whether the $\Delta I = \frac{1}{2}$ rule could be an exact symmetry of the weak couplings, and

b) if this is <u>not</u> so, how this rule can become so prominent dynamically.

Attempts to deduce the $\Delta I = \frac{1}{2}$ rule — 6th LECTURE

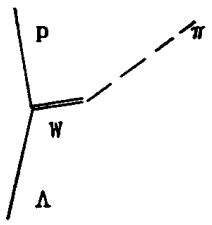
In this lecture we shall present some of the attempts that have been made to deduce the $\Delta I = \frac{1}{2}$ rule for the strangeness changing non-leptic decays. Let me emphasize that this mystery comes just from the isotopic spin properties of the $\bar{J} \cdot J$ form assumed for the weak interactions, and is neither generated nor solved by Cabibbo's choice for the weak interaction current. No answer is known for this mystery. Therefore, we can only list some of the speculations that have been raised and criticize them.

The first interesting question is, if we assume that the strangeness changing non-leptic interaction is of the form

$$(\overline{\sigma^+} p) \cos\Theta \sin\Theta$$

and calculate the amplitudes, then, are the $\Delta I = \frac{3}{2}$ amplitudes dynamically suppressed or are the $\Delta I = \frac{1}{2}$ amplitudes dynamically enhanced, or both? There is no doubt that in the $\overline{\sigma^+} p$ form both the $\Delta I = \frac{1}{2}$ and $\Delta I = \frac{3}{2}$ amplitudes are present with roughly equal weights. However, the strong interactions through renormalization effects will modify the relative weights of the $\Delta I = \frac{3}{2}$ and $\Delta I = \frac{1}{2}$ contributions. It could be that the mesonic corrections are of such a form that the effective strength for the $\Delta I = \frac{1}{2}$ matrix elements is enhanced relative to those for which $\Delta I = \frac{3}{2}$. In addition, we would also like to distinguish between the two cases $\Delta I = \frac{1}{2}$ enhanced, or $\Delta I = \frac{3}{2}$ suppressed. The problem would be very easily answered if we could calculate these renormalizations, but no reliable computation can be made. Therefore, we cannot answer the question in a unique and direct fashion.

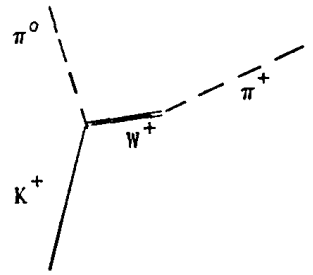
Nevertheless, I would like to report on two little calculations which may or may not be significant for the question. The calculations do represent an attempt to find the order of magnitude of the effects. You may not wish to consider this line of flimsy reasoning; we are becoming very uncertain about this matter, nevertheless, I shall present it. Let us calculate the decay $\Lambda \to p + \pi^-$, presuming that the graph



dominates the amplitude. Both vertices can be obtained from experiment. The $\Lambda p w$ vertex comes from the experimental determination of parameters in the Cabibbo theory of the leptic decays, whereas the $W\pi$ vertex comes from the known decay rate of π^+. If one does this calculation, one finds both $\Delta I = \frac{1}{2}$ and $\Delta I = \frac{3}{2}$ terms of equal magnitude and one gets an estimate for the rate of 8×10^7/sec. The reason that one gets both $\Delta I = \frac{1}{2}$ and $\Delta I = \frac{3}{2}$ terms is that there is no corresponding diagram for the decay $\Lambda \to n + \pi^0$. We now make the assumption that this rate gives a good estimate of the strength of the $\Delta I = \frac{3}{2}$ terms in the amplitudes. Now, if we compare this estimate with the actual rate of 2.4×10^9/sec, we see on the basis of this argument that the $\Delta I = \frac{1}{2}$ amplitude must be considerably enhanced. However, this cannot be the whole story. The ratio of the two rates is one to thirty. That means the amplitude for the $\Delta I = \frac{3}{2}$ is $1/\sqrt{30} = 0.18$, but we already know from the discussion of the experimental validity of the $\Delta I = \frac{1}{2}$ rule that the $\Delta I = \frac{3}{2}$ amplitude is less than 0.045 in magnitude. (There is a way out, perhaps, if the $\Delta I = \frac{3}{2}$ amplitude violates CP. Then there would be no interference term, so that a $\Delta I = \frac{3}{2}$ amplitude this large would not contradict the experimental branching ratio.) The result of this argument is that the $\Delta I = \frac{1}{2}$ term is enhanced and the $\Delta I = \frac{3}{2}$ term is cut down.

However, since the amplitudes from this calculation are off by a factor of five, one way or the other, it is difficult to pretend that this calculation means very much.

Another case in which we can make a similar calculation is the decay $K \to \pi + \pi$. Consider the following diagram



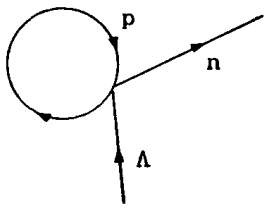
The $K^+\pi^o W$ vertex can be calculated from the $K \to \pi +$ lepton-decay rate and, in a manner similar to the previous case, we find an estimate of 9×10^7/sec, whereas the experimental rate is 1.55×10^7/sec. Thus, the theoretical estimate must be cut down in reality. As explained previously, the K^+ amplitude comes only from the $\Delta I = 3/2$ terms, so that we have another hint that the $\Delta I = 3/2$ amplitudes are suppressed. Furthermore, the experimental rate for $K_1^o \to \pi + \pi$ is considerably greater than what one could get from such an estimate. That seems to indicate that the $\Delta I = 1/2$ amplitude is considerably faster than one estimates from such calculations. Thus, it looks like there is some other type of diagram that must dynamically enhance the $\Delta I = 1/2$ amplitudes.

There are two fundamentally different ways in which people have tried to approach the $\Delta I = 1/2$ rule. The first proposal is that $\Delta I = 1/2$ is an exact rule. To make an exact $\Delta I = 1/2$ rule we have to augment the charged current-current interaction by adding neutral current terms for the strongly interacting particles. If we add a term of the form $(\bar{\sigma}^o n)$ with the right amplitude to the $\bar{J} \cdot J$ form, we then get a perfect $\Delta I = 1/2$ coupling. (We do not want to add neutral currents for the leptons because decays like $K \to \mu + \bar{e}$, $K \to \nu + \bar{\nu}$, and so on are not observed.) This theory results in

two predictions. First, that the $\Delta I = \frac{1}{2}$ rule is exact and secondly, that some day a neutral weak-vector boson might be discovered. Thus, this theory predicts no more than the $\Delta I = \frac{1}{2}$ rule as far as the weak interactions are concerned.

Can the $\Delta I = \frac{1}{2}$ rule be perfect? Since the K^+ does in fact disintegrate into two pions, the $\Delta I = \frac{1}{2}$ rule seems to be violated. People have long speculated that the K^+ decay amplitude results from electromagnetic corrections. If the process is electromagnetic, the rate should be down by a factor $(e^2)^2 \sim 1/20000$, but the observed rate is suppressed relative to the K_1^0 rate by a factor of only 1/500. The question of how to jack up the electromagnetic correction has occupied the minds of many theorists but no one has yet published an explanation that is satisfactory. Thus, no clarity has resulted from the hypothesis that the weak interaction transforms exactly like $\Delta I = \frac{1}{2}$.

The second point of view commonly adopted is that renormalization effects are decidedly different for the $\Delta I = \frac{1}{2}$ and the $\Delta I = \frac{3}{2}$ amplitudes. Many people have tried to argue that there are certain diagrams which are enhanced considerably and which transform only like $\Delta I = \frac{1}{2}$. For example, consider the $\bar{\sigma}^{\mp} p$ cross-term $(\bar{p}\Lambda)(\bar{n}p)$. From such a term there is a diagram of the form



This type of diagram is not defined in perturbation theory, but it certainly may well be that such a term should be included in the correct expansion of the amplitude and it might be that it has a large coefficient. You will note that because the Λ goes directly into the neutron, this diagram gives only a

$\Delta I = \frac{1}{2}$ change. This explains why $\Delta I = \frac{1}{2}$ is so big. At the same time, this argument does not say $\Delta I = \frac{3}{2}$ decays are impossible, but allows such decays to proceed at a small rate with perhaps some renormalizations. However, since we really cannot compute anything, all that such a theory does is restate the experimental data in a different language. This theory thus predicts nothing new.

It seems that the same $\Delta I = \frac{1}{2}$ loop diagrams would result from Cabibbo's complete current. As an exercise you may write out all the terms in the $\bar{\sigma}^{\mp} p$ form that lead to cannibalistic loop diagrams and check to see that they all give $\Delta I = \frac{1}{2}$.

In the octet framework the proposition that $\Delta I = \frac{1}{2}$ terms dominate for strangeness changing non-leptonic interactions has been generalized by many people to the hypothesis that part of the weak Lagrangian transforms like an octet in the eightfold way. In particular the strangeness changing piece transforms like the "n" member of the octet. Since the $\bar{J} \cdot J$ form is the symmetric product of two octets and the only symmetric multiplets contained in the direct product are '1', '8' and '27', the hypothesis consists of completely suppressing '27'. (Terms which transform as '1' do not lead to strangeness changing decays, of course.) Are there any new rules from the assumption of an octet other than those associated with $\Delta I = \frac{1}{2}$? Most unfortunately there are no relations predicted among the reaction amplitudes that can be observed experimentally. Some predictions have been made, but these were based, in addition, on specific assumptions, e.g. there are no gradients in the coupling. This, however, is not in the same spirit as a prediction based only on symmetry principles.

Benjamin Lee derived the relation

$$-A(\Lambda \to p + \pi^-) + 2A(\Xi^- \to \Lambda + \pi^-) = 3A(\Sigma^+ \to p + \pi^0).$$

To derive this relation, however, he assumed R invariance in addition to SU_3. (R invariance is the symmetry based on the transformation $p \to \overline{\Xi^-}$, $n \to \overline{\Xi^0}$, $\Sigma^+ \to \Sigma^-$, and so on. This is simply reflection through the centre of the octet.)

However, because the Ξ's and nucleons have such different masses, R invariance is not in agreement with the facts of nature.

Cabibbo also concluded at one time that the $K_1^0 \to \pi + \pi$ was forbidden by the octet scheme, but this conclusion rested on the additional assumption that there were no gradients in the coupling. Since there is no reason why the matrix element should not involve gradients, the K_1^0 is not forbidden to decay into two pions.

It is rather interesting and disappointing that the assumption of definite SU_3 transformation properties for the non-leptic weak interactions has given us nothing. It seems that a qualitatively different idea is needed to clear up the puzzle of the non-leptic decays. I shall close these lectures with the hopeful speculation that some clever theorist may be able to tie together the fact that CP seems to be violated in weak interactions, and the fact that there are certain unexpected isotopic selection rules in the weak interactions.

Weak decay data summary

Muon Mass, 105.655 ± 0.01 MeV, and lifetime, $2.212 \pm 0.001 \times 10^{-6}$ sec, determine the weak decay constant to be

$$G = (1.0233 \pm 0.0004) \times 10^{-5} \, (\hbar c)^{1/2} \, M_p^{-2}$$

(M_p = proton mass).

Nuclear β decay $O^{14} \rightarrow N^{14}$ (corrected for electromagnetic effects) implies that the value of G for the vector part of the current is 0.980 ± 0.005 times G. Other β-decay experiments, including neutron decay asymmetry and lifetime of 1013 ± 29 seconds, suggest that the axial coupling is -1.26 times the vector coupling.

π^+ [Mass 139.59 ± 0.05; lifetime $(2.55 \pm 0.03) \times 10^{-8}$]. The branching ratio $\pi^+ \rightarrow e^+ + \nu / \pi^+ \rightarrow \mu^+ + \nu$ agrees with predictions of theory to $\pm 2\%$. The Goldberger-Treiman formula for amplitude $\pi^+ \rightarrow \mu^+ + \nu$ is too small by 8%.

K^+ [Mass 493.9 ± 0.2; lifetime $(1.224 \pm 0.013) \times 10^{-8}$ sec]. The branching is $64\% \pm 4\%$ into $\mu + \nu$; $(19 \pm 1)\%$ to $\pi^+ + \pi^0$; $5.7 \pm 0.3\%$ to $\pi^+ + \pi^+ + \pi^-$; $(1.7 \pm 0.2)\%$ to $\pi^+ + \pi^0 + \pi^0$; and $7.6 \pm 1.1\%$ to leptons and π^0. For decay into π + leptons, $K^+ \rightarrow \pi^0 + \mu^+ + \nu$ and $K^+ \rightarrow \pi^0 + e^+ + \nu$, total rate is $6.2 \pm 0.9 \times 10^8$ sec. If the amplitude is written $(p_K + p_\pi) + \xi(p_K - p_\pi)$, then the spectrum of μ indicates that ξ is not large. Ratio of muon to electron decay is: theoretical, $0.65 + 0.13\xi + 0.02\xi^2$; experimental, 0.70 ± 0.06. $K^+ \rightarrow \pi^+ + \pi^- + e^+ + \nu$ is $2.3 \pm 0.7 \times 10^{-5}$ of total disintegration.

K_1^0 [Mass 497.8 ± 0.6, lifetime $(1.00 \pm 0.04) \times 10^{-10}$ sec] goes into $\pi^0 + \pi^0$ or $\pi^+ + \pi^-$. Branching ratio to $\pi^0 + \pi^0$ by two experiments is $26 \pm 2\%$ or $33.5 \pm 1.4\%$.

$\underline{K_2^0}$ (Mass more than K_1^0 by about 0.8 $\hbar/\tau(K_1)$; rate of decay $18 \pm 2 \times 10^6$ disintegration per sec.) For decay into leptons $\pi^{\mp} + \mu^{\pm} + \nu$ and $\pi^{\mp} + e^{\pm} + \nu$ rate is $11.1 \pm 1.2 \times 10^6$ sec^{-1}. For decay into 3π's ($\pi^0 + \pi^0 + \pi^0$ and $\pi^+ + \pi^- + \pi^0$) the rate into charged pions is, by one experiment, $2.9 \pm 0.7 \times 10^6$ sec^{-1} and by another is given as 0.171 ± 0.023 times the total rate into any charged products (therefore pions + leptons). There is some evidence for decay into two pions $\pi^+ + \pi^-$, thereby violating CP invariance at a branching ratio near 2×10^{-3}.

HYPERONS

Hyperon non-leptic decays

Particle	Mass MeV	Lifetime sec	Branching ratio	α	β	γ
Λ	1115.36 ± 0.14	$2.51 \pm 0.09 \times 10^{-10}$	$N + \pi^0$ $1/3$ (to 2%)	$+ 0.56 \pm 0.20$		$+ 0.08 \pm 0.03$
			$P + \pi^-$ $2/3$	$+ 0.62 \pm 0.05$		$+ 0.78 \pm 0.04$
Σ^+	1189.40 ± 0.20	$0.76 \pm 0.06 \times 10^{-10}$	$N + \pi^+$ $51 \pm 4\%$	$- 0.03 \pm 0.08$		
			$P + \pi^0$ $49 \mp 4\%$	$- 0.78 \pm 0.08$		
Σ^-	1197.04 ± 0.30	$1.59 \pm 0.05 \times 10^{-10}$	$N + \pi^-$ 100%	$- 0.10 \pm 0.30$		
Ξ^-	1318.04 ± 1.02	$1.77 \pm 0.05 \times 10^{-10}$	$\Lambda + \pi^-$ 100%	$- 0.51 \pm 0.08 \pm (0.32 \pm 0.17)$		0.80 ± 0.08 (if spin $1/2$)
Ξ^0	1311 ± 8	2.4 to 3.9×10^{-10}	$\Lambda + \pi^0$			

$$\alpha = \frac{2 \text{ Re } S^*P}{|S|^2 + |P|^2}; \quad \gamma = \frac{|S|^2 - |P|^2}{|S|^2 + |P|^2}; \quad \beta = \frac{2 \text{ Im } S^*P}{|S|^2 + |P|^2} \cdot \quad (\alpha^2 + \beta^2 + \gamma^2 = 1)$$

S,P amplitudes for S and P waves.

Hyperon leptic decays branching fraction

	UFI Theory	Experiment
$\Lambda \to p + e + \nu$	$1{\cdot}5 \times 10^{-2}$	$0.81 \pm 0.10 \times 10^{-3}$
$p + \mu + \nu$	0.2×10^{-2}	$\sim 1 \times 10^{-4}$
$\Xi \to \Lambda + e + \nu$	2.0×10^{-2}	$(2 \text{ or } 3) \times 10^{-3}$
$\Sigma^- \to N + e^- + \nu$	5.8×10^{-2}	$1.37 \pm 0.34 \times 10^{-3}$
$N + \mu^- + \nu$	2.6×10^{-2}	$0.66 \pm 0.14 \times 10^{-3}$
$\Sigma^- \to \Lambda + e^- + \nu$	1.0×10^{-4}	$0.07 \pm 0.03 \times 10^{-4}$
$\Sigma^+ \to \Lambda + e^+ + \nu$	0.6×10^{-4}	$0.07 \pm 0.04 \times 10^{-4}$

UFI means rate according to universal coupling constant G times V - A. If actual coupling in G times s(V - tA) then the rate is approximately $s^2/4\ (1 + 3t^2)$ times as much.

Asymmetry and spectrum of Λ decay indicate that t is approximately 1 for this decay.

V.B *Partons — Quarks and Gluons*

After the parity revolution, the great achievement in the field of elementary particle interactions was the discovery of the composite nature of hadrons. Perhaps the earliest suggestion was the Fermi–Yang model of a pion composed of a nucleon–antinucleon pair, later elaborated and extended to strange particles by Sakata and his school by inclusion of the lambda hyperon.[1] A sophisticated approach to compositeness, which for a time was dominant in theoretical physics, was the S-matrix bootstrap hypothesis, or "nuclear democracy," i.e. the doctrine that all hadrons (baryons and mesons) are made of each other. Feynman was following the development of S-matrix theory through dispersion relations, Regge poles, etc., and he learned the lessons they taught about compositeness, but he did not himself participate in this approach. On the other hand, he played an important role in the growth of the strong interaction sector of the Standard Model, according to which hadrons are composed of fermionic quarks interacting through the exchange of bosonic gluons, the quanta of the color gauge field.[2]

Some of the theoretical methods pioneered by Feynman in his studies of QED, meson theory, and quantum gravity, were important, if not essential, for proving the consistency of the gauge theories which are central to both the strong and electroweak sectors of the Standard Model — methods such as Feynman graphs, path integrals, ghost propagators, regularization, and renormalization.[3] While Feynman was not otherwise involved in developing the theory itself, his intuition and wisdom were very valuable to those who were, as were his original insights into gauge theory. We have included a pedagogical account which illustrates this point in the form of lectures that he delivered at a summer school — paper [102].

Feynman's other major contribution to the development and the acceptance of the Standard Model was to analyze the relevant experiments, together with his students and associates at Caltech, and to compare them to theoretical predictions, which were extraordinarily difficult to calculate with accuracy.

Being interested in particle physics at Caltech, Feynman could hardly fail to interact strongly with Murray Gell-Mann and George Zweig, who proposed the existence of the fractionally charged hadronic constituents which they, respectively, called quarks or aces. In terms of these hypothetical spin 1/2 objects they formulated the hadronic currents, introduced by Gell-Mann, through which the weak and electromagnetic interactions of hadrons are expressed. In paper [61], Feynman joined his colleagues in extending Gell-Mann's algebra of quark currents to the group $U(6) \otimes U(6)$. It should be noted that the study of these symmetry groups was the major alternative approach to that of the bootstrap school; Gell-Mann himself was the leading figure in the symmetry school, although he had also been a pioneer in the study of dispersion relations and bootstrapping.

[1] E. Fermi and C.N. Yang, "Are mesons elementary particles?", *Phys. Rev.* **76** (1949): 1739–1743; S. Sakata, "On a composite model for the new particles," *Prog. Theor. Phys.* **16** (1956): 686–688. Other composite models were proposed by M.A. Markov and by M. Goldhaber.

[2] See, e.g., *The Rise of the Standard Model*, edited by L. Hoddeson, L.M. Brown, M. Riordan, and M. Dresden (Cambridge: 1997). Referred to later as *RSM*.

[3] See the theoretical articles in *RSM*, especially those of M. Veltman and G. 't Hooft.

By 1968 a number of physicists had become believers in quarks and were making quark models.[4] Yoichiro Nambu had proposed the color gluon gauge field.[5] Experimentalists working at the Stanford Linear Accelerator Center (SLAC) were on the threshold of discovering quarks.[6] Nevertheless, the theory of quarks was sufficiently new and problematic (for example, the nonobservation of fractionally charged particles and the "paradox" of quark statistics, in the absence of color) that many physicists (including Gell-Mann, but not Zweig) doubted their reality.[7]

A SLAC experimentalist, Michael Riordan, author of a popular history of the quark discovery, wrote that in June of 1968 Feynman "began thinking of each hadron as a collection of smaller parts or entities, which, for lack of any better name, he dubbed simply 'partons'... In this picture the chance of two hadrons colliding was just the sum of the chances of any two of their partons colliding."[8] Riordan reported that when Feynman visited SLAC in August of that year and saw the experimental results on deep inelastic scattering, "Feynman had an epiphany" when he realized that the observations measured "in some way the *momentum distribution* of his partons!"[9] James Bjorken, a theorist at Stanford who advised the SLAC experimenters, had deduced that the unexpectedly large deep inelastic electron cross sections (more like the collision with pointlike charges in the proton than the expected diffuse cloud of mesons) were related to a phenomenon called "scaling," which meant that the results depended on a particular combination of two scattering parameters (energy and momentum transfer), rather than on each independently. However, the theoretical reasoning to arrive at this conclusion was difficult for experimentalists to grasp. Thus the explanation of scaling in terms of Feynman's partons soon became the accepted one.[10]

Papers [75] and [76] deal with very high energy collisions of baryons — say, two protons — analyzed in their center-of-momentum system, with [75] being a much condensed version of [76]. In these papers, Feynman defined two types of measurements which can be made on the results of these collisions: "exclusive," in which all the outgoing momenta are observed, and "inclusive," in which one or a few measurements are made. Because of the high multiplicities which generally occur in these collisions, most of the observations are inclusive and Feynman discussed what can be learned on the basis of such partial observations. His analysis was based on his parton picture, which is a kind of abstraction from quantum field theory. He described his approach as follows:[11]

> It is not that I believe that the observed high-energy phenomena are necessarily a consequence of field theory. Even less do I know what specific field theory could yield

[4]Harry Lipkin, "Quark models and quark phenomenology," in *RSM*, pp. 542–560.

[5]Y. Nambu, "A systematics of hadrons in subnuclear physics," in *Preludes in Theoretical Physics*, edited by A. de Shalit, H. Feshbach, and L. van Hove (Amsterdam: 1966).

[6]For the discovery of quarks, see the 1990 Nobel Lectures of Richard E. Taylor, Henry W. Kendall, and Jerome I. Friedman (general title: *Deep Inelastic Scattering*), *Rev. Mod. Phys.* **63** (1991): 573–595, 596–614, 615–627, respectively.

[7]For the rejection by many theorists of the real existence of quarks, see especially pp. 598–600 of Kendall's lecture (note 6).

[8]Michael Riordan, *The Hunting of the Quark* (New York: 1987), p. 149.

[9]Riordan (note 8), p. 150.

[10]However, Feynman did not publish this explanation until 1972, in item [89]. Bjorken has discussed scaling in "Deep-inelastic scattering: from current algebra to partons," *RSM*, pp. 589–599. This article has two sections headed respectively "Before Feynman" and "After Feynman," referring to the effect of the visit to SLAC mentioned above.

[11]Paper [76], p. 240.

> them. But, rather, I believe that they share some of the properties of field theory, so they might share others... What field theory shall I choose? What shall be the fundamental bare particles that the theory begins with? I do not know and perhaps I do not care. I shall try at first to get results which are more general and characteristic of a wide class of theories and which can be stated in a way independent of the field theory which served as a logical crutch for their discovery.

Papers [83], [90], and [95] are explicitly about the parton theory and its applications to deep probes of hadronic matter and to hadronic collisions. The same is true of most of Feynman's 1972 book *Photon–Hadron Collisions* ([89], not included in the present volume). In paper [90] he extends his analysis of deep inelastic scattering using neutrinos to probe the nucleon, rather than electrons. (Of course, in the former case the probes are virtual massive vector bosons, and in the latter they are virtual photons.) Each time that Feynman lectured he was questioned about whether "parton" was just another name for the quark. Feynman always answered that it was for experiment to decide whether the fundamental constituents carried fractional charge or not. Actually, he liked to point out, the number of fermionic partons was infinite because of the "sea" of virtual pairs which were effective at high excitation energies. In addition, the experiments showed that about half of the longitudinal momentum (in the center-of-momentum system) carried by the proton was in the form of neutral partons, which could be the postulated gluons. The possible relation of quarks and gluons to partons is discussed in some detail in the closing chapters of [89].

When we come to papers [103] and [104], the language of partons has been replaced by that of quarks and gluons. The theory that is used is quantum chromodynamics (QCD), and the fundamental collision processes include not only quark–quark, but also quark–gluon, and even gluon–gluon, scattering. Paper [104], with Rick D. Field and (the later-submitted) paper [103], with Field and Geoffrey C. Fox, are theoretical studies of the production of quark and gluon jets, the latter made before the accumulation of relevant data.[12] These jets arise from cascade processes in which each quark and gluon produced in a collision turns into a rapid stream (i.e. a jet) of hadrons. The analyses of Feynman and his associates exploit the important property of QCD called asymptotic freedom (the vanishing of interaction between quarks having high relative momentum).

Selected Papers

[61] With M. Gell-Mann and G. Zweig. Group $U(6) \times U(6)$ generated by current components. *Phys. Rev. Lett.* **13** (1964): 678–680.

[75] Very high-energy collisions of hadrons. *Phys. Rev. Lett.* **23** (1969): 1415–1417.

[76] The behavior of hadron collisions at extreme energies. In *High Energy Collisions.* London, Gordon and Breach (1969), pp. 237–256.

[83] Partons. In *The Past Decade in Particle Theory.* London, Gordon and Breach (1971), pp. 773–813.

[95] Partons. In *Proc. of the 5th Hawaii Topical Conference in Particle Physics.* Honolulu, Univ. Press of Hawaii (1974), pp. 1–97.

[102] Gauge theories. In *Weak and Electromagnetic Interactions at High Energy.* Amsterdam, North-Holland (1977), pp. 121–204.

[12] For gluon jets, see San Lan Wu, "Hadron jets and the discovery of the gluon," in *RSM*, pp. 600–621.

[103] With R.D. Field and Geoffrey Fox. A quantum-chromodynamic approach for the large-transverse-momentum production of particles and jets. *Phys. Rev.* **D18** (1978): 3320–3343.
[104] With R.D. Field. A parameterization of the properties of quark jets. *Nucl. Phys.* **B136** (1978): 1–76.

GROUP U(6)⊗U(6) GENERATED BY CURRENT COMPONENTS*

R. P. Feynman, M. Gell-Mann, and G. Zweig†
California Institute of Technology, Pasadena, California
(Received 2 November 1964)

It nas been suggested[1-3] that the equal-time commutation rules of the time components of the vector and axial-vector current octets ($\mathfrak{F}_{i\alpha}$ and $\mathfrak{F}_{i\alpha}{}^5$, respectively) are the same as if these currents had the simple form $\mathcal{G}_{i\alpha}$ and $\mathcal{G}_{i\alpha}{}^5$, defined as follows:

$$\mathcal{G}_{i\alpha} = \tfrac{1}{2}i\bar{q}\lambda_i\gamma_\alpha q,$$
$$\mathcal{G}_{i\alpha}{}^5 = \tfrac{1}{2}i\bar{q}\lambda_i\gamma_\alpha\gamma_5 q, \qquad (1)$$

where q is an SU(3) triplet with spin $\frac{1}{2}$—for example, the quarks[2] or aces.[4] Here the matrices λ_i $(i = 1, \cdots, 8)$ are the SU(3) analogs of the Pauli matrices, as defined in reference 1. The operators

$$F_i(t) = -i\int d^3x\, \mathfrak{F}_{i4},$$
$$F_i{}^5(t) = -i\int d^3x\, \mathfrak{F}_{i4}{}^5, \qquad (2)$$

then generate at equal times the algebra of SU(3)⊗SU(3), which may be a very approximate symmetry of the strong interactions,[1,3] while the F_i generate a subalgebra corresponding to SU(3), which is a fairly good symmetry of the strong interactions.

We now propose to extend these considerations to the space components of the currents as well. First we define[1-3] a ninth λ matrix $\lambda_0 \equiv (\frac{2}{3})^{1/2}1$ and a corresponding ninth pair of currents $\mathfrak{F}_{0\alpha}$ and $\mathfrak{F}_{0\alpha}{}^5$ (where $\mathfrak{F}_{0\alpha}$ would be $\sqrt{6}$ times the baryon current in a true quark or ace theory). We then assume that the equal-time commutation relations of all the 72 components of the $\mathfrak{F}_{i\alpha}$ and $\mathfrak{F}_{i\alpha}{}^5$ $(i = 0, \cdots, 8)$ are the same as those of the $\mathcal{G}_{i\alpha}$ and $\mathcal{G}_{i\alpha}{}^5$, at least as far as terms proportional to the spatial δ function are concerned. (There are also, in general, terms[5] involving gradients of the δ function, which vanish on space integration and which we ignore here.) The system of $\mathcal{G}_{i\alpha}$ and $\mathcal{G}_{i\alpha}{}^5$ is closed under equal-time commutation, and the space integrals $\int\mathcal{G}_{i\alpha}d^3x$ and $\int\mathcal{G}_{i\alpha}{}^5d^3x$ generate the algebra of U(6)⊗U(6). Our assumption thus implies that $\int\mathfrak{F}_{i\alpha}d^3x$ and $\int\mathfrak{F}_{i\alpha}{}^5d^3x$ also generate the algebra of U(6)⊗U(6). We assume further that this algebra is a very approximate symmetry of the strong interactions.

We now exhibit some of the structure of the algebra by looking at the $\mathcal{G}_{i\alpha}$ and $\mathcal{G}_{i\alpha}{}^5$. We note that the space integrals of the densities

$$\bar{q}\lambda_i\gamma_4 q = q^\dagger\lambda_i q \quad (i = 0, \cdots, 8)$$

and

$$i\bar{q}\lambda_i\gamma_n\gamma_5 q = q^\dagger\lambda_i\sigma_n q \quad (n = 1, 2, 3)$$

generate the subalgebra corresponding to U(6); the same is then true of the corresponding components of the $\mathfrak{F}$'s. We may refer to the algebra of the space integrals of these $\mathfrak{F}$ components as the A spin, with generators A_γ $(\gamma = 0, 1, \cdots, 35)$. Now the space integrals of

the densities $q^\dagger\lambda_i\frac{1}{2}(1+\gamma_5)q$ and $q^\dagger\lambda_i\sigma_n\frac{1}{2}(1+\gamma_5)q$ also generate a group U(6), and so do the corresponding terms with $\frac{1}{2}(1-\gamma_5)$. The corresponding integrals of $\mathfrak{F}$ components thus give a left-handed A spin A_r^+ and a right-handed A spin A_r^-, respectively, with

$$A_r = A_r^+ + A_r^- \quad (r = 0, 1, \cdots, 35). \tag{3}$$

Those 36 components of $\mathfrak{F}_{i\alpha}$ and $\mathfrak{F}_{i\alpha}{}^5$ (out of a total of 72) that are the densities of the A_r do not go just into themselves under Lorentz transformations, but yield instead the complete system of 72 components of the $\mathfrak{F}_{i\alpha}$ and $\mathfrak{F}_{i\alpha}{}^5$, which form the densities of A_r^+ and A_r^-

We have assumed above that the A^+ and A^- spins are separately very approximate symmetries of the strong interactions. We may now add the further assumption that the total A spin is a good symmetry, nearly as good as the subset that constitutes the F spin. This approximate conservation of A spin is then our way of describing the success achieved by the SU(6) symmetry of Gürsey and Radicati,[6] Sakita,[7] and Zweig,[8] treated further in a series of recent Letters.[9-14] In reference 10, our interpretation of the symmetry is hinted at, but otherwise it is described in different language, which does not make clear the physical identification of the symmetry operators with integrals of components of the vector and axial-vector currents occurring in the weak and electromagnetic interactions. Also, the Lorentz-complete system, obeying the commutation rules of U(6)⊗U(6), is not given.

In a relativistic situation, where a state like ρ exists part of the time as 2π, part of the time as $N+\bar{N}$, part of the time as $\Delta+\bar{\Delta}$, etc., with a different set of channel spins in each case, it is evidently not sufficiently specific to talk of "spin independence" of strong interactions. In contrast, our statement in terms of the approximate conservation of the Gamow-Teller operator $\int\mathfrak{F}_{in}{}^5 d^3x$ $(n = 1, 2, 3)$ does have a definite meaning.

One set of consequences of our approach is that the Gamow-Teller matrix elements within an SU(6) supermultiplet can be exactly computed in the limit of SU(6) symmetry. We adopt the assignments of the $J^\pi = \frac{1}{2}^+$ baryon octet and $J^\pi = \frac{3}{2}^+$ baryon decimet to the SU(6) representation $\underline{56}$, and the assignment of the vector-meson octet and singlet and the pseudoscalar octet to the representation $\underline{35}$; these assignments have explained at least six well-known facts.[15] The axial-vector strength, within the baryon octet, comes out to be $1(D) + \frac{2}{3}(F)$; for the nucleon, this gives $(-G_A/G_V) = 5/3$, as indicated in reference 10, to be compared with an observed value more like 1.2. The agreement is fair, as is the agreement of the D/F ratio with the results on leptonic hyperon decays. The matrix elements of the Gamow-Teller operator between octet and decimet are also exactly specified in the limit of SU(6) symmetry and can be checked by neutrino experiments.

Let us now go on to discuss the badly broken symmetry U(6)⊗U(6), which bears about the same relation to U(6) symmetry as the U(3)⊗U(3) symmetry generated by the time components of vector and axial-vector currents[1,3] bears to the eightfold way. On the way from the full U(6)⊗U(6) down to U(3), we could pass through U(6) or through U(3)⊗U(3) symmetry as an intermediate stage; these are alternatives in somewhat the same way as are L-S and j-j coupling in atomic physics. It seems that the operators of U(6), all of which have nonrelativistic limits, form a much better symmetry system than those of U(3)⊗U(3); hence, the useful procedure is to go from U(6)⊗U(6) to U(6), and then to U(3) and U(2). [Actually U(6) is not much worse than U(3).]

The baryons are presumed to have zero mass in the limit of U(6)⊗U(6) symmetry, as in the limit of U(3)⊗U(3) symmetry.[1,3] The perturbation that reduces the symmetry of U(6) is assumed to transform like $(\underline{6}, \underline{6}^*)$ and $(\underline{6}^*, \underline{6})$ under (A^+, A^-), and like $\underline{1}$ under A. Thus it transforms like a common quark mass term qq, which takes a left-handed q going like $(\underline{6}, \underline{1})$ into a right-handed q going like $(\underline{1}, \underline{6})$, and vice versa.[16] The $J^\pi = \frac{1}{2}^+$ octet and $J^\pi = \frac{3}{2}^+$ decimet belonging to $\underline{56}$ can be placed either in $(\underline{1}, \underline{56})$ and $(\underline{56}, \underline{1})$, or in $(\underline{6}, \underline{21})$ and $(\underline{21}, \underline{6})$, if we restrict ourselves to representations that transform like $3q$. The latter is very attractive, because it splits into a $\underline{56}$ and a $\underline{70}$, where the masses to first order in the perturbation are in the ratio 1:−2; as in reference 3, we must interpret negative mass as positive mass with negative parity, and so we are led to a $\underline{56}$ with unit mass and a $\underline{70}$ with opposite parity and roughly twice the mass. The $\underline{70}$ contains a $\frac{3}{2}^-$ octet, a $\frac{1}{2}^-$ singlet, a $\frac{1}{2}^-$ octet, and a $\frac{1}{2}^-$ decimet. Thus the prediction of reference 3 that the $\frac{1}{2}^+$ octet is accompanied by a $\frac{1}{2}^-$ sing-

let of roughly twice the mass is contained in our present result. The $\frac{3}{2}^-$ octet has probably been seen [including $N(1512)$], but the $\frac{1}{2}^-$ octet and decimet have not so far been identified.

In the limit of U(6)⊗U(6) symmetry, the vector and pseudoscalar mesons of the $\underline{35}$ can be put into either of two pairs of representations that transform like $q+\bar{q}$. The mesons could go like $(\underline{35},\underline{1})$ and $(\underline{1},\underline{35})$, or else like $(\underline{6},\underline{6}^*)$ and $(\underline{6}^*,\underline{6})$. If they belong to the adjoint representation pair $(\underline{1},\underline{35})$ and $(\underline{35},\underline{1})$, as the current components do, then the usual $\underline{35}$ is accompanied by another $\underline{35}$, consisting of a normal axial-vector octet and singlet and an abnormal scalar octet. [Here, "normal" means that the $Y=0$, $I=0$ member of an axial vector, scalar, or pseudoscalar SU(3) multiplet is even under charge conjugation; "abnormal" means it is odd.] If the mesons belong to $(\underline{6},\underline{6}^*)$ and $(\underline{6}^*,\underline{6})$, then the usual $\underline{35}$ is accompanied by a $\underline{1}$ (a normal pseudoscalar singlet), another $\underline{1}$ (a normal scalar singlet), and a $\underline{35}$ consisting of an abnormal axial-vector octet and singlet and a normal scalar octet. In either case, the perturbation that reduces U(6)⊗U(6) to U(6) does not split the mesons into U(6) multiplets in first order; in second order, they are split. The assignment to $(\underline{6},\underline{6}^*)$ and $(\underline{6}^*,\underline{6})$ is appealing because the pseudoscalar singlet could be identified with $\eta(960)$, the scalar octet may include $\kappa(725)$, and the abnormal axial octet may include the meson at about 1220 MeV with $I=1$ that decays into $\pi+\omega$.

*Work supported in part by the U. S. Atomic Energy Commission. Prepared under Contract No. AT(11-1)-68 for the San Francisco Operations Office, U. S. Atomic Energy Commission.

†Work partially supported by the U. S. Air Force Office of Scientific Research under an NAS-NRC Fellowship.

[1]M. Gell-Mann, Phys. Rev. 125, 1067 (1962).

[2]M. Gell-Mann, Phys. Letters 8, 214 (1964).

[3]M. Gell-Mann, Physics 1, 63 (1964).

[4]G. Zweig, CERN Reports No. 8182/TH.401 and No. 8419/TH.412, 1964 (unpublished).

[5]See, for example, J. Schwinger, Phys. Rev. Letters 3, 296 (1959).

[6]F. Gürsey and L. Radicati, Phys. Rev. Letters 13, 173 (1964).

[7]B. Sakita, Phys. Rev. (to be published).

[8]G. Zweig, unpublished.

[9]A. Pais, Phys. Rev. Letters 13, 175 (1964).

[10]F. Gürsey, A. Pais, and L. A. Radicati, Phys. Rev. Letters 13, 299 (1964).

[11]T. K. Kuo and Tsu Yao, Phys. Rev. Letters 13, 415 (1964).

[12]M. A. B. Bég and V. Singh, Phys. Rev. Letters 13, 418 (1964).

[13]M. A. B. Bég and V. Singh, Phys. Rev. Letters 13, 509 (1964).

[14]M. A. B. Bég, B. W. Lee, and A. Pais, Phys. Rev. Letters 13, 514 (1964).

[15]The ratio of p and n magnetic moments, the relation of octet and decimet spacing, the initial degeneracy of φ and ω, the amount of mixing of φ and ω, the equality $m_K{}^2-m_\pi{}^2=m_{K^*}{}^2-m_\rho{}^2$, and the absence of appreciable mixing between η and $\eta(960)$.

[16]At the end of reference 3, it is suggested that perhaps the perturbation in the energy density that reduces U(3)⊗U(3) symmetry to U(3) symmetry generates, together with the algebra of U(3)⊗U(3), a small algebra, which could be that of U(6). Such a use of U(6) is not the same as the use we are discussing in this Letter. However, we might consider the analogous possibility that the perturbation in the energy density that reduces U(6)⊗U(6) to U(6) generates, together with the algebra of U(6)⊗U(6), the algebra of U(12), corresponding to all unitary transformations on the four Dirac components and the three unitary-spin components of a quark field. Even if this is true, of course, U(12) need not be a useful symmetry of strong interactions.

VERY HIGH-ENERGY COLLISIONS OF HADRONS

Richard P. Feynman
California Institute of Technology, Pasadena, California
(Received 20 October 1969)

Proposals are made predicting the character of longitudinal-momentum distributions in hadron collisions of extreme energies.

Of the total cross section for very high-energy hadron collisions, perhaps $\frac{1}{3}$ is elastic and 10% of this is easily interpreted as diffraction dissociation. The rest is inelastic. Collisions involving only a few outgoing particles have been carefully studied, but except for the aforementioned elastic and diffractive phenomena they all fall off (probably as a power of the energy at high energy). The constant part of the total inelastic cross section cannot come from them. And we know that at such energies, the majority of collisions lead to a relatively large number of secondaries (perhaps the multiplicity increases logarithmically with energy). These collisions have not been studied extensively because, with the large number of particles, so many quantities or combinations of quantities can be evaluated that one does not know how to organize the material for analysis and presentation.

It is the purpose of this paper to make suggestions as to how these cross sections might behave so that significant quantities can be extracted from data taken at different energies. These suggestions arose in theoretical studies from several directions and do not represent the result of consideration of any one model. They are an extraction of those features which relativity and quantum mechanics and some empirical facts[1] imply almost independently of a model. I have difficulty in writing this note because it is not in the nature of a deductive paper, but is the result of an induction. I am more sure of the conclusions than of any single argument which suggested them to me for they have an internal consistency which surprises me and exceeds the consistency of my deductive arguments which hinted at their existence.

Only the barest indications of the logical bases of these suggestions will be indicated here. Perhaps in a future publication I can be more detailed.[2]

Supposing that transverse momenta are limited in a way independent of the large z-component momentum of each of the two oncoming particles in the center-of-mass system (so $s = 2W^2$), an analysis of field theory in the limit of very large W suggests the appropriate variables to use for the various outgoing particles in comparing experiments at various values of W in the c.m. system. They are the longitudinal momentum P_z in ratio to the total available W, i.e., $x = P_z/W$, and the transverse momenta Q in absolute units.

Differential cross sections for $dx\,d^2Q$ of the various outgoing particles will then have simple properties as a function of W. Negative x means particles with P_z negative.

First we must distinguish exclusive and inclusive experiments. In exclusive experiments, we ask that certain particles, with given x and Q, be formed, and no others. An example is a two-body charge-exchange reaction. A typical exclusive reaction is

$$A+B\rightarrow\sum_1^n C_i+\sum_1^{n'} D_j,$$

where A is to the right, B to the left, and $C_1, C_2, \cdots, C_n$ are definite particles with definite Q's and x's all moving to the right ($x>0$), whereas the $D_1, D_2, \cdots, D_n$ are moving to the left ($x<0$). The cross section should then vary, for sufficiently high W, as $s^{2\alpha(t)-2}$ or $(W^2)^{2\alpha(t)-2}$, where $\alpha(t)$ is the α of the highest Regge trajectory capable of converting the quantum numbers of A into those of the sum of the C's, and t is the transverse momentum difference of A and the sum of the C's.

This is an evident expectation from Regge theory and should be as approximately valid as that theory is for two-body reactions. The additional point here is the clarification of using the variable x to compare experiments at various energies. If no unitary quantum number is required to be exchanged, so the C group has the same quantum numbers as A, this C group can rise from A via diffraction dissociation and the cross section should approach a constant ratio to the elastic cross section at this value of t (i.e., it should be constant if the elastic cross section is).

Next, an inclusive experiment is one in which we look for special particles with x, Q in the final state but we allow anything else to be produced also. An example is a measurement of the mean number of K^+'s produced with given Q, x in a pp reaction. Such cross sections should approach a constant as $W\rightarrow\infty$.

How can these be reconciled? Why does the cross section fall if, for example, in a two-body reaction we must exchange 3-component of isotopic spin? Because under such circumstances, the current of 3-component isospin must suddenly reverse from right moving to left moving. Thus if any fields are connected to such currents as sources, they would be expected to radiate (in a manner analogous to bremsstrahlung). To be an exclusive experiment (say, pure two-body), we require that no such radiation occur, a condition becoming more and more difficult to satisfy as the energy rises and the current reversal is sharper.

This leads us to expect, in the majority of collisions, many particles over a wide range of x, but their characteristics for the smaller values of x are easy to envision. By Lorentz transformation, the fields to be radiated are becoming narrower and narrower in the z direction as W rises. The energy in this field is therefore distributed in a δ function in z. Fourier analyzed, this means that the field energy is uniform in momentum, dP_z. Since each particle of mass μ carries energy $E=(\mu^2+P_z{}^2+Q^2)^{1/2}$, if we suppose that the field energy is distributed among the various kinds of particles in fixed ratios (independent of energy W), we conclude that the mean number of particles of any kind and of fixed Q is distributed as dP_z/E for not too large x. That is, the probability of finding, among all the emitted particles, a particle of kind i, transverse momentum Q, and mass μ_i is of the form $f_i(Q, P_z/W)dP_z d^2Q/(\mu_i{}^2+Q^2+P_z{}^2)^{1/2}$, where $f_i(Q,x)$ is ultimately independent of W and has a limit $F_i(Q)$ for small x. As $W\rightarrow\infty$, for any finite x, dP_z/E becomes dx/x, of course.

Because of this dx/x behavior, the mean total number (or "multiplicity") of any kind of particle rises logarithmically with W. We need not decide what are "primarily emitted units" and what are secondaries arising from their decay, for the results so far stated are in a form that does not depend on that. If we imagine some primary independently emitted units, however, their number $\bar{n}$ would also rise logarithmically with energy, and the probability that none of them would be emitted might be $e^{-\bar{n}}$ (as suggested by a Poisson distribution) which would then fall as a power of the energy, accounting for the Regge expressions which we are supposing are valid for such exclusive collisions.

We can extend this idea to other amplitudes which involve a similar $\bar{n}$. In particular, we find that the probability that $A+B\rightarrow C+$ anything should vary as $(1-x_C)^{1-2\alpha(t)}dx_C$, where C is moving to the right with almost all the momentum of A (that is, for $1-x_C$ small). Here $\alpha(t)$ is the highest trajectory (excluding the Pomeran chukon) which could carry off the quantum numbers (and squared momentum transfer t) needed to change A to C.

Thus the Regge trajectory function $\alpha(t)$ appears not only in an interaction (as $s^{2\alpha-2}$) but also in an emission process, reminiscent of the close re-

lationship of virtual interaction and real emission that Yukawa emphasized.

Finally, for those special reactions which are partially exclusive, in which anything can be emitted except that a hadron must be transferred from the right-moving to the left-moving system (carrying its fermionic and half-integral spin character) the cross section should vary as $1/s$. Of this last conclusion, I am less sure than of the others.

---

[1]I have taken the approximate existence of Regge poles, the constant total cross sections, and the constant transverse-momentum distributions as empirical facts.

[2]For a somewhat more detailed description, see R. P. Feynman, in Proceedings of the Third Topical Conference on High Energy Collisions of Hadrons, Stony Brook, N. Y., 1969 (to be published).

The Behavior of Hadron Collisions at Extreme Energies

RICHARD P. FEYNMAN

California Institute of Technology, Pasadena, California

Talk given at

Third International Conference
on High Energy Collisions

State University of New York, Stony Brook

September 5-6, 1969

high energy hadron scattering. Firstly, most theoretical inventions are based on analysis of simple collisions, in which only a small number of particles come out. But it is at once realized that questions of unitarity, the asymptotic behavior for high energy in dispersion integrals, etc., require some ansatz be made for the higher energy collisions, in order to close the infinite hierarchy of equations which result. Secondly, experiments at high energies usually yield many particles, and only by selecting the rare collision can we find those about which the theorist has been speaking. For the highly multiple inelastic collisions (to which the major part of the inelastic cross section is due) so many variables are involved that it is not known how to organize or present this data. Any theoretical suggestion (even if it proves to be not quite right) suggests a way that this vast amount of data may be analyzed. For this reason I shall present here some preliminary speculations on how these collisions might behave even though I have not yet analyzed them as fully as I would like.

Hadron phenomena possess a number of remarkably simple properties. Besides the well-known agreements with relativity, analyticity, unitarity, etc., there are of course the conservation rules of isospin and strangeness and an approximate agreement with SU_3 symmetry. In addition to these, however, we have some special regularities which appear to be true empirically which apply to very high-energy behavior and which may point the way to an ultimate dynamical theory. A partial list of these regularities at very high energy are:

a constant fraction of this.

(2) Exchange reaction cross sections fall with a power of the energy $s^{2\alpha-2}$.

(3) This power, α, seems to be t-dependent, and for those $t = m^2$ where $\alpha(t)$ is an appropriate integer (or half-integer) there are resonant particles of mass m (t is the square of the four-vector momentum transfer in the collision).

(4) $\alpha(t)$ varies with t as a straight line $\alpha(t) = \alpha_o + \gamma t$. Although α_o varies with the quantum numbers which must be exchanged, the value of γ is the same (0.95 per $(\text{GeV})^2$) for all.

(5) Cross sections fall very rapidly with transverse momentum transfer. The average transverse momenta of the particles in inelastic collisions are limited (to about 0.35 GeV).

There are, of course, very many rough approximations in this brief summary. For example, we do not know if the total x-section might not vary very slowly (for example, like $\ln s$); all the slopes of Regge trajectories are not exactly equal; whether the pion nonet is on a trajectory of such a slope is unknown; there are corrections to the simple "Regge expectations" (2), (3), presumably for absorption; some of the inelastic cross sections can be associated with diffraction dissociation of the elastic part; etc. Nevertheless, the list above contains a number of main phenomena whose general behavior must ultimately be understood. It will be noticed that in discussing the power laws associated with Regge behavior, I have explicitly separated the behavior of the total x-section which is often described as the Pomeranchuk trajectory,

for it is not certain that this is a typical trajectory. In trying to understand the meaning of these regularities of high energy behavior, I have been led to suggest certain further regularities accompanying them. I should like to present these guesses here to see if they are possibly true, or, if some of them are obviously in disagreement with experiment, to learn where I may have already gone off the track in my thinking.

I shall, for completeness, first explain how I am trying to go about analyzing these things; second, describe some special considerations dealing with Regge exchange; and finally, present my suggestions for the limiting behavior of cross sections at high energies.

I. FIELD THEORY AT HIGH ENERGY

In order to think about these questions I wish to use concepts which will immediately insure that the most fundamental properties of relativistic invariance, quantum superposition, unitarity, etc., will automatically be satisfied. The only theoretical structure I know which has a chance of doing this is a quantum field theory (I say, a chance, because we are not sure if unrenormalized field theories can give finite answers, or if renormalized theories are still unitary). It is not that I believe that the observed high-energy phenomena are necessarily a consequence of field theory. Even less do I know what specific field theory could yield them. But, rather, I believe that they share some of the properties of field theory, so they might share others. Therefore, I wish to study the behavior expected from field theory for collisions of very high energy. What field theory shall I choose? What shall be the fundamental bare particles that

the theory begins with? I do not know and perhaps I do not care. I shall try at first to get results which are more general and characteristic of a wide class of theories and which can be stated in a way independent of the field theory which served as a logical crutch for their discovery. Only later, possibly, might it be worth trying to see if certain special experimental details imply something about a special theory which underlies all these phenomena.

In the meantime I call the fundamental bare particles of my underlying field theory "partons" (which may be of several kinds, of course). For example, in quantum electrodynamics the partons are bare electrons and bare photons. Imagine this theory to have a Hamiltonian H which may be separated into two pieces $H = H_{free} + H_{int}$, one to represent free partons and the other the interaction between them. This H is, at first, expressed in terms of creation and annihilation operators (a^*, a) of these partons (or, if you prefer, local field operators of the parton fields in space). Next, for special application to collisions of extremely high energy, W, incoming in the z-direction in the center-of-mass system, only the operators of finite transverse momenta (i.e., x, y components) are kept as $W \to \infty$. The ones with positive z-component of momentum of order W (i.e., $P_z = x\,W$, where x is a finite quantity as we take a limit as $W \to \infty$) are separated from those of negative z-component. The first are called *right movers* (a_R) and the second *left movers* (a_L). If this is done, and the Hamiltonian reexpressed, we get $H = H_R + H_L + H_I$ where H_R is the Hamiltonian involving right movers (a_R^*, a_R) only (containing, of course, interaction terms among these right movers coming from H_{int}), H_L involves left movers (a_L^*, a_L) and H_I contains both a_R and a_L, and

represents an interaction between the objects moving the left and those moving to the right. But H_I, as $W \to \infty$, becomes a very simple expression (depending on the theory of course). For example, a collision of a right moving proton with a left moving neutron yielding right and left moving particles can then be analyzed in a simple way. The proton is an eigenfunction of H_R, the neutron of H_L. Neglecting H_I, the system of states before and after collision are complete eigenfunctions (not simple partons) of $H_R + H_L$. The operator H_I makes the transition between them (not necessarily in first order, of course).

I leave to a later publication a more detailed description of how this might be carried out, but here I need only make some remarks about the variables on which things appear to depend in the limit as $W \to \infty$.

To describe a proton of momentum $\underset{\sim}{P}_o$, energy E_o, ordinary field theory gives a state function or wave function giving the amplitude that a number of partons of 3-momenta $\underset{\sim}{P}_1$, $\underset{\sim}{P}_2$, . . . etc., are to be found in it. The total momentum of these partons $\sum_i \underset{\sim}{P}_i$ equals $\underset{\sim}{P}_o$ the momentum of the proton, but their total energy $\sum_i E_i$ (where each energy E_i is calculated from the mass μ_i of the parton via $E_i = \sqrt{\mu_i^2 + \underset{\sim}{P}_i \cdot \underset{\sim}{P}_i}$) is not equal to that of the final proton E_o. In fact, the amplitude to find this state contains, among many other factors, one which is inversely proportional to this energy difference

$$A \sim (E_o - \sum_i E_i)^{-1} \qquad (1)$$

Knowing this wave function completely for some P_o, say at rest, how can we find it at some other momentum? It is very difficult to do, and in fact it requires knowledge of the entire Hamiltonian operator H, for the wave function is not a relativistic invariant. This is emphasized

by the point that the momentum is the sum of the momenta of the parts but the energy is not. The wave functions that should be useful for us are those in which $\underset{\sim}{P}_o$ is very large in the z-direction and finite in the directions perpendicular to that. If we take $P_{oz} = x_o W$ and measure the parton's momentum in the z-direction in the same scale $P_{iz} = x_i W$, then the wave function has a definite limiting form as $W \to \infty$ for x_o, x_i finite. (x_o, of course, is arbitrary; it may be taken to be unity, for example.) We have

$$x_o = \Sigma x_i \quad . \tag{2}$$

Let the components of momentum perpendicular to z be called $\underset{\sim}{Q}$, a two-dimensional vector; then

$$\underset{\sim}{Q}_o = \Sigma \underset{\sim}{Q}_i \quad . \tag{3}$$

Finally we can see how A varies (insofar as its denominator behaves) by writing it as

$$A \sim ((E_o - P_{oz}) - \sum_i (E_i - P_{iz}))^{-1}$$

which equals (1) since the total z-momentum is the sum of its parts. However, for large z-momentum,

$$E - P_z = (m^2 + P^2 + P_z^2)^{1/2} - P_z = (m^2 + Q^2 + x^2 W^2)^{1/2} - x W$$

$$\approx \frac{1}{2W} \left(\frac{m^2 + Q^2}{x} \right) \quad ,$$

where m is the mass of the particle. Therefore, the amplitude becomes

$$A \sim 2W \left[\left(\frac{m_o^2 + Q_o^2}{x_o} \right) - \sum_i \left(\frac{m_i^2 + Q_i^2}{x_i} \right) \right]^{-1} \quad . \tag{4}$$

When the numerator as well as the denominator are expressed in these variables x, Q only a simple power of W appears in A which can be removed by a renormalization of the scale of P_z in phase space integrals. The conclusion that I wish to remark for our present purposes is: when expressed in terms of Q, the transverse momentum in absolute scale, and x, the longitudinal momentum in relative scale, the wave functions have definite limits as W, the energy scale of the longitudinal momentum of the state goes to ∞.* In this limit the values of x are positive only. They represent the fractions of the momentum x_o which each parton has (thus, if $x_o = 1$, all x_i run from 0 to 1). The reason x_i must be positive is that if a parton has a negative momentum $P_z = Wx$ with x negative or $P_z = -|x|\ W$ the energy E_i is approximately $+|x|\ W$ so $E - P_z$ is $2\ |x|\ W$ which is of order W^2 larger than the $E - P_z$ of positively moving partons. Appearing in the denominator of amplitudes, such amplitudes are of order $1/W^2$ smaller and vanish in the limit W infinite.

It is seen that this rule (that x be purely positive and also that the amplitude depends only on x) is not valid in an asymptotic way (for $W \to \infty$) when $|x|$ is as small as order c/W. The masses and transverse momenta seem to be of the order of 1 GeV so we shall call an $|x|$ of order (1 GeV)/W a "wee" x. For example, for $P_o = W = 100$ GeV in the center-of-mass system (taking $x_o = 1$), $x = 1/10$ is a small x, but it is not wee; a wee x would be $x = 0.005$ say. The behavior of the amplitude for wee x is not without interest, as we shall see, but for the present we shall discuss only x which is not wee. Then the amplitude as $W \to \infty$ appears

*The statement is not precisely correct. What is meant is the density matrix has definite limits.

to be a function of the Q's and x's of the parts.

In a high-energy collision, the initial state consists of one of these groups of partons moving to the right interacting with another similar group moving to the left. What is the interaction? It will not be enough to just naively say that one parton has a cross section for colliding with another, for, in field theory, interaction is represented by mediation of a field; in fact, by the exchange of just another parton coupled by the piece H_I of the Hamiltonian. But this H_I in its form, is not entirely independent of the form of H_R for they both come from operations on the original Hamiltonian H. Thus the amplitude to interact via the exchange a parton is closely related to the amplitude that there is some parton in the right moving system in the initial state that can be "mistakenly" considered as really being a parton belonging to the left moving system. (Just as two electrons interact in first order by exchange of a photon, so it is also true that a right moving physical electron of $x_o = 1$ has a first order amplitude to be an ideal electron of longitudinal momentum $1 - x$ (times W) and a photon of momentum x. Interaction results if we consider that the left moving electron (insofar as it is bare) has some amplitude to be a left moving outgoing physical electron if we make it up of a left bare electron and the aforementioned photon.) However, a parton of momentum xW to the right would be moving backwards in the left system and would have practically no amplitude (as $W \to \infty$) to be "mistakenly" considered as belonging to this left system. This is true, of course, only if x is not wee (of order 1 GeV/W). If x is wee, a right moving parton and a left moving parton are very similar in appearance. Thus interaction occurs only through exchange of partons or systems of partons of wee longitudinal momentum.

The energy dependence of reactions thus depends upon the probability of finding wee partons of a certain nature. A great deal of information on wee partons can be gotten from a knowledge of the partons where x is not wee, but only small. For the small x and wee x behavior must join in a continuous fashion. For example, suppose, for small x the amplitude to find a parton system with x small varies as $x^{-\alpha} dx$ where x is some constant ($\alpha < 1$). Then the amplitude to find a wee parton of momentum $\sim 1/W$ in dx would, to fit on, have to vary as $(1/W)^{-\alpha}$, but the range of x that such wee partons occupy is of order $1/W$ so that the amplitude to find a wee parton must vary as $W^{\alpha-1}$, if W is the momentum of the right moving object. If E_1 is the energy of the right mover, this amplitude is $(E_1)^{\alpha-1}$. If this is to exchange with a similar system moving to the left with energy E_2 the amplitude that this parton system is acceptable to the other system is $(E_2)^{\alpha-1}$. The amplitude for exchange therefore is $(E_1)^{\alpha-1}\ (E_2)^{\alpha-1}$ or varies as $S^{\alpha-1}$ since $S = E_1\ E_2$. The cross section (there is always a problem of convention of the normalization of amplitudes) varies as the square of this, or $S^{2\alpha-2}$. A constant cross section means $\alpha = 1$ or the amplitude to find partons of small x must vary as dx/x. The amplitude to find a wee parton is not $\int_{0?}^{1/W} dx/x$ because this dx/x law fails below $1/W$. The amplitude to find a wee parton is just constant, independent of energy since the curve $1/x$ cuts off for x below $1/W$ at a value of order W and the integral below $x = 1/W$ is finite in this event. Since cross sections (such as the total x section) are constant, we see that this must actually happen. It is, therefore, not strictly true that as $W \to \infty$ if we keep all x and Q constant there is a definite limiting wave function (as we said earlier) for there is always

a finite amplitude for wee partons. However, the finite x part of the probability distribution of partons has a definite limit, in this limit there are probabilities of finding partons varying as dx/x. The apparently diverging character of this distribution for small x is cut off at wee x / (that is at x of order $1/W$). (A complete theory would have to describe this cutoff region in detail. We shall say more about it later.)

The equations for x not wee simplify if one concentrates on the small x part. It is then seen that there is an approximate scaling law for small x (the approximation improving as x decreases) so that solutions with special distributions of partons with a power law scale dependence ($x^{-\alpha}$) are eigenfunctions natural to field theory.

It may help to give a few, nearly trivial examples. First, according to first order perturbation theory in the expression (4) for the amplitude, the numerator does not depend on x, Q for scalar partons (couplings involve no momenta). If one of the partons has an especially low x, the term $(\mu^2 + Q^2)/x$ belonging to it dominates and we get an amplitude proportional to x (times the scale dP/E, or dx/x, of relativistic phase space). This corresponds to $\alpha = 0$ for the scalar meson. Likewise, it can be shown that the amplitude for (longitudinally polarized) vector partons varies as constant (times dx/x). In this case a factor $1/x$ comes from the numerator couplings. For spin 1/2 particles coupled in the simplest ways, the amplitude varies as $x^{1/2}$ (times dx/x). In general, α equals the spin of the particle. In perturbation theory, these agree with well-known results for the energy dependence of x sections, in particular that vector meson exchange as in electrodynamics lead to constant cross sections in perturbation theory.

The deep inelastic behavior of the electron-proton scattering has been looked at from the point of view described here. It can be argued that the curve of $\frac{\mu^2}{Q^2} W_2$ vs. $Q^2/2M\nu$ (in the variables of Bjorken*) is the distribution in x of charged partons (each weighed by the square of its charge). A behavior like dx/x for the low momentum partons is indicated by experiment.

II. REGGE BEHAVIOR

By "Regge behavior" is meant the second item of our list of regularities, that the cross section for exchange reactions vary as an (inverse) power of the energy, which power depends on the momentum transfer.

The original expectation of Regge behavior were the results of a brilliant induction from Regge's non-relativistic studies by Gell-Mann and Frautschi. Now, however, I should like to consider it as an established empirical fact and to try (in this section) to understand physically why it should be so.

Let us consider a typical exchange reaction, for example, a charge exchange reaction like $p + \pi^- \rightarrow n + \pi^o$, in which the π^-, π^o are right moving, and the p, n left moving. The easiest view of this is that a negative charge has been exchanged from the π system to the nucleon system. The cross section falls about as s^{-1} or according to the best estimates as $s^{2\alpha_o - 2}$ with $\alpha_o = 0.43$ or $s^{-1.14}$. It might at first be thought that an exchange via a vector meson such as a ρ^- would lead

*SLAC publication No. 571.

(as it does in perturbation theory) to a constant cross section. However, it is to be noted that an important current density (the 3 component of isotopic spin) has been suddenly reversed. Initially (π^-, p) / this current density had fast moving components of -1 to the right, +1/2 to the left. Afterward (π^o, n) it has 0 to the right, -1/2 to the left. Although the total 3 component of isospin is not changed, a motion of a part of it (-1 unit) is suddenly changed from right to left motion. In electrodynamics we are aware that a sudden reversal of electric current density induces a copious Bremsstrahlung -- a sudden reversal of the direction of an electron carrying a photon field to the right, leaves the field coasting on to the right in the form of photons (and, of course, the new motion of the charge to the left generates left moving photon Bremsstrahlung).

The hadrons may act similarly. These currents are of considerable importance in our present theories and in fact we believe there are particles (ρ mesons in fact) strongly coupled to them. Thus the strong current reversal in a charge exchange will tend to shed ρ mesons. Perhaps we can guess some of the behavior of the high-energy inelastic collisions by working by analogy to Bremsstrahlung.

In studying the pure reaction $p + \pi^- \rightarrow n + \pi^o$ at high energy, we insist, first that a current be suddenly reversed and, second, that no Bremsstrahlung actually occur. This latter is because we insist that the reaction have only the two particles n, π^o in the fixed state. We can interpret the fact that in such an exchange the cross section falls (relative to the main behavior of the majority of inelastic scatterings -- for which a constant cross section is empirically more appropriate) as the energy rises, by the observation that as the energy rises it

becomes increasingly less likely that the current reversal can be accomplished without Bremsstrahlung.

The theory of Bremsstrahlung with strong coupling and with the "photons" of the field carrying the very type of currents which are sources of further Bremsstrahlung has not been worked out in detail. Nevertheless, we may boldly try to guess that certain analogies to electromagnetic weak interaction Bremsstrahlung exist. Some hope for sense here comes from noting that many features can be seen from a classical view which takes $\hbar \to 0$ so $e^2/\hbar c$ is large. Therefore some properties are understandable both for $e^2/\hbar c$ large, and for $e^2/\hbar c$ small, may have more general validity. This is especially likely if we understand the reasons for them clearly.

First, the spectrum of the particles in longitudinal momentum is dP_z/E (or dx/x for x's which are not wee). This is because the Lorentz contracted field is so sharp in z that the energy in it is distributed uniformly in P_z (the Fourier transform of a pulse being a constant). The energy distribution of the radiated particle is therefore dP_z, or if the individual particles have energies E their longitudinal momenta are distributed as dP_z/E. Also, such a distribution is stable under further disintegration of the particles, or of interaction between the particles. If we write $\tanh w = P_z/E$ for z component of the velocity of a particle, the relativistic rule for the addition of velocities becomes, as is well known, simply the addition of rapidities. Suppose a particle in its rest system can disintegrate or yield a new particle with rapidity u. Then if the old particle has rapidity w, instead, the new particle appears with rapidity $w' = w + u$. Therefore, if the old particles are distributed

uniformly in w (as dw) the new particles appear with a distribution also uniform in rapidity because $dw' = dw$. Thus this feature (a spectrum $dw = dP_z/E$) is to be expected independently of whether we are seeing what was originally radiated (say, ρ mesons) or are observing some other secondary particles that these may have changed into (e.g., if ρ's go into kions). Hence I believe we should expect it for our inelastic distribution of particles of small (and wee) x in our strongly interacting systems also. (We cannot expect this to be valid for large x, say $x = 1/2$ also, because if that much energy is taken from the primary particle by radiation of one emitted particle, subsequent emissions are severely affected. The dP_z/E spectrum for photons in electrodynamics is precisely valid only for smaller values of x.)

Next, the energy in the field thus radiated is some fraction of the energy of the particle which radiates. Thus the particle may be found after the radiation to have lost on the average some fixed fraction of its energy. This is found experimentally, in some cases. For example, in pp collisions which yield a forward proton, its average momentum is about 0.60 of the incident momentum* (in individual collisions, its value fluctuates widely).

For weak coupling electrodynamics, the vector field particles are emitted independently into a Poisson distribution with mean number $\bar{n}$ emitted. The probability that none are emitted is $e^{-\bar{n}}$. The sum of the chance of emitting none, one, two, etc., (that is, the total cross section) is much like it would be without coupling to the photons. Here we know the

*Report on the Topical Conference on High Energy Collisions, CERN 68-7, February 1968, Turkot, p. 316.

total x section is constant, and so can try to interpret the energy fall-off $s^{2\alpha_o - 2}$ of the pure two-body charge exchange reaction as the factor $e^{-\bar{n}}$ the probability of / for no emission, where $\bar{n}$ is the expected mean / number of primary particles emitted. This multiplicity $\bar{n}$ must rise logarithmically with energy then as $\bar{n} = (2-2\alpha_o)\ \ell n\ s$. The particles we observe are not, of course, the primary field particles emitted, but rather the observed particles are secondary disintegration products of these unknown primaries. But if each primary produces on the average a fixed number of secondaries, we see that the expectation is that the multiplicity grows logarithmically with E.

This is necessary if our various ideas are to fit together. Because we have already suggested that the mean number of any kind of particle emitted is to vary / with x as $c\ dx/x$ for small x and, for a given x, not to vary otherwise with the energy W of the collision (so that c is a constant). The total mean number emitted, then, is $c \int dx/x$. The upper limit of x is of finite order (for the formula fails as $x \to 1$ and x cannot exceed 1) but the lower limit is of order of wee x, (i.e., order $1/W$) where the dx/x fails. Thus the mean number emitted / to the right must vary as $c(\ell n\ W + \text{const})$. (Actually we can do the integral all the way to zero, for we expect the integrand to be $c\ dP_z/\sqrt{\mu^2 + Q^2 + P_z^2}$ where μ is the mass and Q is the transverse momentum of a typical particle. Putting $P_z = xW$, this is

$$\int_0^{x_1} c\ \frac{dx}{\sqrt{x^2 + (\mu^2 + Q^2)/W^2}} = c\ \ell n \left(\frac{2W\ x_1}{(\mu^2 + Q^2)^{1/2}} \right)$$

for finite x_1. To go further, we should have to know the transverse momentum distribution.)

The one respect in which the electrodynamic analogy leads us astray is in the transverse momentum distribution of the photons. These large transverse momenta fall off slowly (like $\frac{dP_z}{E}\ \frac{d^2Q}{Q^2}$ for sudden current reversals) but in the strong collisions this is empirically not true. Some characteristics of the distribution of charge across the face of the interacting particle is involved here.

I have not yet studied the regularities involving the transverse momenta (items (4) and (5) in our Introduction) from the viewpoint being developed here. In the meantime, we can take these as empirical facts to be included in any expectations. In the same way we leave for further research strangeness and isospin character of these effects. We should notice, however, that, although we discussed a charge exchange arising from an exchange of a particle of the quantum numbers of the ρ^-, the exchange of any current of the usual octet would have analogous effects on the possible Bremsstrahlung of particles coupled to other (non-commuting) currents.

In a pure two-body exchange reaction, since the probability of not Bremsstrahlung depends on exactly what currents are reversed, the value of α_o will, from this point of view, depend on the quantum numbers of the particle exchanged (which, of course, it does). The α_o here referred to is evidently only the largest for a given set of quantum numbers exchanged.

The quantum numbers exchanged may involve not only currents of unitary symmetry, for baryon number (and possibly spin) may be exchanged, and we do not know if there are special couplings to baryon currents (or spin currents) which are also involved in determining α_o. However, I should like to hazard the guess that baryon number cannot be exchanged without

the transfer of a fundamental part of spin 1/2. Such an ideal part already has $\alpha = 1/2$ which would imply a $1/\sqrt{s}$ behavior of amplitudes before corrections to Bremsstrahlung. Therefore, if we do an experiment which freely allows the emission of wee mesons, except that the quantum numbers are controlled so that a baryon must be exchanged between right and left systems this cross section probably approaches a $1/s$ behavior, instead of the constant expected for similar experiments in which no baryons need be exchanged.

A final question is that of the distribution of correlations among the various emitted particles in the average inclusive collision. In the perturbation theory these are emitted independently and at random in a Poisson distribution so the probability that there will be k of them of momenta $x_1, x_2, \ldots x_k$ is just $(cdx_1/x_1), (cdx_2/x_2), \ldots (cdx_k/x_k)\ e^{-\bar{n}}/k!$ where $\bar{n}$ is the mean number emitted. I have not yet found a good reason whether this would be true in our non-perturbative case or not, but if one insists on comparing the experimental distributions and correlations to some theory, perhaps this is the first thing to try: that the pion distribution results from an original Poisson distribution of ρ's, each distributed for small and wee x as $c\ dP_z/E_z$ with c near 1.1 or 1.2 (c is $2\text{-}2\alpha_o$ for the ρ trajectory). In fact, if we suppose two pions for every ρ, the multiplicity of ρ's would be $c\ \ell n\ s$, and the multiplicity of the pions twice this, or about $2.3\ \ell n\ E_{LAB}(\text{GeV})$. Surprisingly, this fits observations very well (see Table 1). As an additional coincidence, this value c is what one gets from perturbation theory if one uses the coupling constant determined for ρ nucleon coupling. It must be admitted that in this paragraph we have gone much further than we should -- for our precise numerical result

depends upon a choice of which particles are fundamental and which they disintegrate into. All our other predictions were of those features which were independent of such specific choices.

The probability that the total momentum of all the emitted right moving ρ's is less than y is proportional to y^c so that (aside from diffraction dissociation) the momentum distribution of the ongoing particle, when it takes a fraction of momentum x close to 1 should vary as $(1-x)^c$ where 1-x is small.

In Section I / we discussed the longitudinal momentum distribution of partons expected in a hadron wave function, but we have not seen how this might lead us to expectations for the distributions of momenta of real hadrons in a collision, for partons are not real hadrons. Nevertheless, we shall suppose that when a hadron is disturbed via interaction, so that its distribution of partons is no longer exactly that of a single real hadron state, it must be compounded of a series of real hadron states, but the distribution of longitudinal momentum of these real hadrons is qualitatively like those of the partons which we described before (in Section I). I have no way to see why this must be true, but the features of the distributions discussed in that section seem, firstly, to rely mainly only on qualities of relativistic transformation; secondly, the principle is right in perturbation theory; and thirdly, the results of assuming this fit very well with the qualitative predictions of the Bremsstrahlung analogy. Finally, for one reason or another -- empirical or theoretical, good or bad -- I suspect that the high energy collisions to have a number of features which we summarize in the next section.

III. EXPECTED BEHAVIOR AT HIGH ENERGY COLLISIONS

In a collision of very high energy / between two hadrons each of momentum W, in the center-of-mass system, the outgoing particles of the collision should be described by two variables Q, the transverse momenta in absolute units and x the longitudinal momentum as a proportion of W (so that $P_z = xW$). We intend to describe cross sections for various processes as W increases without limit, keeping the x, Q's of the particles constant. If for large W an x is of order / 1 GeV/W (so that its momentum in the C.M. system is in the BeV range or less) we say the particle has a wee momentum. Small x simply means a value of x much below one, but higher than order / 1 GeV/W for extremely large W. Finally we should like to characterize experiments as being of two classes -- exclusive and inclusive.

An exclusive experiment is one in which it is asked that only certain particular particles of fixed Q, x, and character be found in the final state, and no others. In particular, it excludes the emission of particles with wee momenta in the limit. Examples are: two-body reactions $AB \to CD$; an experiment which uses missing mass to try to select events which are virtually two-body reactions, etc. For such collisions, the cross sections should fall off inversely as a power of / $s = 2W^2$, the power being $2\alpha(t) - 2$ where $\alpha(t)$ is the α appropriate to the highest Regge trajectory capable of carrying the necessary exchange of quantum numbers between the right moving and left moving system, and where t is the negative square of the transverse momentum which must be exchanged. (/ With these x, Q these variables fixed, the longitudinal momentum transfer and the energy transfer go / inversely as W to zero as $W \to \infty$.) In the special case that no quantum numbers need be exchanged, the cross section is constant (empirically, at least, if $t = 0$); it does not fall as $W \to \infty$. This phenomena is called diffraction dissociation,

and is sometimes expressed as the exchange of the "Pomeranchuk trajectory".

An inclusive experiment is one in which certain particles are looked for at given Q, x, but one also permits any number of additional particles. More precisely, it does not in any way exclude the emission of arbitrary numbers and kinds of particles with wee momenta. Examples are the total inelastic cross section, the mean number of Λ mesons emitted with momentum x in range dx, the probability that no single particle is moving right nor left in the (center-of-mass system) with more than 1/2 the original momentum W, etc. Such cross sections should approach constant finite values as $W \to \infty$.

The mean number of mesons of a given kind formed with small x in a high-energy inclusive experiment should vary as dx/x. This should extrapolate right through the wee x region in the form dP_z/E where E is the energy $\sqrt{\mu^2 + Q^2 + P_z^2}$ of the meson of momentum P_z, at fixed Q. (This suggests that more appropriate variables for the small x region, would be w, Q where w is the z-component rapidity $w = \tanh^{-1}(P_z/E)$. The distribution should be uniform in dw (for each Q) and ultimately independent of W.) As a consequence, the multiplicity of a given kind of hadron should rise logarithmically with W.

It is this dx/x distribution with its logarithmically divergent character for small x which makes it possible that the probability of finding any specific set of particles with given x, Q values (except the elastic or diffraction dissociation ones) falls with energy as a power, and yet the total cross section can be constant.

In an inclusive experiment of A colliding from the right, with B from the left, the probability that some particle C comes out moving to the right with an x close to unity should vary as $(1-x)^{2-2\alpha(t)}$,

as long as $(1-x)^{1/2}$ is not wee. Here $\alpha(t)$ is the value appropriate to the trajectory of highest α (excluding the Pomeranchon) which could, upon emission, carry away the quantum numbers and transverse momentum needed to turn A to C.

In a special kind of partially exclusive process in which a baryon must be exchanged to get the reactants of finite x, but no wee baryons appears among the particles of wee momentum, then I believe the cross section will vary as $1/W$ but this is not on as firm a basis as the other suggestions.

TABLE 1

Multiplicity of Pions in High Energy Collisions

E_{LAB} (BeV)	Mult.	2.3 ℓn E_{LAB}
30	7	7.8
470	13 ± 1	14.1
1500	18 ± 2	16.8
12300	24 ± 4	21.6

PARTONS*

Richard P. Feynman

Department of Physics
California Institute of Technology
Pasadena, California 91109

Many of us are working on the same problem, which is to understand the behavior of hadrons, the strongly interacting particles. Among the various attempts to understand it in the past decade, there is a recent one that goes under the name of "partons". It sounds very mysterious, but I will try to show you that you knew it all the time, you just called it something else. Work on partons has been done by myself, Paschos, Drell, Bjorken, and others.

Almost everybody supposes that the strongly interacting particles obey a number of principles like the quantum-mechanical superposition of amplitudes, relativistic invariance, unitarity, and analyticity in some form or other. The problem is to

*An invited talk presented at the Symposium "The Past Decade in Particle Theory" at the University of Texas in Austin on April 14-17, 1970.

make a theory which is consistent with all of these principles at the same time. It would be very convenient if we could write our equations in a mathematical form such that we do not have to keep checking them for unitarity, relativistic invariance, quantum mechanics, or something else. We would then not get new equations if we add some of these features, for the special mathematical form would automatically have all the equations in it. The thing to do is to find a model which satisfies all of these conditions simultaneously. The only model that we have is quantum field theory. The idea is to take quantum field theory, despite its divergences and difficulties, and try to see what kinds of mathematical formulations it would have in it. We will start with something which satisfies all the conditions and later on perhaps we can modify some parts.

What is implied by quantum field theory? Of course you would have to make for field theory a specific model, yet in every quantum field theory there are fundamental fields and operators which create and annihilate some kind of particles that underlie the theory. We need a name for the objects which the fundamental operators a* and a create and annihilate, objects that are not the final states of the system,

such as a complete proton, which may have parts inside. We need a name for what one would call the "elementary particles" of the theory. We do not know if there are any such particles in the end, but we will start by supposing that there are because otherwise we would have no field theory at all. I will call these things "partons". In the case of electrodynamics, for example, the partons are the ideal, what one sometimes calls "bare", electron and the ideal proton. The creation and annihilation operators create and annihilate these. But, a physical electron is an ideal electron with some photon field around it, and is therefore a combination of various partons.

What does a field theory look like? In quantum field theory, instead of representing things by scattering amplitudes and Regge poles, we use a wave function, and it gives the amplitude for finding a parton moving this way or that, or two partons going in a certain direction, etc. So we shall see if we can get anywhere by discussing what the character of the wave functions of the protons, the neutrons, the pions might be, and shall try to represent the experimental results in terms of properties of the wave functions. Ultimately, maybe we will see some special properties of the

wave functions and perhaps get some idea of what the partons are, whether they have spin one-half, or they are quarks, if in fact they are anything. Also, maybe by using the wave function, one might be led, psychologically, to suggest a certain principle or find a general proposition that can be deduced outside the realm of abstract field theory, without talking about the objects inside. For example, one might be able to deduce it by looking at the commutation law of currents, or something like that, which does not say anything about the underlying machinery. In such a case one would establish something that is more satisfactory than if it were model dependent. But, it does not bother me too much if, at first, it is model dependent.

The characteristic of a wave function is that if you have an amplitude for finding a lot of particles or pieces around, the amplitude is given by

$$\text{Amp} = \frac{N}{E_o - \sum_i E_i} ,$$

where E_o is the state energy, say of a proton, and E_i are the constituent energies of the partons inside. The numerator is some kind of matrix element. Now an interesting

thing about a wave function is that it does depend on the coordinate system in which one looks at it. It is an unrelativistic idea, really, because it cuts space-time at a given moment and asks, "What partons do you see?" If, say from a Lorentz transformation, it cuts space-time at a different angle, one would see other combinations. If one has never tried these things, the fact that the wave function is not a relativistic invariant may be bothersome for a moment. But you know that the sum of the momenta of the partons inside the wave function is always the same as the total momentum of the state, and that the momentum is conserved, but that this is not so for the energy. It is the lack of energy conservation which gives the strength of a certain amplitude, so it is not relativistically invariant.

It is not easy to transform a wave function, because one has to know the Hamiltonian. Therefore, there may be one system or another in which it is most convenient to look at the wave function and the usual suggestion is to look at the wave function in the rest system of the entire state. I prefer, instead, to look at the wave function in a system in which the particles are moving very fast, all extremely relativistic, in fact, in the limit. Thus, I would like to look

at the wave function of a proton when it is moving extremely fast in the z-direction.

Now if such an object moves very fast, then in order to understand what happens it is convenient to rewrite the same expression as

$$\text{Amp} = \frac{N}{(E-P_{oz}) - \sum(E_i-P_{iz})} ,$$

since $\sum P_{iz} = P_{oz}$ where P_{oz} is the center-of-mass momentum in the z direction. When the thing is moving at extreme velocities in the z-direction, let me suppose that

$$P_{iz} = x_i W .$$

For example, W may be the energy in the center-of-mass of the original state, in which case x_o would be one. Or W could just be a scale, and we could leave x_o open. Then the energy of a particle (denoting the transverse component of a particular particle by Q_i) is

$$E_i = \sqrt{P_{iz}^2 + Q_i^2 + M_i^2} = \sqrt{x_i^2 W^2 + Q_i^2 + M_i^2} .$$

It has appeared experimentally in collision after collision of hadrons with each other, that the side-ways momenta are

all limited -- they seem to be of the order of 340 MeV. As the energy of the collision rises to the higher and higher GeV's, there is no increase in the transverse momenta of the outgoing particles in the collision and they are finite. I am therefore going to guess that inside the wave function, it is also true of the partons. Therefore, as W goes to infinity, Q stays the same. Let us try it. Using $E - P_z \simeq \frac{M^2 + Q^2}{2Wx}$, we get

$$\text{Amp} \simeq \frac{2WN}{\frac{M_o^2 - Q_o^2}{x_o} - \sum \frac{M_i^2 - Q_i^2}{x_i}}$$

with $\sum x_i = x_o$. Now it turns out that as you take the limit for certain matrix elements, 2WN falls and the amplitude of those states is very low. But for various states 2WN approaches a constant, at most, and therefore, interestingly enough, the wave function has a definite limit as long as one stays away from x's near zero. That is, if you go to infinity with W, everything will be all right if the x's remain finite. There is a small technical point to straighten out as to what happens as x approaches zero.

However, we have already discovered something interest-

ing. It is that if we represent transverse momenta in absolute units, and longitudinal momenta in the scale of the energy of the collision, the wave functions that come in are invariant as W goes to infinity. This leads to a suggestion about the energy behavior of highly inelastic collisions in general, namely, that expressed in terms of the variables Q's and x's, the laboratory experiments might approach invariance at extremely high energies.

A little more can be said about the wave function. I want to represent this fast moving photon by a number of particles moving at different speeds, sharing the total x_o, the total longitudinal momentum in scale, and let us say $x_o = 1$. All these particles share the total momentum of the system. How about some of them becoming negative? No, they cannot! If x were negative, when I take the square root of E, which has to be positive, it would be -xW, and so for particles going backwards, $E-P_z$ is very large, not very small. It is the $E-P_z$ in the denominator that makes the amplitude of that term negligible. So the wave function of a fast moving hadron consists of a lot of little particles moving forward, sharing the longitudinal momentum and having a finite, confusing, transverse momentum.

I have tried to use this concept of the wave function to understand a number of phenomena. I would ultimately like to be able to understand, as much as possible, all the characteristics of the things that we observe from such a point of view, such as where the Regge expressions come from, etc. So far, I have had but a limited success. However, even the crudest situation leads you to certain ideas. Consider the elastic form factor of a proton. That means that we hit the bunch of forward-moving partons transversely with a tremendous momentum by a photon, and ask with what amplitude do we get a proton back again. It would be represented by a number of partons moving in the direction of the final proton. Now the electromagnetic field, I am going to suppose, interacts with only one parton at a time, because most field theories have a propagator for the parton obtained by the substitution $p \to p - eA$. So what we are really asking is this: Suppose we have a number of these partons and we kick one of them sideways, what is the probability that the rest of them look like they were moving in the final direction? Now, for each of these initial partons there is a certain amplitude that it was not moving exactly in the forward direction, that it had a little transverse momentum, and there is some ampli-

tude for the final proton state that its corresponding partons look the same way, so there is an overlap integral between the two. But the higher the momentum transfer is, the harder it is to get the overlap, and, therefore, the form factor should fall off with Q depending on how likely it is that the non-kicked partons can be seen to be going in the final direction. I would guess that there is a kind of universal function for such problems of high momentum transfer. There is, of course, a matrix element for the photon on one parton that happens not to be Q dependent, but the main Q-dependence, the rapid fall-off, must come from trying to get the fast moving particle going in one direction to look like a very quiet simple thing, a proton.

Since we know experimentally that this form factor falls off toward zero as Q goes toward infinity, perhaps as Q^{-3}, we learn immediately that there is no amplitude to find only one proton-like parton inside of a proton. Because if there were, there would be a certain finite amplitude, A let us say, to find a proton consisting of a proton-like parton all alone. Then when you kick it, it would be a deflected proton-like parton, and the amplitude of that for the proton is also A, so A^2 would be the ultimate limit of the form factor

as Q goes to infinity. So we have to look for more complicated wave functions. This amplitude could be zero for several reasons, one of which would be that there is no parton in the fundamental field theory which has the same quantum numbers of charge and spin as the proton. Or it may be for reasons somewhat like the fact that the electron is never found with absolutely no photons around it. There is always a field around it. But I do not think that this case is realistic here because in QED it depends on the zero mass of the photon.

The next application which I want to describe is called "inelastic electron-proton scattering." In this experiment, we take a proton and hit it with an electron, and there is an exchange of a photon. There is also a small correction for the exchange of two photons, but that is removed by the theoretical people who analyze these experiments in order to have an interaction of first order in e^2 between the electron and proton. We can control the variable $q^2 = -Q^2$ and the variable $P_o \cdot q = M\nu$. The momentum transfer, q, is space-like, and ν is the energy loss of the electron in the proton rest system. M is the mass of the proton. In principle, we can measure the probability amplitude to have different final

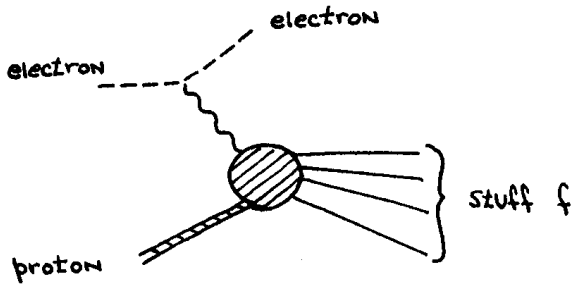


Figure 1

Inelastic Electron-Proton Scattering

states as a function of these two variables. In fact, however, at the present time, without looking at the final states of the proton we observe the total cross section by looking only at the final electron. So we have to discuss the total probability that a proton, being hit by a photon, absorbs a certain energy and a certain momentum.

First, we will describe the experiment. It turns out that the differential cross-section for inelastic electron-proton scattering can be expressed in terms of the angles in the laboratory and two functions, W_2 and W_1, which are functions of Q^2 and ν.

$$\frac{d^2\sigma}{dE'd\Omega} = \frac{e^4}{4E'^2\sin^4\theta/2}\left[W_2\cos^2\frac{\theta}{2} + 2W_1\sin^2\frac{\theta}{2}\right]$$

Here E' is the final electron's energy. The amplitude for the scattering is

$$\text{Amp} = 4\pi\alpha\frac{(\bar{u}_2\gamma_\mu u_1)}{q^2}(f|j_\mu(q)|P) \ .$$

With the known factors all factored out, we can express everything in terms of the quantity

$$K_{\mu\nu} = \sum_f (P|j_\nu(-q)|f)(f|j_\mu(q)|P), \quad \nu > o \ .$$

From relativistic invariance, we have

$$K_{\mu\nu}(q) = P_\mu P_\nu W_2(q^2, \nu) - \delta_{\mu\nu}W_1(q^2, \nu),$$

where both W_2 and W_1 are positive, and

$$\left(1 + \frac{\nu^2}{Q^2}\right)W_2 \geq W_1 \ .$$

Furthermore, since they are both positive, there is another relationship.

Because the angles used were small, the results are sensitive only to W_2, and we will plot the complete result as though we were plotting W_2, although this is a little bit erroneous. Let us look at the complete scattering, first at

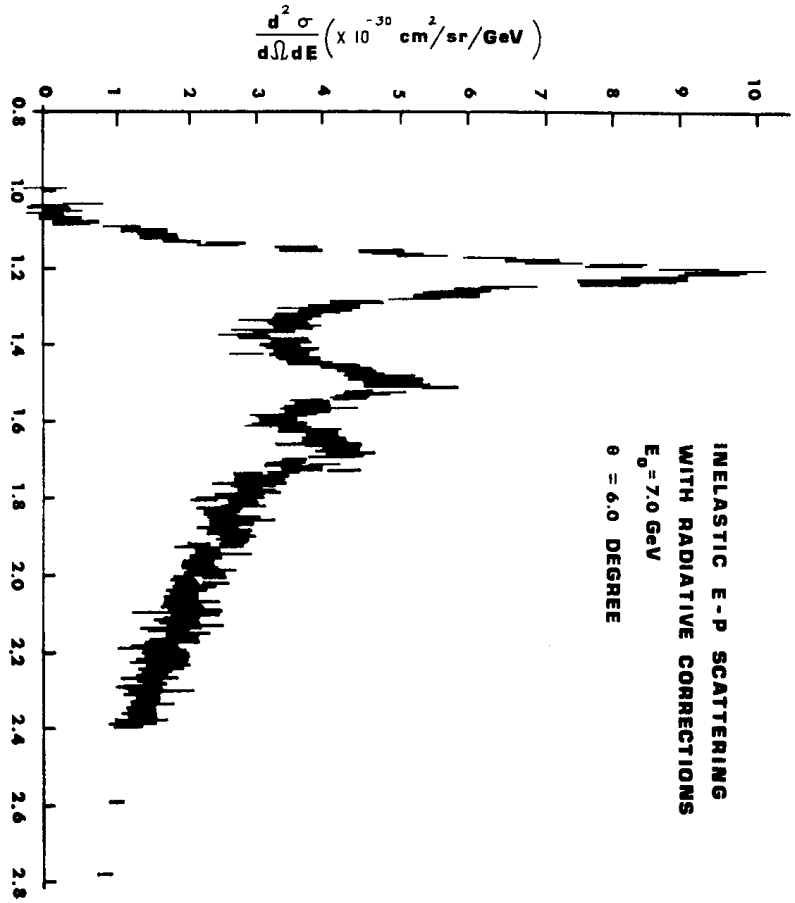


Figure 2. Inelastic e-p scattering cross section versus missing mass.

low values of momentum transfer and energy.

The missing mass is simply another way of expressing the variable ν,

$$M_f^2 = (p+q)^2$$
$$= M^2 + 2M\nu - Q^2 .$$

But it measures the total four-momentum squared of all the products added together, such that if the object which is made in the collision is not a lot of pieces, but happens by accident to be one little ball of slightly excited "goop", there will be a resonance if that "goop" has a definite mass. In Figure 2 there are the famous resonances like the 1238 and there is a 1535 and something higher. The resonance at 1410 does not appear but it is very interesting that these resonances appear so beautifully.

Now if we hit the proton a little harder, it will be a little bit harder to get all those pieces to come back together again to form a resonance. We should still see the resonances but without so much glory.

In Figure 3, as expected, the resonances are much weaker and there is a rather large smear in the back which is sometimes called the "deep inelastic scattering." It is obvious

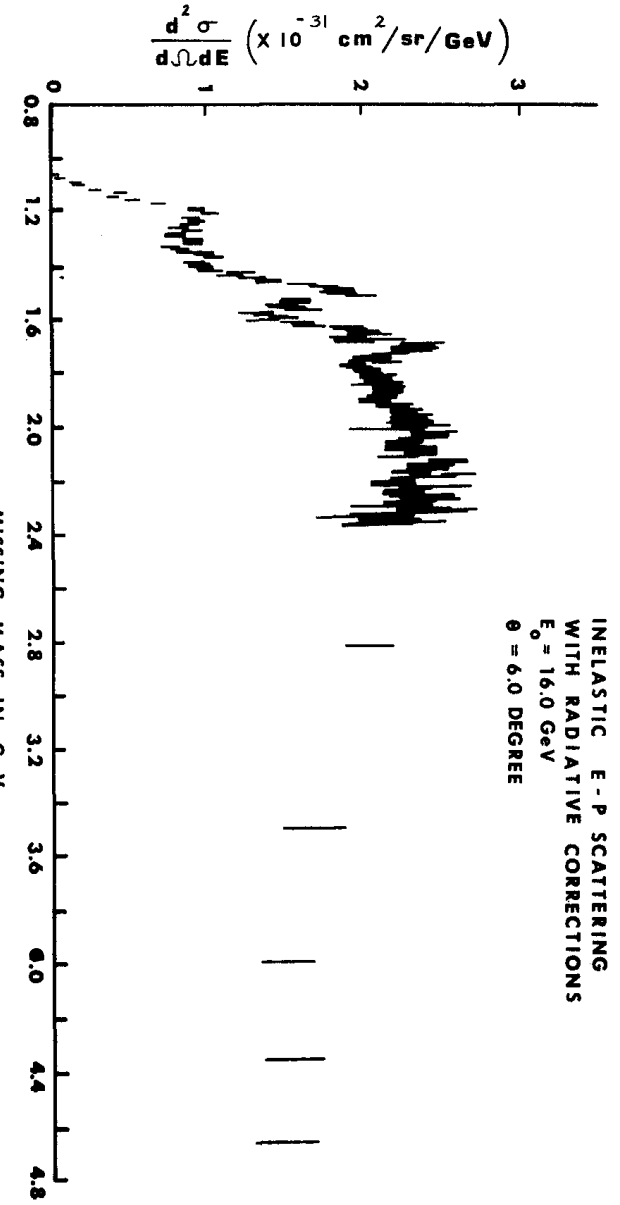


Figure 3. Inelastic e-p scattering cross section versus missing mass.

what that is: that is when you hit the proton and the pieces do not stay together, they just fly into a lot of pieces, but the chance that they are resonances becomes less and less.

Incidentally, it is very interesting that the way these resonances fall as Q^2 increases is almost exactly the same as the way the elastic scattering form factor goes out. There is a certain amplitude, that is not on these curves, which should be represented by a δ-function at exactly the mass squared of the proton, but that has been taken off so as to make this curve look good, otherwise this curve's elastic term is very big, and the size of it is the elastic scattering, which is the ordinary form factor. The funny thing is that all these fall off in proportion to the elastic form factor. In other words, the probability that the thing holds together to form a proton and the probability that it holds together, as you give it a higher momentum transfer, to form some excited state is a ratio roughly independent of Q^2. As Q^2 gets harder and harder, it is difficult to line things up, but if they are more or less lined up, they might as well be an N* as a proton.

Now let us consider deep inelastic scattering. Plotted together on Figure 4 are the data for several different

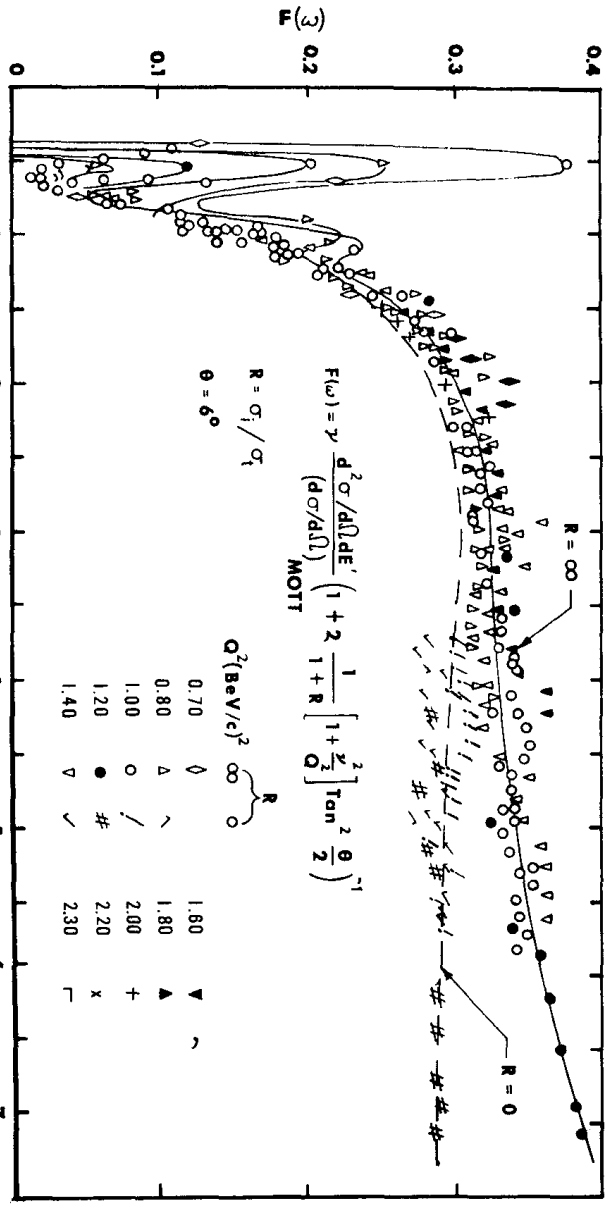


Figure 4. Scaling of the structure function.

momenta and energies as a function of ν/Q^2. What is plotted vertically is not W_2, but νW_2, and strictly speaking it is a little bit mixed up with W_1. The way that you plot it depends upon how much you assume W_1 is. If $W_1 = 0$, you get the upper curve, in which the dots do not fit on top of each other, while if W_1 is at its maximum value, then you get the curve marked R = 0 which is more or less a constant. What is interesting is that νW_2 plotted against ν/Q^2 seems to be approaching a universal curve so we should like to explain why that is a natural thing to expect.

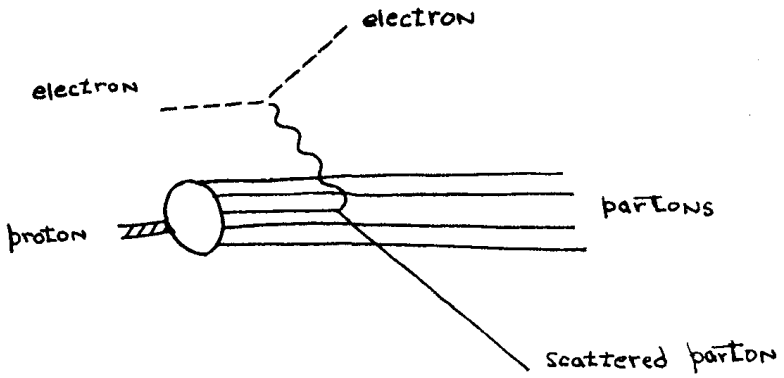


Figure 5

Parton Scattering Out of Incoming Proton

As before, in Figure 5 we represent the proton coming in

by a lot of partons. Now one of the partons is hit by an enormous momentum. The proton is moving along fast, but the sideways momentum from the photon is also very great and it knocks the parton off in a crazy direction, and leaves the other partons undisturbed. It might appear that the partons are all interacting with each other but they are not. When a thing is moving slowly, there is a lot of interaction between the parts. When the thing is moving very fast, the Lorentz time shift makes all the interaction go like a slow clock, and therefore the momentum increases so fast that it does not have much time to find out what happened. That is to say, we do not have to worry about the interactions ahead of time, nor do we have to worry about them very much afterwards, because when you sum over all the final states of the system there is a principle of completeness which says that the sum over all the final states is the same, whether they interact or not. (This is not absolutely true because there may be a shift in the energy. And when we insist experimentally that there be a given loss ν, it may not be exactly the same as the loss of a single parton because its energy may be somewhat changed by the interaction. At any rate, we can deal with these objects as free particles.)

The probability for this process involving the parton is

$$\text{Prob} = \int \frac{f(x)\,dx}{2\varepsilon}\, 2p_\mu 2p_\nu\, \delta((p+q)^2 - ?, \text{ say } p^2)$$

where f(x) is the momentum distribution of the partons inside, each weighted by the square of the charge that it carries (in units of the electron charge), and x is the fraction of the z-momentum of the incoming proton that the parton has. The scattering probability for a photon on a parton is the "square" of the current, $2p_\mu 2p_\nu$, if the partons had spin zero; and for the other cases, you get more complicated things. But if you just look for W_2, you only need this factor. The delta function is over the square of the four-momentum before and afterwards. Let us suppose that afterwards it is equal to approximately p^2. So expanding,

$$\text{Prob} = \int \frac{2f(x)\,dx}{Ex}\, P_\mu P_\nu x^2 \delta(q^2 + 2xP\cdot q + \text{transverse momenta, etc.}),$$

where we are taking $p_\mu = xP_\nu$ for the whole four-vector, so our notation for the p's is a little inconsistent. Inside the delta function we have a term $q^2 + 2xP \cdot q$, where P is in the z direction, plus some transverse momenta squared. There is an uncertainty as to the exact form for these extra

terms and there may be little energies of interaction of the final object with the original object, so there are finite errors here. However, we are going to take the limit as Q^2 goes to infinity and as $2M\nu$ goes to infinity. As the energies and Q^2's get higher, we get deeper inelastic scattering, and the contribution of these finite terms to the delta function becomes less, and we have

$$\text{Prob} = \frac{2P_\mu P_\nu}{E} \int x\delta(-Q^2 + 2M\nu x)\ f(x)\ dx$$

$$= \frac{P_\mu P_\nu}{EM} \cdot \frac{1}{\nu}\ x\ f(x)\ ,$$

thereby getting,

$$\nu W_2 = xf(x),\ x = \frac{Q^2}{2M\nu}\ ,$$

which shows that ultimately, as Q^2 and $2M\nu$ go to infinity, νW_2 is an invariant function independent of any variable except $Q^2/2M\nu$.

I can discuss the degree to which this should be correct. It turns out that it should be much more correct than we would guess. The errors are the order of something over ν, and then are only important if $xf(x)$ varies rapidly near the origin. But $xf(x)$ is constant at the origin, so if you misplace the value of x that you are looking at, it does not

make much difference to νW_2, so I can say that this is a pretty good approximation. Actually, the energy that the things are measured at is not so terribly high, so it is rather nice that the universality of νW_2 is already showing up.

Now that I have explained the properties of the experiment in terms of the parton model, I would like to discuss the results of the experiment to find out what we know about the wave function for the proton. In the first place, we find that νW_2 is universal. What can we conclude from that? We can conclude that the charged partons are either spin 0 or spin 1/2. The coupling to a spin 1 particle increases with energy so rapidly that if there were spin 1 partons in there, the universality would be lost. We can even go a little bit further. If the spin were 0 for the partons, W_1 is zero. If the spin is 1/2, then W_1 reaches its maximum as expressed by the inequality we gave earlier. The question is, "What is W_1?" Now, here we go into a circle depending on how energetic a theorist we are. If we are sure that W_2 is universal, then we are also sure that the case corresponds to spin 1/2, because that is the case in which the curve is more universal when you plot it. But if, for some other reason, the curve is not universal, you would be hard pressed to argue which

one is the right curve. There are going to be some more experiments at other angles and higher energies, that will make this a little bit clearer. At the present time, I would say that the experiment does favor the spin 1/2 and that, in principle, it could decide the question as to whether the charged objects which contribute to the current, the partons, are spin 1/2 or spin 0. I will conclude myself from what I have seen, that they are spin 1/2, and not spin 0.

Now it is interesting to turn these statements into a universal language that does not depend on our model. This is always a good thing to try to do . The model is just a scaffolding to discover something, but it is not the house. The first question is: What does it mean that νW_2 is universal? The $K_{\mu\nu}$ which we are measuring can be expressed directly in terms of things that do not involve the field theory

$$\int K_{\mu\nu}(q)\ \exp(iq\cdot x)\, d^4q = \langle P|[J_\nu(x),\ J_\nu(0)]|P\rangle \quad .$$

Therefore, in measuring $K_{\mu\nu}$ we are measuring the property of the commutator of two currents' expectation for a proton (minus expectation for a vacuum). Everybody knows that when two points in space-time are separated space-like the commutator shall be zero, and we also know that when the two

points are time-like related, then the commutator is not zero, except in a trivial case. So what type of singularity is it as we sweep across from one region to another? What happens as we cross the light cone? According to the simplest theories, there is a delta function type singularity on the light cone for the above expression. It turns out that for a single particle in perturbation theory, particles of spin 0 and 1/2 have an ordinary delta function across the light cone, whereas particles of spin 1 have a gradient of a delta function. The proposition that νW_2 is a universal function is equivalent to the statement that the singularity on the light cone for the expectation value of the commutator is a simple delta function across the light cone and involves no gradient of the delta function. That is, if $\langle P|[J_\mu(x), J_\nu(0)]|P\rangle$ varies as

$$P_\mu P_\nu g(t)\frac{1}{r}\,[\delta(r-t) - \delta(r+t)]$$

on the light cone, then

$$\nu W_2 \to G(x) = \int g(t)\, e^{-iMxt}\, dt \quad .$$

We can therefore forget about partons if we want to, and state everything entirely in terms of commutators.

What does it mean that the parton has spin 1/2? Gell-

Mann, in writing his current commutation relationships assumed a kind of minimal electromagnetic coupling which is equivalent to the statement that the singularity across the light cone is the least possible. He also proposed that the commutation rules of the currents involve axial currents and that the axial currents and the regular currents commute, just as they would if they were coupled to a spin 1/2 object. So I think that with a little bit of luck and a little bit of fiddling around, I could conclude that the idea that the W_1 and W_2 are related in this limiting case for the spin 1/2, is related to the statement that the currents are part of a "thing" x "thing"; it does have to be SU(3), but could also be say SU(3) x SU(3). That means a coupling with γ_μ exists and that the opposite parity coupling, or the axial current, exists and has a coupling very similar to the normal coupling.

Let us go on to some other properties of this distribution F(x). We have seen it plotted in the reciprocal variable, but I would like to plot x F(x) in the regular variable $x = \frac{Q^2}{2M\nu}$.

Figure 6, when suitably normalized, yields

$$\int_x f(x)dx = 0.16 \text{ or } f(x) \sim \frac{dx}{x}\, 0.16 .$$

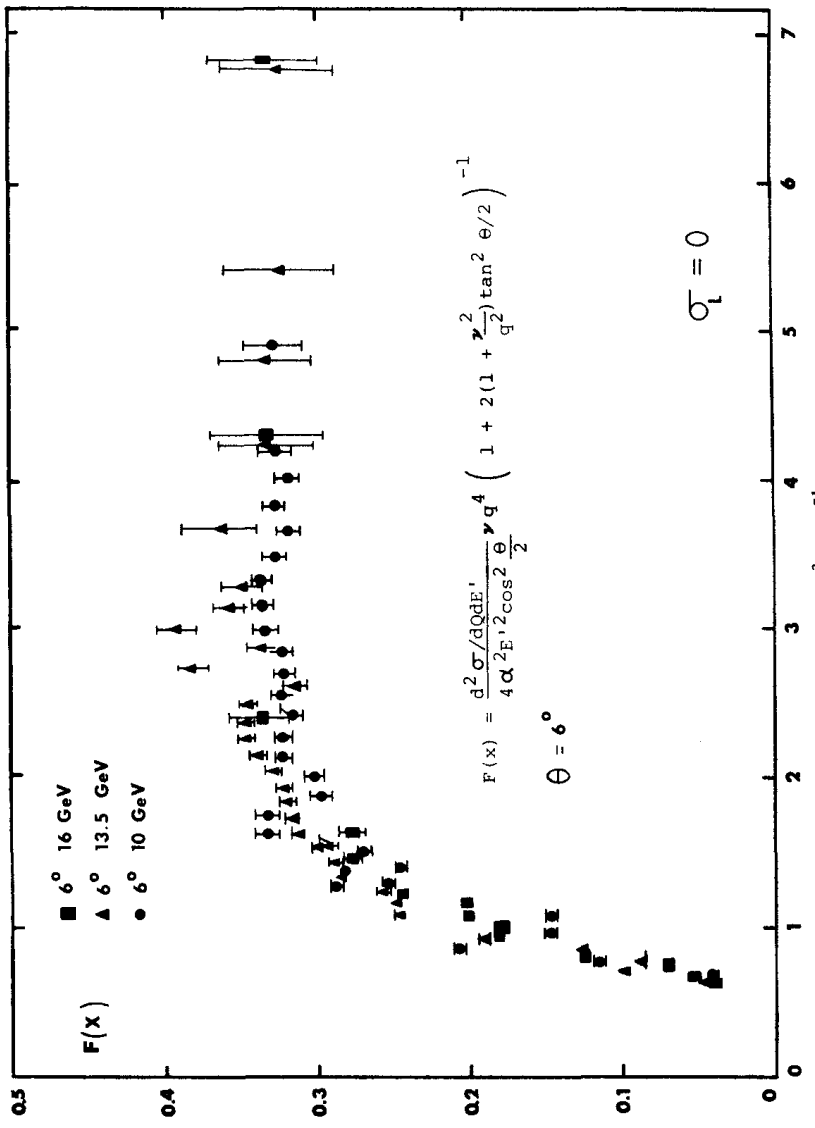


Figure 6. Structure function versus the scaling variable.

The first thing it shows is that there is an infinite number of partons because the mean number of charged partons goes as dx/x, and so the total number of them in an infinitely fast-moving proton would be logarithmically infinite at the low end. That is perfectly all right. However, it is very interesting to take a simple perturbation theory and make a model, say of proton going along that disintegrates into a parton of spin 1/2 and another parton of spin 0. If we then ask what is the distribution of this parton of spin 0, we will find that for spin 0, the distribution must always be xdx; for spin 1/2, dx; and for spin 1, $\frac{dx}{x}$. This is another representation of the angular momentum group and the angular momentum properties. For a single particle with nothing else around, perturbation theory gives the distribution $xdx/x^{2\alpha}$, for small x where α is the particle's spin, and a higher angular momentum would be more singular. Now, here we have a little peculiarity because dx/x is what you get from a vector particle, and dx/x is what we found; but a vector particle is the very one that we could not allow to couple. However, our arguments were based on the erroneous idea that a single particle could be in the small x region and that it is all right to do perturbation theory. There

is an infinite number of particles for small x, and one cannot proceed by assuming a single particle and perturbation theory for the strong interactions.

However, very much like the particles in electrodynamics, when an electron is moving fast the field is condensed into a narrow splat by the Lorentz transformation, and if we analyze that in momentum space we get a uniform distribution in momentum space of the energy; the energy of each photon times the number of photons goes as dk/ω. So the reason for the dx/x is this dk/ω -- the field is squashed. (Incidentally, $\frac{dx}{x}$ looks like dk/ω, and if the photon had a finite mass as k goes to 0, it would have a finite limit. So does dx/x. It doesn't look like a finite limit, because the x is on a momentum scale which has been increased by the factor W to an enormous amount.)

Now the next thing of interest is the integral, $\int xf(x)dx = 0.16$. What is the normalization and what does 0.16 mean? I can express it this way: Suppose for a moment that all the partons were charged. If they all had unit charge, the square of their charges is also 1.0, and then this integral would be simply the sum of the momenta of all the particles. It is so normalized that it would be the total

momentum of the original system and would be 1.0. It is remarkable that this comes out to be only 0.16, which is very small. For example, it seems to say that the number of neutrals in the system is five times the number with charges which is a little hard to accept. It could mean that the partons have a non-integral charge. There has been a lot of talk about quarks, and the proton may be made of two quarks of charge 2/3 and one of charge 1/3, in which case that number would come out 0.33, if all there was in the proton were three quarks. But, we just discovered that there must be an infinite number of partons in there. As Paschos has said, "Yes! Those are millions of pairs of quarks and antiquarks." So what we should do is take the statistical average of the 2/3 squared, the 2/3 squared, and the 1/3 squared, every kind of quark being equally likely. That comes out 0.22. But Paschos has disregarded the fact that the quarks will probably interact if there is going to be an infinite number of them. They are not going to be free, and there must be something which carries momentum back and forth between the quarks through interactions. This would be equivalent, I believe, to having a neutral quark around. So the interactors would lower it still further, and maybe it is quarks that are in-

teracting. On the other hand, as long as the interactions are lowering it you can say that there is ordinary charge and an awful lot of interaction, such that there are indeed five neutrals for a single charged one. However, it is very interesting that these points of view are leading us into "semi-difficulties", or if you like, indications that the quarks and the partons are related.

One might ask: What is meant by the statement that the charge is an integer? It turns out that it can only be expressed as a property of four currents. Two current operators can make a commutator, but four currents can make commutators of commutators and things like that. The statement that the charge must be an integer on all the partons is a statement about a commutator with four currents in it. Experimentally, one can get a four current commutator by measuring the inelastic Compton effect from a proton at very high energy if the parton view is right. This is because in the Compton effect the amplitude goes as the charge squared; and when one works this model, one will be weighing charge to the fourth power. So by having another number on the charge, we can compare the charges. For example, if all are unit charges, then the same integral, where the weight is the

charge to the fourth power, would give the same number. But, if the charges were less than a unit, then the integral for the Compton effect would give a smaller number. Of course, it is technically not quite feasible to do the Compton scattering. So I know of no direct way of testing it.

An attempt has been made to apply these ideas to hadronic scatterings, and I have made a number of speculations in print. For example, I can talk about the distribution of partons inside the hadrons and also about the distribution of hadrons which come out of a collision, and they are likely to be related but it is not obvious exactly how. One interesting feature, though, is this: Suppose for a moment that a proton moving to the right has a wave function to be partons and suppose that the exact wave function has virtually zero, or very low, amplitude to find any partons of very small x. That is, suppose all the partons are moving to the right and there is none standing still. Suppose also that another guy is moving the other way with exactly the same property. His partons are all moving to the left and nobody is nearly standing still. I am now talking about the wave function, not for the infinitely high speed, but for a very high speed where the limit has not quite been taken. Then each wave function is a

solution for the whole Hamiltonian. The right moving solution involves only the right moving piece of the Hamiltonian, but it is a solution of the complete Hamiltonian. However, all the states in the right moving object are different from those in the left moving object, that is all the parton excitations are different. So if you will multiply these two wave functions together, it will still be a solution of the Hamiltonian equation. In short, there is no interaction!

There will be an interaction if one particle in the right moving proton can be thought of as being in the left moving proton. When we say that the distribution is dx/x, where xW is equal to the momentum of the parton in the z direction and W is the total momentum, it is the same as saying dp_z/p_z as long as p_x is large. But if x is of the order 1/W, then p_z is no longer large, and I could not have taken out the square root. It is true that it is dx/x as long as x is bigger than a small quantity 1/W; however, if x is of the order of 1/W, then the argument that the partons cannot go backwards and that for dx/x distribution begin to fail. So in this coordinate system, there is a certain amplitude for finding partons at rest and that amplitude is approximately the integral of dx/x from 0 to 1/W. The integral ex-

actly is not of dx/x, but rather is of something like dk/ω. It is a cutoff, it is a function that looks like dx/x for large x, and then it cuts off somewhere of the order 1/W. So the area under this cutoff is approximately the cutoff height W times the width 1/W. In other words, the integral below 1/W is finite and independent of W. So a dx/x distribution of partons permits an interaction which leads to a cross section independent of the energy in the higher hadronic collisions. This may not be exactly right, there may be some errors, maybe a slight variation logarithmically or something.

If I suppose that the dx/x distribution extends also to the real particles which are produced in the collision, then the probability to have an exchange collision is the probability that there are no emitted particles. If a proton comes in, say, and exchanges a charge to become a neutron, it is the probability that the neutron is quiet. If I suppose that the probability that there is an emitted particle is also distributed as dx/x up to order 1/E, then

$$\text{Prob} = \int_{1/E}^{1} \frac{dx}{x} = c\ \ell n E, \quad c = \text{constant}\ ,$$

and the mean number of emitted particles is $\bar{n} \sim c\ \ell n E$, that is the mean number of pions and junk to be expected in an in-

elastic collision. But for charge exchange where no pions come out, the probability that none of them does come out is $e^{-\bar{n}} \sim E^{-c}$. This is the reason why when you have nice exchange reactions, where you ask for a specific exchange, things fall off with a power of the energy. They fall because the probability that they do not shatter the product into a lot of tiny partons, each with a mean number logarithmic will fall. I also expect the multiplicity in high energy collisions to go logarithmically, which apparently it does, as the number slowly increases in cosmic ray experiments.

In closing, the theoretical concept that I would like to emphasize is not so much the "partons", but that it is the wave functions of fast moving particles which might be useful for analyzing strong interactions, especially at high energies. By the way, there is no approximation made here and that is the complete wave function, so you could deal with any energy. For a different reason I would like, on the other hand, to emphasize for further study the interesting results of the experiments on the deep inelastic cross-sections. There are quite a few numbers associated with them, as well as simple relations, from which we ought to be able

to learn something more, whether we do it by partons or by any other method.

DISCUSSION OF PROFESSOR FEYNMAN'S TALK

JULIAN ROSENMAN (University of Texas):

Can individual, free partons ever be seen?

FEYNMAN: No, I don't believe so and, at any rate, I'm supposing not. Just like the proton cannot have any amplitude to be a pure parton, the pure parton would itself disintegrate into real particles. This would be true, for example, for the parton kicked out by the photon in the deep inelastic scattering. There also might be some interaction with the remaining partons on the way out but I don't think that's as important as its disintegration. It could be that the parton's energy is higher than any of the fractions into which it can go.

ROSENMAN: Doesn't this mean that partons must have an integral charge?

FEYNMAN: Yes. So one worries if partons are quark-antiquark pairs, as Paschos suggests. What can we do, because certainly such a parton cannot be coming out? One of the ways that we can fix it is to assume that the interaction between them, say, a harmonic force to produce the linear masses, so when you pull a quark away, it snaps back by the harmonic force. Likewise, the kicked parton would be pulled

back and all mixed up with the other ones, and the net result would still be pions and protons, not quarks. But it seemed to me at first that this large force would obviate the argument about the small energies. It doesn't. What matters is not the strength of the forces, but rather the spacing between the energy levels. The large force only means it takes longer to come around, and I've already summed over all the states and that's independent of things because of completeness. In fact, using the harmonic oscillator potential, I analyzed this and found that this theory should be right as $f(\frac{Q^2 \pm (300)^2}{2M\nu})$ for a spacing between the levels of 300 MeV, as for the proton excitation spectrum.

EDWARD M. MACKHOUSE (University of Texas):

If the partons are attracted with a quadratic force, do you postulate some other force to keep the self-energies from blowing up, to avoid the divergences?

FEYNMAN: What I postulate is that I can go as long as possible so long as I evade the question! It's true that any field theory has its divergences and it is also true that the proposition that the transverse momenta are limited is not quite right for a field theory. If you do it right field-theoretically, when you integrate over the transverse momen-

ta, you get divergences, and so I think that there is something limiting the transverse momenta in a way that I do not understand, which makes convergence a lot easier. In other words, you leave your formulas with integrations over transverse momenta and just interpret them as numbers without actually carrying out the integration.

So I don't know how to cure the field theory and, therefore, I do not really propose that at the present time we should start with some specific field theory. I'm trying to meet both ways across the middle in order to have unitarity, and on the other hand to avoid divergences.

ROBERT J. YAES (University of Texas):
Can particles interact by the exchange of a single parton which would give an appropriate peaking in the crossed channel?

FEYNMAN: No! The two things interact due to the overlap between the partons that have zero momentum, but that's not one parton; it's a characteristic of the system that the mean number of partons in there is infinite, not one. Perturbation theory is wrong. Here, diagrams of one, two, ... interacting partons, all going across, are to be added together to produce a resonance "gluck-glock" also known as a

Regge particle.

ROSENMAN: Doesn't the fact that partons decay imply that they have structure?

FEYNMAN: No! Such a decay only has to do with the energies. It is possible theoretically to have, say, a structure of three much lighter than just one, so you can't produce only one, for it would disintegrate into two three's or into some other combination of complicated objects. But these underlying objects must not be like quarks with a carrier quantum number of one-third, because then nothing can disintegrate. The quark picture is different, there you have to invent something to hold them together if you don't see them.

E. C. G. SUDARSHAN (University of Texas):
Could it be that the partons are like the ether for light waves -- that the vibrations in the partons are like vibrations in the ether and may never be found?

FEYNMAN: Yes, and in the case of "the ether never being found", it was ultimately realized that there wasn't any ether at all, the ether was one of these scaffoldings to create a theory. It was later realized that the ether was an irrelevant complication and it may be that the partons are also nonexistent. The partons were only put in there to have

a field theory but the world, in fact, is not field theory, rather it's something else. However, the world has unitarity, analyticity and relativistic invariance, ..., and so we use the field theory with partons. Thus, "this" and "that" is right but the real view is wrong, just like with the case of the ether.

PARTONS*

Richard P. Feynman
Lauritsen Laboratory of Physics
California Institute of Technology
Pasadena, California

*Notes edited by Peter N. Dobson, Jr., with the help of Leo Pilachowski.

LECTURE 1

PARTONS, SCALING, AND REGGE BEHAVIOR

The general outline of these lectures will be something like this: first, I will describe the general ideas of the parton concept. Then I will point out how the experiments on deep-inelastic scattering of electrons and neutrinos by nucleons give us information about the nature and distribution of the partons. This will be followed by a discussion of the applications of parton ideas to hadronic collisions. I shall unfortunately not have time to discuss predictions of the products (the particles that come out) in deep-inelastic scattering. I will also discuss a number of theoretical ideas that have never been worked through to the end.

It should be pointed out that the parton theory is a conglomeration of rather imprecise ideas. In particular, I want to emphasize that in hadronic collisions parton ideas have not been very effective. They did strongly suggest that there would be scaling and a plateau in the rapidity plot, and these suggestions appear to be correct, but we haven't gotten anywhere with the detailed questions--proportions of π's and K's to be expected, and that sort of thing--and that has been a disappointment to me. The idea of partons was originally conceived in an attempt to

understand high-energy hadronic collisions, and I worked intensively on this in the summer of 1968. At that time I made a visit to SLAC and learned about the results on deep inelastic scattering of electrons. I saw that the experiments were tailor-made for investigating partons, and were easy to interpret in terms of the pictures I had already developed for the strong interactions. The deep-inelastic scattering experiments can identify the kinds of partons and how they are distributed, whereas, as yet, I see no clear way to do this from a study of hadronic collisions.

The idea of partons is an old one--it is a field theory idea. In a field theory the physical state of a particle, say a proton, is described by a wave-function which gives the amplitudes for finding various configurations inside the proton. Configurations of what? If we consider quantum electrodynamics, the wavefunction for positronium contains some amplitude for finding an ideal Dirac electron-positron pair, an amplitude for finding those plus a virtual photon, etc. There is a certain amplitude--small because the coupling is small--for finding two electrons and two positrons. Now the constituents whose distributions are given by these amplitudes are the objects created by the field operator ψ, which are not physical electrons. They are what we sometimes call "bare" electrons. The physical electron

we find in the laboratory has some amplitude to be a bare electron, but also some small amplitude to be a bare electron plus virtual photons, additional pairs, etc. It is unfortunate for our language of description (although very fortunate for mathematical simplicity) that the bare and physical electrons are approximately the same. That is, the physical electron has around 99% probability ($\sqrt{.99}$ in the amplitude) for being a bare electron and only a small probability for being more complicated. In strong interactions, the latter probability will not be small, and the physical particles and bare objects of the underlying field theory need not be so closely connected. So we introduce the word "parton" to refer to these bare constituents. The assumption is that we have an underlying field theory which contains operators of various types, which carry spin, isospin, strangeness, and so on. The partons are the objects created by these basic field operators. They are the quanta of the fields of the underlying field theory.

In this picture, the wavefunction of a proton contains the amplitudes to find configurations of partons with various momenta. In principle, the wavefunction contains an amplitude C_0 to find nothing at all (for a proton, $C_0 = 0$ since the vacuum does not have the quantum numbers of a proton), an amplitude $C_{1i}(\vec{p})$ to find one parton of type i (which index

specifies such things as spin, isospin, etc.) with momentum $\vec{p}$, an amplitude $C_{2ij}(\vec{p}_1,\vec{p}_2)$ to find two partons of types i and j with momenta $\vec{p}_1$ and $\vec{p}_2$, and so on. Mathematically, we could write

$$|\psi> = C_0|0> + \sum_i \int d^3\vec{p}\ C_{1i}(\vec{p})a_i^*(\vec{p})|0>$$

$$+ \sum_{i,j} \int d^3\vec{p}_1 d^3\vec{p}_2\ C_{2ij}(\vec{p}_1,\vec{p}_2)a_i^*(\vec{p}_1)a_j^*(\vec{p}_2)|0> + \ldots,$$

where $|0>$ represents the vacuum, and $a^*(\vec{p})$ is the creation operator for a parton of type i and momentum $\vec{p}$. We will ignore all complications of field theory which may make the rigorous basis of such an expression suspect.

If such an expression does make some sort of sense, there will be a wavefunction of this sort for a proton in any circumstance: at rest, moving along the z-axis with a certain momentum, and so on. For collisions at high energy, we need the wavefunction for a particle with a large momentum. This is not easily obtained from the wavefunction for the particle at rest, however. The wavefunction is not relativistically invariant, and the transformation which connects one Lorentz frame with another involves the Hamiltonian, which we do not know--it is our hope that we can understand many features of high-energy collisions without a detailed knowledge of the Hamiltonian. Thus we are led to try directly to guess or

understand the wavefunction of a rapidly moving particle.

One of the first features we come across is the scaling behavior of the wavefunction. Let us consider a proton, say, of large total momentum P, and consider the amplitude that this proton consists of just two partons with momenta $\vec{p}_1$ and $\vec{p}_2$. We write the longitudinal momentum of one parton as x_1P and of the other as x_2P. Then we have, of course, $x_1 + x_2 = 1$. The partons will have transverse momenta $\vec{Q}_1$ and $\vec{Q}_2$. In that part of the wavefunction where $\vec{Q}_1$ and $\vec{Q}_2$ are finite (and remain finite as $P \to \infty$), the amplitude for finding x_1, x_2, $\vec{Q}_1$, and $\vec{Q}_2$ is independent of P in the high P limit. If, in fact, there is no substantial amplitude for finding partons with large transverse momenta, then the wavefunction itself may be regarded as scaling in this way. That this may be true was suggested to me by the observed behavior in high-energy collisions, where we find the products limited to about 300 MeV/c in transverse momentum. This makes it reasonable to suppose that the wavefunction does not contain partons with large transverse momentum.

The simplest elementary argument to support the general notion of scaling of the wavefunction if the transverse momenta are limited is the following. Suppose the wavefunction refers to a particle with

total energy E. In perturbation theory, the amplitude for finding this as two subsystems of energies E_1 and E_2 will be proportional to an energy denominator factor

$$A/(E-E_1-E_2) .$$

Let P be the longitudinal momentum of the particle, and $\vec{Q}_o$ its transverse momentum. The partons are assigned longitudinal momenta x_1P and x_2P and transverse momenta $\vec{Q}_1$ and $\vec{Q}_2$. We are interested in the limit of very large P, with the Q's fixed. Since conservation of momentum requires $x_1 + x_2 = 1$, we can rewrite the energy denominator in the form

$$(E-P) - (E_1-x_1P) - (E_2-x_2P) .$$

Letting m be the particle mass, and m_1 and m_2 the parton masses, we have

$$E = \sqrt{m^2+P^2+Q_o^2}$$

and

$$E_i = \sqrt{m_i^2+x_i^2P^2+Q_i^2} .$$

When P is very large, we can approximate these expressions by

$$E \cong P + (m^2+Q_o^2/2P)$$

and

$$E_i \cong x_iP + (m_i^2+Q_i^2/2x_iP) .$$

Inserting these expressions in the energy denominator, the amplitude becomes

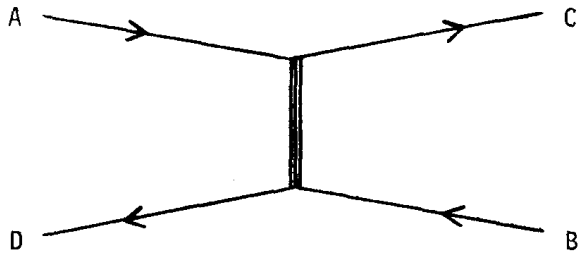
$$\frac{AP}{\frac{m^2+Q_o^2}{2} - \frac{m_1^2+Q_1^2}{2x_1} - \frac{m_2^2+Q_2^2}{2x_2}}$$

There will also be normalization factors like $1/\sqrt{2\omega}$, etc. In general, as $P \to \infty$, these factors multiplied by the P in the numerator will either go to zero or approach a non-zero constant. The limiting amplitude thus depends only on the x's and the transverse momenta. Note that in the approximations used for the energy we have assumed the x's are not too small; clearly, the argument breaks down for x's less than order 1/P. This is what we will call the "wee" region, and we will find that in the wee region of x there are a large number of partons, which complicate the interpretation of the wavefunction. We also note that there is negligible amplitude to find partons with negative (and non-wee) values of x. For if x_1 is negative, E_1 is $|x_1|P$ and $E_1 - x_1P$ is $2|x_1|P$, not of order 1/P as before but of order P. So the large denominator as $P \to \infty$ makes the amplitude negligible for finite negative x.

I am going to interrupt this discussion at this point to discuss one of the ideas I mentioned earlier that only partially worked. This concerns the ideas about high-energy collisions that come from the Regge pole approach. We find that various cross sections fall as inverse powers of s, the square of the total

center-of-mass energy, in such a manner that the powers depend on the quantum numbers exchanged but not on the precise process involved. For example, in a collision that involves a ρ exchange, the cross section as a function of s at fixed momentum transfer t determines a power of s, $\alpha(t)$. This power $\alpha(t)$ is (supposedly) the same as found in any other collision involving ρ exchange. This is analogous to the case of resonances: the position of the resonance is independent of the process in which it is formed. Now suppose I gave you the strong interaction Hamiltonian and asked you to compute the various powers that we find in Regge theory. You could simply use the Hamiltonian to work out the details of a particular process, say $\pi^+ + p \to \pi^0 + \Delta^{++}$, and extract the powers from the high-energy behavior of the cross section. But that would be something like getting the mass of the Δ^{++} from the same process by computing all the details of $\pi^+ + p \to p + \pi^+ + \pi^0$ and looking for the peak in the $p\pi^+$ mass distribution. That isn't the way we calculate energy levels. We expect, instead, to be able to solve an eigenvalue equation, $H\psi = E\psi$. In other words, we use the Hamiltonian directly to find an eigenvalue which does not depend on a specific process. The question is whether or not the Regge parameters can also be found in some such deeper way.

Suppose we have a scattering process A + B → C + D, which we represent schematically by a diagram:



This is supposed to indicate that something (a "Reggeon?") is exchanged in the process. From our point of view, some "pieces" of the fast moving particle A fall off and are absorbed by B. Now in the center-of-mass system, as the momentum P of A and B goes to infinity, both the longitudinal momentum and the energy that are transferred by the exchange will be of order 1/P, and so we neglect them in high-energy collisions. The pieces of A that are emitted are those with momentum not too far from zero. Let us take F(E) as the amplitude for A emitting this "stuff" when A has energy E in the center-of-mass. We let β be the amplitude for propagating it to the other particle, and G(E) be the amplitude for B's absorbing it. The overall amplitude is

$$F(E)\beta G(E) \ . \qquad (1)$$

Now we suppose that the amplitude F(E) that A emits this stuff does not depend on what particle B absorbs it and how it is done. That is, the amplitude above factors: F(E) does not depend on B.

Now this time suppose particle A has energy E_1 and it emits the "thing" with amplitude $F(E_1)$, but now particle B has a different energy, say E_2, (moving, of course, in the opposite direction to A) so the proper absorption amplitude is $G(E_2)$, and our amplitude is

$$F(E_1)\beta G(E_2) \ . \qquad (2)$$

But relativity tells us that this must not depend on the coordinate system, and hence can only depend on the invariant E_1E_2 ($s = 4E_1E_2$). For example, we may have transformed by velocity v from the center-of-mass system, so A's energy is increased by a factor $f = \sqrt{(1+v)/(1-v)}$, and B's decreased by the same factor: $E_1 = fE$, $E_2 = f^{-1}E$, and (2) must not depend on the factor f. If (2) is to depend only on the product E_1E_2, then $F(E_1)$ and $G(E_2)$ each must be proportional to a power of E, the same for each:

$$F(E) \sim E^{\alpha}, \ G(E) \sim E^{\alpha}$$

and the overall amplitude goes as s^{α}. (One might object that the "stuff" being exchanged in the system at velocity v is not the same as that in the center-of-mass, so β depends on f. But the stuff has zero energy and longitudinal momentum, so when transformed it is the same.) We are led in this way to associate Regge behavior with the exchange of something with very low momentum. This involves the probability for finding slow partons in a wavefunction for a particle

with large momentum, which in principle can be calculated directly given the Hamiltonian.

In order to see how this might be done, I analyzed the situation as follows. Let F be some large energy scale, and consider a wavefunction for a particle with longitudinal momentum $P_z = x_0 F$. (We are going to keep x_0 fixed and let $F \to \infty$. x_0 is arbitrary, of course; you may choose it to be 1.) This wavefunction is a solution of $H\psi = E\psi$, where H is the (unknown) Hamiltonian of the system. ψ is also an eigenfunction of the operator for longitudinal momentum, P_z; that is, $P_z\psi = x_0 F\psi$. We will assume that $x_0 F$ is very large; then the energy E is nearly equal to $x_0 F$, and using the same approximation as earlier we have

$$E = x_0 F + (m^2+Q_0^2/2x_0 F) \ ,$$

with $\vec{Q}_0$ being the transverse momentum. Now we can write

$$2F(H-P_z)\psi = (m^2+Q_0^2/2x_0)\psi \ .$$

If we define an operator

$$W = \lim_{F \to \infty} 2F(H-P_z) \ ,$$

then ψ at infinite momentum is a solution of $W\psi = w\psi$, where the eigenvalue is

$$w = (m^2+Q_0^2)/2x_0 \ .$$

Applying this idea with various model Hamiltonians suggest W operators, which depend only on x's and Q's,

which is how I arrived at the scaling idea for the wavefunction ψ.

Let's take a very easy example--the scalar ϕ^3 theory, described by the Lagrangian

$$\mathcal{L} = (\nabla\phi)^2 - \mu^2\phi^2 + g\phi^3 .$$

The Hamiltonian for this theory has the form

$$H = \sum_k \omega(\vec{k})a^*(\vec{k})a(\vec{k})$$

$$+ g \sum_{k_1k_2k_3} \frac{\delta^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3)}{\sqrt{2\omega_1}\sqrt{2\omega_2}\sqrt{2\omega_3}} [a(\vec{k}_1)+a^*(-\vec{k}_1)]$$

$$\times [a(\vec{k}_2)+a^*(-\vec{k}_2)][a(\vec{k}_3)+a^*(-\vec{k}_3)] ,$$

where $a^*(\vec{k})$ is the creation operator for a scalar particle of momentum $\vec{k}$, $\omega(\vec{k}) = \sqrt{k^2+\mu^2}$, and $\omega_i = \omega(\vec{k}_i)$. The operator for longitudinal momentum is

$$P_z = \sum_k k_z a^*(\vec{k})a(\vec{k}) .$$

Now we want to take $k_z = xF$ and study the limit $F \to \infty$. If x is positive, we have $\omega \to k_z$ in this limit and so we can use the usual approximation

$$\omega - k_z \cong (\mu^2+Q^2)/2xF .$$

For x negative, however, ω approaches $-k_z$ and so

$$\omega - k_z \cong 2|k_z| = 2|x|F .$$

Now we see that if we write $2F(H-P_z)$ in terms of the x's and transverse momenta, we get

$$\sum_{x>0} \frac{\mu^2+Q^2}{x} a^*(x,\vec{Q})a(x,\vec{Q}) + \sum_{x<0} 4F^2|x|\ a^*(x,\vec{Q})a(x,\vec{Q})$$

for the kinetic energy contribution to the operator. The second term here looks like the largest part, but it is really unimportant except in the following sense. If we start with a wavefunction containing positive x's, and in a perturbation treatment some operator generates something with a negative x, this term will give an enormous contribution to the denominator in the 1/(H-E) factor in the perturbation expansion. So the overall effect of this term is simply to kill all negative x's--to get rid of things moving backward.

The operators $a^*(x,\vec{Q})$ introduced in this expression are not the same as the original creation operators $a^*(\vec{k})$. The latter obey commutation relations

$$[a(\vec{k}_1),\ a^*(\vec{k}_2)] = \delta^{(3)}(\vec{k}_1-\vec{k}_2),$$

whereas we want the former to obey

$$[a(x_1,\vec{Q}_1),\ a^*(x_2,\vec{Q}_2)] = \delta(x_1-x_2)\ \delta^{(2)}(\vec{Q}_1-\vec{Q}_2)\ .$$

Since $\delta(k_{z1}-k_{z2}) = F^{-1}\delta(x_1-x_2)$, this implies we must introduce a factor $\sqrt{F}$ in the definition of the new operators; that is,

$$a^*(\vec{k}) \to F^{-\frac{1}{2}}a^*(x,\vec{Q})\ .$$

On the other hand, the sum over longitudinal momenta will be F times a sum over x, so that we also have the replacement

$$\sum_{k} \to F \sum_{x,Q}\ .$$

We see that the scale factors cancel in the terms involving $\Sigma a^* a$. When we consider the interaction term, converting the sums over $\vec{k}_1$, $\vec{k}_2$, and $\vec{k}_3$ to sums over the x's and transverse momenta introduces a factor F^3, while transforming the a's gives $F^{-3/2}$. The factors $1/\sqrt{2\omega} = 1/\sqrt{2|x_i|F}$ give another $F^{-3/2}$, and converting the δ-function of longitudinal momentum to a δ-function of the x's brings in $1/F$. Overall, we have just the factor $1/F$ left, which is canceled by the F in $2F(H-P_z)$. Putting this all together, the interaction contribution is

$$(g/\sqrt{2}) \sum_{x_i, Q_i} \frac{\delta(x_1+x_2+x_3)\ \delta^{(2)}(\vec{Q}_1+\vec{Q}_2+\vec{Q}_3)}{\sqrt{x_1|x_2|x_3|}}$$
$$\times [a(x_1,\vec{Q}_1)+a^*(-x_1,-\vec{Q}_1)][a(x_2,\vec{Q}_2)+a^*(-x_2,-\vec{Q}_2)]$$
$$\times [a(x_3,\vec{Q}_3)+a^*(-x_3,-\vec{Q}_3)] \ .$$

When we compute the limit $F \to \infty$ to obtain the operator W, we can make use of the fact that negative x's in the wavefunction are killed by the $4F^2$ term discussed earlier. Therefore, the interaction terms with three creation or three annihilation operators can be discarded, since the δ-function requires the sum of the x's to be zero. Thus we are finally led to the effective W operator:

$$W = \sum_{\substack{x,Q \\ x>0}} \frac{\mu^2+Q^2}{x}\ a^*(x,\vec{Q})a(x,\vec{Q})$$

$$+\frac{g}{\sqrt{2}}\sum_{\substack{x_1x_2Q_1Q_2\\ x_1,x_2>0}}$$

$$\times\frac{[a(x_1,\vec{Q}_1)a(x_2,\vec{Q}_2)a^*(x_1+x_2,\vec{Q}_1+\vec{Q}_2)+\text{h.c.}]}{\sqrt{x_1x_2(x_1+x_2)}}\ .$$

We can see how scaling is implied by this limit, since the wavefunction is an eigenfunction of W, which has no F dependence, with eigenvalue w, which is also independent of F. (Strictly speaking, W is singular near $x = 0$ and in general cases has no strictly scaling ψ solutions--the complications this produces in the wee region will be discussed later.)

Now we notice that this equation is homogeneous of degree -1 in x. It is not obvious that W has this property, because of the square roots, but it follows from the definition of the a's. This implies that under $x \to \lambda x$,

$$a^*(x,\vec{Q}) \to \lambda^{-\frac{1}{2}}a^*(\lambda x,\vec{Q})\ ,$$

since this is equivalent to changing the energy scale via $F \to F/\lambda$. It is then easy to verify that if we replace x by λx, $W \to (1/\lambda)W$. (That this must be true in general can be seen since the eigenvalues for different x_o vary as $1/x_o$.) The operators that change the scale of the x's are just the boosts along the longitudinal direction. The infinitesimal generator of these boosts may be written

$$B = \sum_i x_i(\partial/\partial x_i)\ ,$$

and the fact that W is homogeneous in x^{-1} implies the commutation relation

$$[W,B] = WB - BW = W \ .$$

We would like to extract from our knowledge of W something about the small (but finite) x region. Suppose we started with an approximate wavefunction containing one fast moving parton. In the next approximation, the interaction would produce some partons moving slower. These, in turn, would give rise to some more moving slower still, and so on. The slow partons would result in this way from a long cascade. Now we come across cascades like this in the theory of cosmic ray showers, and the general qualitative feature of such showers is that when there are a large number of steps in the cascade, we ultimately produce a distribution with power law behavior in the fraction x of the original energy or momentum, which is independent of how the shower started. In shower theory, there are many different eigenvalues for the power, which do not depend on the starting point, although the amount of each power present does. If we concentrate on the lowest power, the one that is most effective as the shower becomes deeper, the shape of the distribution becomes independent of how the shower started.

How can we make this qualitative idea more explicit? Looking at small x's is equivalent to

keeping x fixed and letting x_0 become large. If we let x_0 go to infinity, the eigenvalue w goes to zero, and the wavefunction becomes a solution of $W\psi = 0$. In this case it is possible to choose ψ to be simultaneously an eigenfunction of B (which is not possible for $w \neq 0$ since B and W do not commute); that is, we take $B\psi = b\psi$, where b is some number. But this means the wavefunction must have the property

$$\psi(x_1, x_2, \ldots) = \lambda^b \psi(\lambda x_1, \lambda x_2, \ldots) .$$

If we ask for the probability for finding particles at x in range dx, we could calculate this by scaling down the wavefunction used in a calculation of the probability for some other value of x. Thus the probability will have to depend on some power of x related to this eigenvalue b (the relationship is determined by how the wavefunction is used to compute the probability). In any case, we will get a distribution of the general form $x^\beta dx/x$. This will be valid for small x down to the wee region; i.e., down to x of order 1/E. The smooth joining of the wee region onto this distribution implies that this region contributes a total probability of order $(1/E)^\beta$, which is the sort of thing we were looking for to solve the problem originally proposed. Namely, to find the powers that show up in the Regge theory from a knowledge of the Hamiltonian, we could construct the limiting operator W and find

simultaneous eigenfunctions of W with eigenvalue 0 and of B. The eigenvalues of B would then determine the powers in question. In fact, there are some technical difficulties with this idea. In particular, to get simultaneous eigenfunctions of W and B we must let x_0 become infinite. But this is not compatible with the usual boundary conditions for the wavefunction whereby we must start somewhere with a finite number of particles. At best, we have here an outline of a theory that might lead us from the Hamiltonian to a theory of Regge poles.

The main point we want to remember from this discussion is that it is reasonable to expect a distribution at low x that goes like a power of x. The value of the power may depend on the quantum numbers being considered, so that exchanges involving different quantum numbers can have different power law dependence on 1/E. If total cross sections go to constants as $E \to \infty$, then the lowest allowed value of β is zero. Unfortunately, it is not at all clear how this particular value follows in any simple way from the theory.

LECTURE 2

PARTON DISTRIBUTIONS AND DEEP-INELASTIC SCATTERING

In the first lecture, we tried to describe the character of the wavefunction for a proton moving with a very high momentum, and came to the conclusion that the amplitude to find a parton with a certain momentum inside should depend only on the fraction, x, of the proton's longitudinal momentum carried by the parton, and on its transverse momentum. We also discussed how the distribution of partons might be expected to behave in the region of small x, and decided by analogy with shower theory that it was reasonable to expect power law behavior; that is, a distribution consisting of a superposition of terms like $x^{\beta}dx/x$. In the early development of these ideas, I came to the conclusion that it was most likely that the lowest eigenvalue for β would be zero. There were two sorts of arguments which indicated this. One was based on an analogy with bremsstrahlung: if you compute the distribution of photons in the field of a fast moving electron, you get this sort of distribution. The other argument comes from the notion outlined last time that it is the wee partons that are involved in high-energy hadronic collisions, and $\beta=0$ is required to give

total cross sections that are constant at high energy. To get a constant cross section the mean number of wee partons must then be constant, independent of P. This is because the wee region has a width of order 1/P (the distribution must fall rapidly to zero for $x < 0$), and it must match onto the small x distribution in the neighborhood of $x = 1/P$. For a dx/x distribution, this implies a height of order P, and a total number of partons in the wee region that is constant, independent of P. If we go to higher momentum P, the x distribution remains the same, according to scaling, except that now we can use it down to smaller values of x, so it climbs to a higher value before turning over. If we had a small x distribution like $x^{\beta}dx/x$, the value at the peak would be of order $(1/P)^{\beta-1}$, giving an area in the wee region of order $(1/P)^{\beta}$.

The operator W we discussed in the last lecture suggests the power law, which we take in the present case to be dx/x, but it does not tell us how things vary in the wee region (where the approximation $\omega = k_z + (\mu^2+Q^2)/2k_z$ is inadequate). To get some rough idea of how a wavefunction could describe such a region, we look at a simple example of a model which is so simple that we can see the effect of our approximation of ω. It is a model with scalar particles coupled to a c-number source, described by

the Hermitian operator

$$H = \sum_k f(\vec{k})a^*(\vec{k})a(\vec{k}) + \sum_k [s(\vec{k})a^*(\vec{k}) + s^*(\vec{k})a(\vec{k})],$$

where $s(\vec{k})$ is just a numerical valued function, and $a^*(\vec{k})$ is the creation operator for a scalar field. It is easy to verify that

$$|\bar{\Psi}> = \exp\{\sum_k c(\vec{k})a^*(\vec{k})\}|0>$$

is an eigenvector of H if we choose $c(\vec{k}) = -s(\vec{k})/f(\vec{k})$. If, as in the example considered last time, we want to look at the difference between the Hamiltonian and the longitudinal momentum operator, we must take $f(\vec{k}) = \omega - k_z$. Also, to get scaling behavior, the function $s(\vec{k})$ must scale as $F^{-3/2}$, where F is the energy scale. To get the desired dx/x behavior, we choose it to be proportional to $\omega^{-3/2}$, giving

$$c(\vec{k}) = \alpha/(\omega - k_z)\omega^{3/2} .$$

The mean number of particles with momentum $\vec{k}$ is found by computing the expectation value of $a^*(\vec{k})a(\vec{k})$, which is just $|c(\vec{k})|^2$. Thus the distribution function for particles with transverse momentum $\vec{Q}$ and longitudinal momentum k_z in range $dk_z d^2\vec{Q}$ is

$$\alpha^2 dk_z d^2\vec{Q}/(\omega - k_z)^2\omega^3 .$$

Now suppose we have $k_z = xP$ and consider P very large. For $x > 0$,

$$\omega - k_z \to (Q^2 + m^2)/2xP$$

and

$$\omega \to xP \ ,$$

so the distribution becomes

$$Pdx(2xP/Q^2+m^2)^2 d^2\vec{Q}/x^2P^3 = [4\alpha^2/(Q^2+m^2)^2]d^2\vec{Q}(dx/x).$$

This has just the dx/x behavior we have been talking about. (With a $\vec{Q}$ dependent coefficient, but since α could depend on $\vec{Q}$, we know nothing about it.) For $x < 0$, $\omega - k_z \to 2|x|P$, so the distribution becomes

$$(\alpha^2/4P^4)(dx/x^5) \ ,$$

which falls rapidly to zero as $P \to \infty$. Of course, a model like this doesn't prove anything, but it does give us some mathematical expressions to manipulate. One of the great troubles of the parton model is that there isn't any mathematical expression, even to play with, that is of any rigor.

The wee region complicates the wavefunction very much. It is a purely technical point, but it is incorrect to speak of the wavefunction scaling. The wavefunction contains the amplitudes for all kinds of configurations, including partons in the wee region, and these pieces of the wavefunction do not scale. It is the density matrix that exhibits

scaling--the probability for finding a particular value of longitudinal momentum not in the wee region scales, for instance. The probability for finding four partons with various x's and Q's will approach a limit as $P \to \infty$, if you ignore where the other partons are. But, of course, you cannot ignore where the "other" partons are when talking about a wavefunction, for a wavefunction is necessarily a function of all the particles present, and not just a few selected ones, as a density function may be. This is analogous to another situation in physics which has technically the same sort of difficulty. Consider a one-dimensional liquid which has a "surface" at one end that is affected by some kind of surface tension forces, and a surface at the other end with some (perhaps different) forces acting. If the liquid is thick enough--the distance between the two surfaces large enough--what happens at one end will not affect what happens at the other. But this independence is not apparent in the wavefunction, which must give the amplitude for every configuration of all the particles in the liquid. However, if we compute a density matrix for the probability, say, of finding two particles at two specific locations in the liquid, the independence will show up in a simple way--the probability will become a product of independent single particle probability distributions

when the separation is large enough.

A variable like the logarithm of x is useful to make a closer correspondence with the theory of a liquid. It is the quantity called rapidity, defined by

$$y = (1/2)\log[(\omega+k_z)/(\omega-k_z)].$$

In the large P limit,

$$(\omega+k_z)/(\omega-k_z) \cong 4P^2x^2/(Q^2+m^2),$$

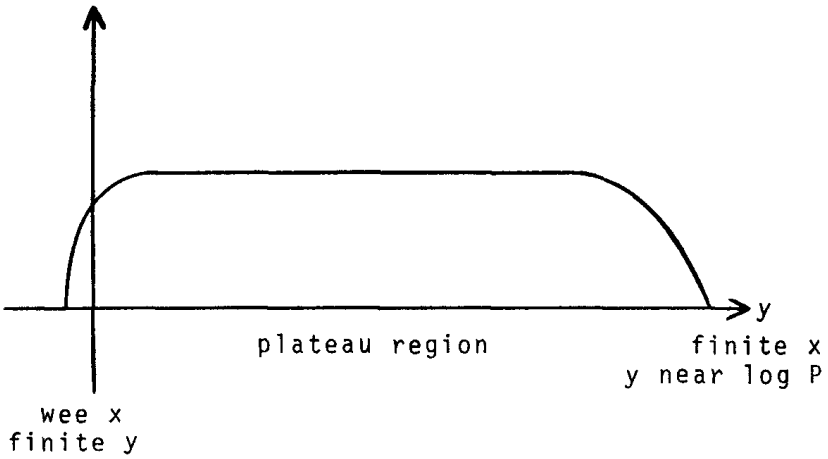
so that

$$y \cong \log x + \log P.$$

This will be valid even for small x, so long as x is larger than order 1/P (not wee). Since $dy = dk_z/\omega \to dx/x$, we see that for y in the region corresponding to small x, the assumed dx/x distribution implies a flat "plateau" in the rapidity plot. The rapidity is related to velocity via

$$v/c = \tanh y.$$

This has the useful property that a Lorentz transformation along the longitudinal direction simply changes y by an additive constant. The distribution at the ends of the plot (x near 1 and x wee) may be complicated, but our analogy with shower theory suggests the two ends are independent, like the two surfaces of the one-dimensional liquid (see figure).



Density to find a parton vs. rapidity

We can ask what the probability is of finding no partons in the small x region--that is, we look for configurations that have many wee partons, and partons with x greater than some x_0 (where x_0 stays finite as P goes to infinity), but none in-between. Our picture of how the small x region is built up by a long cascade process leads us to conjecture that the probability of finding no particles in some region should be proportional to $\exp(-c\bar{n})$, where $\bar{n}$ is the mean number of particles in that region. The constant c would be 1 for a Poisson distribution, but forces between partons may make it easier to make one parton when another is present, etc., so that successive creations are not statistically independent. We could have some sort of Markovian chain in the plateau region in which some number of partons

makes up the statistically independent unit--and so we introduce the constant c to represent these deviations from Poisson distribution behavior. In the small x region, the distribution as a function of rapidity is flat, so the mean number in a gap of width Δy will be proportional to Δy. For large P we see that $\Delta y \stackrel{\sim}{=} \log x_0 + \log P$, and so it increases as the logarithm of the momentum. This means the probability of finding no partons in the gap will vary with momentum as $\exp(-c \log P) = P^{-c}$. We will consider a number of applications of this basic principle later. The value of c probably depends on the quantum numbers "carried across the gap"; that is, on the way in which we apportion the quantum numbers of the wavefunction among the partons on the two sides of the gap.

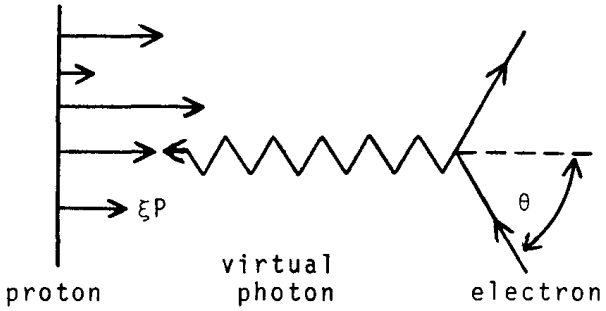
We have a picture at this point in which we have a reasonable guess as to the average distribution of partons, and we have an idea that the probability of finding no partons in some large range Δy of y will behave exponentially with Δy as $\exp(-c\Delta y)$. It is a naive one-dimensional picture--we are missing information about transverse momentum distributions. Instead, we simply guess from experimental considerations that the probability of finding partons of high-transverse momentum is small and falls rapidly with increasing transverse momentum. This does

indicate that parton-parton interactions are significant only among partons of low-relative momentum. In other words, there can't be a big amplitude that two partons interact to produce a large change in their relative momentum, since this would lead to large transverse momentum partons. Then it is just as true that there will not be an exchange of a large amount of longitudinal momentum. Basically, the argument implies that interactions occur only between partons separated by a finite amount of rapidity.

So far, we have not even said what the partons might be like--spin 1/2, spin 0, charged, neutral...? Unfortunately, we are as yet unable to deduce anything from high-energy hadron collisions (except that the original idea of scaling in longitudinal momentum works). On the other hand, I discovered to my great glee that the SLAC experiments on deep-inelastic scattering were looking directly at the parton probability distributions, and that Bjorken scaling was exactly the same as my conclusion that the wave-function scaled. Furthermore, the parton interpretation of these experiments gave a dx/x distribution, which confirmed my opinion which I was then forming on this point (it wasn't really certain enough in my mind to call it a prediction).

Let us see how the deep-inelastic scattering experiments look in the parton-model interpretation.

We have a proton with a large momentum P, which we regard as made up of a large number of partons carrying various fractions, ξ, of the total longitudinal momentum. (We temporarily use ξ instead of x for this quantity.) The electron produces a virtual photon which interacts with one of the partons:



This interaction involves a simple local coupling, consistent with the whole approach of expanding the proton's wavefunction in terms of the fundamental pointlike bare objects, each having the usual "minimal" coupling to electrodynamics (putting $\nabla \to \nabla - e\vec{A}$). The virtual photon carries energy and momentum; we can arrange to use a coordinate system in which its energy is zero, and write its 4-momentum:

$$q = (0, 0, 0, -2Px).$$

In this system, the proton's 4-momentum is

$$p = (P, 0, 0, P),$$

assuming that P is large enough that we can neglect the proton mass. Thus we have

$$-q^2 = 4P^2x^2$$

and

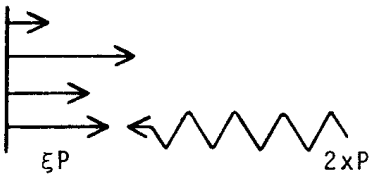
$$p\cdot q = 2P^2x.$$

The last of these is usually called $M\nu$, and the idea of Bjorken scaling is that the high-energy cross section should depend only on the dimensionless ratio

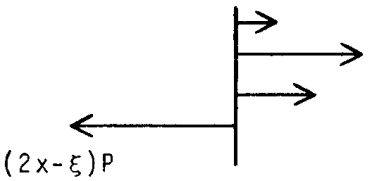
$$-q^2/2M\nu = x.$$

In the parton picture the scattering process looks like

Before:



After:



The energy of the hadron system before and after the collision is the same in this frame, since the photon brings in no energy. Now, the idea discussed earlier that the parton-parton interactions are limited to a finite range of rapidity means that these interactions contribute only a finite amount to the energy of the state. (The best mathematical way to state the principle is to say the unknown interactions have the same effects in a problem as finite uncertainties in parton masses would have.) The parton kinetic

energies, however, are of order P, as is the total energy of the proton. To order P, then, the energy of the hadrons in the final state (as well as in the initial proton state) is just the sum of the parton kinetic energies. In the example we are considering, conservation of energy (to this order) requires $2x-\xi = \xi$, or $\xi = x$. So the experiment directly samples partons with various values of x.

In these scattering processes, the photon is virtual and there are longitudinal as well as transverse components to the photon's polarization. The proportion of these varies with the electron's angle, θ, and so the cross sections for scattering of transverse photons, σ_t, and longitudinal photons, σ_s, can each, in principle, be separately measured. If the partons had spin 1/2, then at sufficiently large P (such that parton masses can be neglected), helicity will be conserved in the scattering, as we know from the theory of photon coupling. A parton whose direction was reversed as in the figure would have to absorb a unit of angular momentum in order to maintain the same helicity, and hence would contribute only to σ_t. Similarly, a spin 0 parton would contribute only to σ_s. The data are consistent with the assumption that all the charged partons have spin 1/2. That is, σ_s/σ_t is small. The early data indicated that σ_s/σ_t could be of the

order of 18±10%, so this fraction could be spin zero. A recent thesis by Riordan contains a more careful analysis of the possibility that all the charged partons have spin 1/2. This means that the residual σ_s is due to the finiteness of P, and consequently that helicity is not exactly conserved. In this case, σ_s must vary as $1/\nu$; in particular,

$$R = \sigma_s/\sigma_t = (1/\nu)f(x) = (1/q^2)g(x).$$

The data fit this hypothesis quite well, if, for example, $g(x)$ is roughly constant.

In these experiments we are measuring $f_1(x)$--the mean number of partons with a given value of x, weighted by the square of the charge. The square of the charge comes in, of course, for this is the probability a given parton interacts with the photon. Neutral partons would not contribute to the scattering f_1. The function $f_2(x) = xf_1(x)$ tells us how much momentum is being carried by the partons at x (with the same weighting). This is the function usually called νW_2, and the experiments indicate that it becomes constant at small x, consistent with a 1/x behavior of the basic distribution $f_1(x)$.

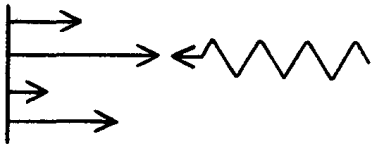
Another interesting region is near x = 1, as pointed out by Drell and Yan. Let us ask for the probability of finding a parton with x near 1; or equivalently, with x' = 1-x small. This means that

one parton will carry nearly all the proton's momentum, with the rest sharing a total of x'P. We will therefore have a gap in the rapidity plot from y about log P + log x' up to y near log P. The size of the gap is thus -log x', independent of P, so we still have scaling. Using the basic principle outlined earlier, we expect this configuration to occur with probability

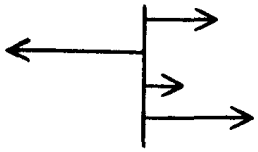
$$\exp(-c\Delta y) = \exp(c \log x') = (1-x)^c.$$

The distribution near x = 1 should behave as the differential of this; i.e., like $(1-x)^{c-1}dx$. Now we can connect this exponent to another experimental quantity. Consider elastic scattering with large q^2. We cannot have a picture like

Before:



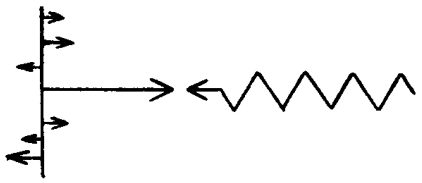
After:



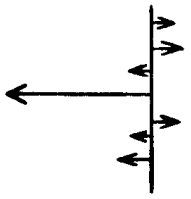
since in an elastic event, the final particle will still be a proton, and will not have the backward moving partons shown in the "after" picture. That is, the "after" picture has zero amplitude to be found in a proton moving to the left. So to get elastic scattering, we must require the photon to

interact with a configuration containing just one fast-moving parton, with all the rest in the wee region:

Before:



After:



Only in this case will there be a reasonable amplitude to find each configuration in a proton. Thus the probability of an elastic scattering event will be proportional to the probability of finding such a configuration, and this configuration has a gap in rapidity Δy of order $\log P$. This should go like

$$\exp(-c \log P) = P^{-c},$$

with c the same number as found from the parton distribution function near $x = 1$. If the elastic form factor really has the famous dipole form $(1/q^2)^2$, this would predict the distribution function near $x = 1$ to vary as $(1-x)^3$. Unfortunately, neither the elastic nor deep-inelastic data is sufficiently accurate to confirm this connection; but there is no inconsistency.

LECTURE 3

PARTONS AS QUARKS

Up until now we have been talking about the possibility that the wavefunction for a rapidly-moving particle consists of amplitudes for fundamental constituents, which we call partons, having limited transverse momentum and which scale in longitudinal momentum. If this picture is correct, then the determination of the characteristics of the partons becomes a really fundamental question in the theory of hadrons. In the last lecture we discovered that the experiments on deep-inelastic scattering of electrons by nucleons indicate that it is possible all the charged partons have spin 1/2. In this lecture we will investigate a specific hypothesis--that the charged partons have the quantum numbers of quarks--and see what the data have to say about this possibility. Identifying partons with quarks tells us that we have three kinds of fundamental constituents with quantum numbers summarized by the table:

Symbol	Charge	I_3	Strangeness	Name
u	2/3	1/2	0	"up"
d	-1/3	-1/2	0	"down"
s	-1/3	0	-1	"strange"

In the simple non-relativistic quark model, the proton is supposed to be made up of three quarks--two u's and a d. The parton picture, which is relativistic, doesn't allow a fixed number of constituents (which are bare objects, and not physical particles or quasi-particles anyway). So the "partons as quarks" are not quite the same thing as the quarks in the non-relativistic model.

With three basic types of partons (and their anti-particle counterparts), the distribution functions for the partons in protons and neutrons will be constructed from six functions describing the distributions of the specific types. In particular, the function $f_2^{ep}(x)$, measured in deep-inelastic scattering of electrons by protons, will have the form

$$f_2^{ep}(x) = (4/9)x[u(x) + \bar{u}(x)]$$
$$+ (1/9)x[d(x) + \bar{d}(x)] + (1/9)x[s(x) + \bar{s}(x)] .$$

In this expression, the function u(x) is the probability distribution of finding an up quark in the proton with fraction of longitudinal momentum x, while $\bar{u}(x)$ is the distribution function for the u-type anti-quark. The factors 4/9 and 1/9 are the squares of the charges of the appropriate quarks. We can write a similar expression for the neutron; but by isospin symmetry, the number of up quarks in the neutron is the same as the

number of down quarks in the proton, so we can write

$$f_2^{en}(x) = (1/9)x[u(x) + \bar{u}(x)]$$

$$+ (4/9)x[d(x) + \bar{d}(x)] + (1/9)x[s(x) + \bar{s}(x)].$$

The distribution functions u(x), etc., continue to be defined as above; i.e., they refer to the distributions in the proton.

So far we have two experimentally measurable functions expressed in terms of six unknown distributions; but there are some theoretical arguments which we can make from our theory of the wavefunctions. For example, the total strangeness due to strange quarks will be

$$-\int_0^1 s(x)dx,$$

and that due to the anti-particles will be

$$+\int_0^1 \bar{s}(x)dx.$$

Adding these gives the total proton strangeness, which is zero, of course, and we get the sum rule:

$$\int_0^1 [s(x) - \bar{s}(x)]dx = 0.$$

This also means that the s and $\bar{s}$ quarks will contribute net zero charge to the proton, so we can write a sum rule for the charge in terms of u(x), $\bar{u}(x)$, d(x), and $\bar{d}(x)$, as follows:

$$(2/3)\int_0^1 [u(x)-\bar{u}(x)]dx - (1/3)\int_0^1 [d(x)-\bar{d}(x)]dx = 1.$$

Similarly, for the total I_3, we get:

$$(1/2)\int_0^1 [u(x)-\bar{u}(x)]dx - (1/2)\int_0^1 [d(x)-\bar{d}(x)]dx = 1/2 .$$

We can combine the last two sum rules to obtain the simpler expressions

$$\int_0^1 [u(x) - \bar{u}(x)]dx = 2 ,$$

and

$$\int_0^1 [d(x) - \bar{d}(x)]dx = 1 .$$

Note the connection with the usual non-relativistic quark model: the net number of up quarks is 2, the net number of down quarks is 1, and the net number of strange quarks is 0.

We also believe from our theory that the wee quark region should be independent of whether the wavefunction describes a proton, a neutron, an anti-proton, or an anti-neutron. Our picture is that this region is built up from a long cascade and doesn't depend on whether we started with an up or down quark or anti-quark. So we believe that for x small enough, we should have

$$u(x) = \bar{u}(x) = d(x) = \bar{d}(x) \tilde{=} c/x \text{ (small } x) .$$

Since SU_3 is not an exact symmetry, it is possible that the wee region distribution of strange quarks is somewhat different, and we only require that

$$s(x) = \bar{s}(x) \tilde{=} c'/x \text{ (small } x) .$$

If we look at the ratio f_2^{en}/f_2^{ep}, and note that the distributions are all positive functions, we see that it must lie between 1/4 and 4. Near x = 0, the theoretical argument about the wee region implies the ratio goes to 1, in agreement with experiment. As x approaches 1, the experimental value of the ratio seems to be approaching the lower limit of 1/4 (it has fallen to about 0.4 by x = 0.8). This can be explained if the proton becomes almost pure up quark in the region near x = 1. I have an argument that suggests why this ratio might most likely be either 1/4 or 3/2, but I confess I thought it up after I saw the results. The idea is the following: near x = 1, we have one fast quark with the rest squeezed into the wee region. As we discussed in the last lecture, the probability of this goes as $(1-x)^{\gamma}$, where the exponent γ depends on the quantum numbers squeezed into the wee region. Now the net quark number is a good quantum number, and for the wee region it will either be 2 (if the fast parton is a quark) or 4 (if it is an anti-quark). We guess that it is harder to squeeze four quarks into the wee region than two, so that the four-quark case gives a larger exponent γ. Then the leading behavior will correspond to the wee region having the quantum numbers of a two-quark system. This can carry isospin 0 or 1, for strangeness zero. (If

you wish to guess that the lowest γ is for the wee region to carry strangeness, the argument generates the additional possibility that the ratio be 1.) If it is isospin 1, then the fast parton has to be a mixture of up and down quarks (so that the quark isodoublet is combined with the wee parton isotriplet to give an $I = 1/2$, $I_z = 1/2$ state). This gives a probability 2/3 for the fast parton to be a down quark, and 1/3 for it to be an up quark, so that $d(x) = 2u(x)$ for x near 1. With s, $\bar{s}$, $\bar{u}$, and $\bar{d}$ zero in this region we get $f_2^{en}/f_2^{ep} = 3/2$, which is ruled out by the data. The case where the wee quarks have $I = 0$ implies the fast parton is pure up quark, and gives the ratio 1/4.

It is also interesting to consider the sum of the e-p and e-n distributions--I'll call this the "e-d" distribution

$$f_2^{ed}(x) = f_2^{ep}(x) + f_2^{en}(x).$$

If we define

$$q(x) = x[u(x) + d(x)]$$

to be the momentum carried by non-strange quarks, and

$$\Sigma(x) = x\ s(x),$$

that carried by strange quarks (with similar definitions for the anti-quarks), then

$$f_2^{ed}(x) = (5/9)[q(x) + \bar{q}(x)] + (2/9)[\Sigma(x) + \bar{\Sigma}(x)].$$

Now common sense leads us to believe that the strange quarks do not carry as much of the momentum as, say, the down quarks. The net excesses are two up quarks and one down quark, with no excess of strange quarks. It doesn't seem reasonable that there should be a large number of pairs of s and $\bar{s}$ in the sea of quarks and anti-quarks without a correspondingly large number of pairs of d and $\bar{d}$. Anyway, it seems a fair guess that $\Sigma(x) \lesssim (1/2)q(x)$, particularly when we are not in the sea region--say, for x greater than about 0.2. So we conjecture that the strange quark contribution to f_2^{ed} is at most of order 10-20% and likely smaller. Experimentally,

$$\int_0^1 f_2^{ed}(x)dx = 0.31 \ .$$

Now if all the proton's momentum were carried by the charged partons--in this model, the quarks--then we should have

$$\int_0^1 [q(x) + \bar{q}(x) + \Sigma(x) + \bar{\Sigma}(x)]dx = 1 \ .$$

Solving these two equations for the amount of momentum carried by strange and non-strange quarks, we find that the strange quarks carry 74% of it. This is in violent disagreement with our common sense ideas. The alternative is to suppose that there are neutral partons, carrying roughly half the proton's momentum. Various models have been proposed

in which these neutral constituents are the particles that intermediate the quark-quark interactions, and are usually called gluons.

This brings up another question that deserves a short comment--namely, the question of understanding how the quark-quark interactions can account for the absence of free particles with quark quantum numbers. The idea that partons are quarks was originally suggested by Paschos and Bjorken--it was not my idea. When they suggested it, I was worried how it could fit into my picture. When an interaction knocks a parton backwards, it interacts only weakly with the other partons, and appears to be carrying quark quantum numbers off into the distance. One way out is to have long-distance harmonic potentials; then the interaction is weak at short distances and our picture of a collision is valid, while after the collision the force grows with distance and can keep the quark from getting out. (It is only necessary to have q^2 or $2M\nu$ large compared to the spacing of levels and not to the ultimate size of the potential itself.) I do not favor this picture. Another possibility is that, as the Hamiltonian acts to generate the evolution of the system to the asymptotic final state, it creates large numbers of quark-anti-quark pairs at intermediate rapidity, and these somehow associate to produce the usual quantum

numbers of hadrons. It is clear that neither of these proposals is like an ordinary field theory. If the quark theory proves to be right, we will have to learn a lot of exciting, interesting, and semi-paradoxical things.

Now let's turn to the scattering of neutrinos by nucleons. In the case of electron scattering, we knew there was a photon propagator going as $1/q^2$; in the neutrino case, we don't know whether or not there is an intermediate propagator or a direct coupling. That's an experimental question. For the sake of analysis, let's assume a direct 4-fermion coupling of the quarks to leptons. We will also ignore strangeness-changing interactions (that is, we set the Cabibbo angle to zero), and neglect the muon mass. With these assumptions, we can calculate the differential cross section for the scattering of a neutrino of energy E, producing a muon of energy $E-\nu = E(1-y)$ coming out. In units of $Gs/2\pi$, the result is

$$d\sigma/dy = 2$$

for a neutrino on a particle, and

$$d\sigma/dy = 2(1-y)^2$$

for a neutrino on an anti-particle. When a neutrino scatters from a nucleon to produce a μ^-, it must change a d into a u or a $\bar{u}$ into a $\bar{d}$. Thus, for neutrinos on protons,

$$d^2\sigma^{\nu p}/dxdy = 2xd(x) + 2(1-y)^2\, x\bar{u}(x).$$

(The factor x comes from the fact that the s against the parton is x times the s against the proton as a whole.) For neutrinos on neutrons, recall that the distribution for down quarks is u(x), etc., so

$$d^2\sigma^{\nu n}/dxdy = 2xu(x) + 2(1-y)^2\, x\bar{d}(x).$$

If we knew these two cross sections sufficiently well as functions of x and y, we could determine the four functions u, $\bar{u}$, d and $\bar{d}$. Then we would be able to predict the anti-neutrino cross sections, which are similarly given as

$$d^2\sigma^{\bar{\nu}p}/dxdy = 2x\bar{d}(x) + 2(1-y)^2\, xu(x),$$

and

$$d^2\sigma^{\bar{\nu}n}/dxdy = 2x\bar{u}(x) + 2(1-y)^2\, xd(x).$$

These formulas imply many relations we can predict and check, but we discuss here only those tests which can be made in the most immediate future. The targets would then be heavier nuclei which we can say contain roughly equal numbers of protons and neutrons. Therefore we define the cross section for scattering on an "average nucleon" by taking half the sum of the cross sections on proton and neutron. This gives

$$d^2\sigma^{\nu}/dxdy = q(x) + (1-y)^2\bar{q}(x),$$

and

$$d^2\sigma^{\bar{\nu}}/dxdy = \bar{q}(x) + (1-y)^2 q(x),$$

where $q(x) = x[u(x)+d(x)]$ is the momentum carried by non-strange quarks, as defined earlier. What do we know about these functions? First, we have a sum rule

$$\int_0^1 [q(x) - \bar{q}(x)]dx/x = 3$$

which simply states that the net number of quarks in the nucleon is three (called the Gross-Llewellyn-Smith sum rule). Furthermore, we have

$$q(x) + \bar{q}(x) = (9/5)f_2^{ed}(x) - (2/5)\,[\Sigma(x) + \bar{\Sigma}(x)].$$

We expect the strange quark contribution to be small, say, less than 1/6 of the total, except for x near zero. In any event, $(9/5)f_2^{ed}(x)$ is a strict upper bound for $q(x)+\bar{q}(x)$ (an observation due to Llewellyn-Smith).

The sum rule for net quark number is difficult to test since the low x values are rather uncertain and are weighted by 1/x; in any case, the present experiments described at this conference by Perkins are at any energy which I could have thought too low. If we integrate $d^2\sigma^{\nu}/dxdy$ over y to get $q(x)+(1/3)\bar{q}(x)$, and $d^2\sigma^{\bar{\nu}}/dxdy$ to get $\bar{q}(x)+(1/3)q(x)$, then 1.5 times the difference of the integrated cross sections is just $q(x)-\bar{q}(x)$. Perkins reports that it looks like the current data give a value of 3.5 ± 0.5 for the sum rule--so maybe we're

beginning to learn something about the real world. Let's also look at total cross sections. If we define

$$Q = \int_0^1 q(x)dx, \text{ and } \bar{Q} = \int_0^1 \bar{q}(x)dx,$$

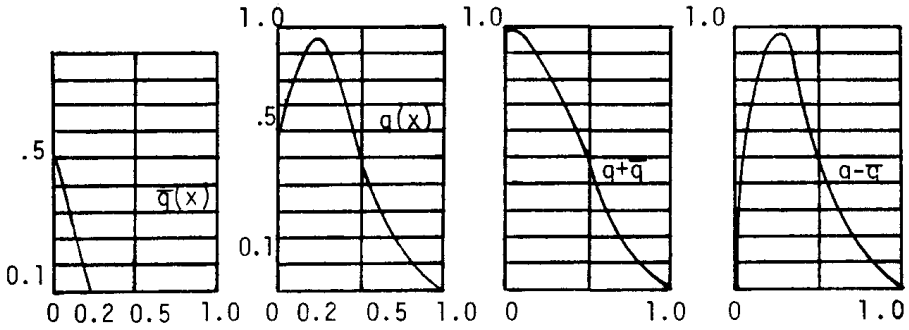
then

$$\sigma^{\nu}_{TOT} = Q + (1/3)\bar{Q}$$

and

$$\sigma^{\bar{\nu}}_{TOT} = \bar{Q} + (1/3)Q.$$

This means that the ratio $\sigma^{\bar{\nu}}_{TOT}/\sigma^{\nu}_{TOT}$ can never be less than 1/3, and it will only be 1/3 if $\bar{Q} = 0$; that is, only if the wavefunction has no anti-quark contribution. The preliminary data give $\sigma^{\bar{\nu}}_{TOT}/\sigma^{\nu}_{TOT} = 0.38 \pm 0.02$, which means $\bar{Q}/Q = 0.05 \pm 0.02$. Now in our model, $q(x)$ and $\bar{q}(x)$ are equal near $x = 0$. Experimentally, their sum is around 1, so they both start at about 0.5 at $x = 0$. The small $\bar{Q}/Q$ ratio indicates that $\bar{q}(x)$ must fall off rapidly as x increases, giving a picture like:



This means, that except for small x, we have

$$d^2\sigma^{\bar{\nu}}/dxdy = (1-y)^2 d^2\sigma^{\nu}/dxdy ,$$

with $d^2\sigma^{\nu}/dxdy$ being independent of y. Close to x = 0, on the other hand, the two cross sections become equal to about

$$0.5[1 + (1-y)^2] .$$

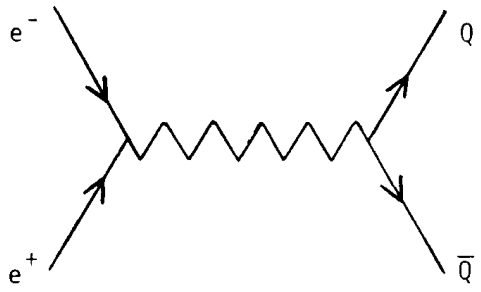
The sum of the two total cross sections is $(4/3)(Q+\bar{Q})$, and we can get a prediction for this from the electron scattering data. For we have

$$Q + \bar{Q} = (9/5) \int_0^1 f_2^{ed}(x)dx - (2/5) \int_0^1 [\Sigma(x) + \bar{\Sigma}(x)]dx .$$

Using the experimental result of about 0.31 for the integral of f_2^{ed}, then $(4/3)(Q + \bar{Q})$ is about 0.74 if we neglect the strange quark contribution. The latter should not lower this by more than 10%. The data at the present time fit this number very well. Of course, when more detailed data are available, we will be able to extract the u(x), etc., and make more detailed predictions. I have discussed here mostly results for which data are now being taken, in particular by Barish et al., at NAL. Therefore we expect soon to get more accurate data at higher energy, so that the specific forms for $d^2\sigma^{\nu}$ or $^{\bar{\nu}}/dxdy$ expected by the model that partons are quarks may be tested. There is enough redundancy to independently check scaling and assumptions about the q^2 dependence of the weak interaction. We are,

therefore, at a very intriguing moment in high-energy physics where we can do an experiment for which we have such definite predictions and are on the verge of being able to decide about a very fundamental property of strongly interacting systems. We have put this property in the language "charged partons are quarks," but other theoretical interpretations of the same results, say, in terms of current commutator rules, will not alter the fact that if the predictions work, something profound has been established that must be directly incorporated into future theory.

Another experiment that poses some problems for the quark parton model measures $e^+e^- \to$ any hadrons. We would expect this process to be described by the diagram



where the quarks subsequently generate the final hadrons. But in this case, the ratio of the cross section for this process to that for $e^+e^- \to \mu^+\mu^-$ should be equal to the sum of the squares of the

quark charges; namely, 4/9 + 1/9 + 1/9 = 2/3. The data indicate a value greater than 2. Of course, if there were three "colors" of quarks, the prediction would be 3x(2/3) = 2, which is closer to the experiment. There are other reasons for introducing the color quantum number, and we might as well review a few of them here. One has to do with the spin-statistics problem. Consider a Δ^{++}: in the non-relativistic quark model this is made up of three u's in a relative s-state with all their spins aligned. This is impossible for real fermions according to the spin-statistics theorem. Now maybe if quarks can't get out, the theorem doesn't apply, but another way out is to make them distinguishable by, say, coloring one of them chartreuse, one lavender, and one beige. (This is not the standard color notation.) The three values of the color quantum number can be incorporated into an SU_3 symmetry. This then can be used to explain why the physical hadrons have 0 or 3 quarks, but not 1, 2, 4, ..., by assuming that we only get binding for a state that is a color singlet. The baryons will then always have one quark of each color.

LECTURE 4

HIGH-ENERGY HADRONIC COLLISIONS

In this lecture, I am going to discuss high-energy hadronic collisions. The idea of partons--the study of the wavefunction at high energy--was, as I explained, originally invented to deal with hadronic collisions. The development of the ideas was guided by some of the experimental features of hadron physics, and the net result was to predict some properties to be expected in these collisions at high enough energies. The observed fact of limited transverse momentum was used to get the idea that partons do not interact strongly when their relative rapidity is high. As we have seen in previous lectures, this led to the notion that the wavefunction was likely to scale and therefore, after the collision is over, it was a very good guess that the products would scale. So the scaling principle was established in my mind in this way, and the idea of a plateau in the rapidity plot. These two ideas are really about all that I have been able to extract as far as hadronic collisions are concerned. Nevertheless, in thinking about these things, I also considered some other physical views of high-energy collisions, which corroborated with the ideas I was

getting from the parton view, some of which I will discuss here.

For example, later on in this lecture I shall consider some suggestions about what is known as Regge behavior, by which I mean the power law fall-off with s of cross sections for processes involving an exchange of quantum numbers. I specifically exclude from that discussion processes in which no quantum numbers are exchanged; that is, those involving what is called "Pomeron exchange." These I regard as being elastic scattering plus diffraction dissociation. I want to separate the phenomena into diffraction effects and "true" inelastic processes:

$$\sigma(\text{TOTAL}) = \sigma(\text{DIFFRACTION}) + \sigma(\text{TRUE INEL.}),$$

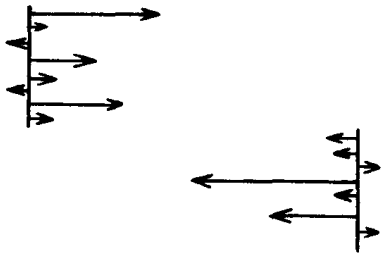
where

$$\sigma(\text{DIFFRACTION}) = \sigma(\text{ELASTIC}) + \sigma(\text{DIFF. DISS.}),$$

This is based on a picture in which an incoming wave is absorbed in some finite region. The absorption is related to the true inelastic processes, and the elastic scattering is connected to it through the optical theorem--it is a shadow effect in which the incident wave undergoes diffraction into the shadow region. If the incident wave is not simply a fundamental particle, but is structured, the various parts may be diffracted differently, leading to an excitation or breaking up of the wave. This we call diffraction dissociation, and will regard as distinct

from the fundamentally inelastic processes. I will discuss the diffraction effects in more detail in the next lecture, and give a mathematical definition of the separation of the cross section into diffractive and true inelastic parts. The ideas I am going to discuss do not take into account the possibility (suggested by experiments in the last year) that the total cross section may be rising as some power of log s; I have not modified the arguments to include this kind of behavior.

Now let's discuss the ideas that come from the parton picture of the wavefunction. The observation of finite transverse momenta has led us to a notion which has been used again and again--the idea that when the relative momentum of the partons is very large, their interaction is very small. Only finite relative rapidity and finite transverse momenta are generated by the interaction, which means the wee partons come from a long cascade and become independent, since there are so many steps in the cascade, of how it started. It seems to me that all these ideas are coming from one physical principle, which is that the couplings in the Hamiltonian are such that direct interactions among partons of high-relative momentum are very small. If this is the case, then in a very high-energy hadronic collision, pictured as below:



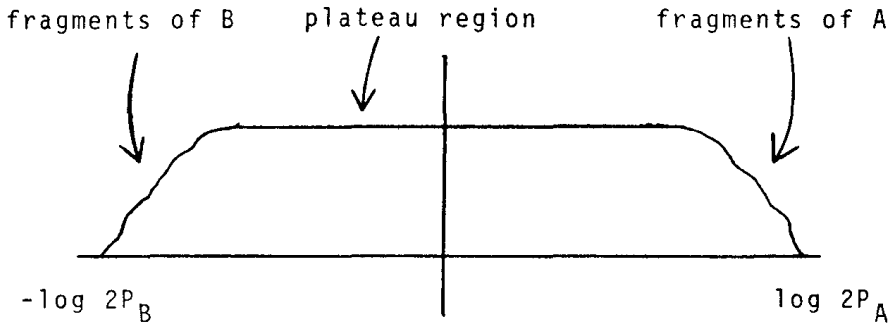
the interaction proceeds through the wee partons since only these have finite relative momenta. My picture of how the collision proceeds is that the interaction of the wee partons disturbs the phase relationships in the hadronic wavefunctions, so that they come apart in some sense. That is, after the interaction, the relative phases between the partons are no longer those of the original hadron state, so the final state is a complicated superposition of hadron configurations. Of course, the partons themselves cannot come out since they are not eigenfunctions of the Hamiltonian. The outgoing partons must distribute themselves in some way to form real hadrons. The way in which this occurs is generally local in relative momentum--not exactly, since an outgoing fast parton must use up some of the wee partons to form an outgoing hadron. But since the wee region is universal, the character of the hadrons that come out will depend on the character of the fast partons going in the same direction. This is essentially the idea of limiting fragmentation, which, when it was suggested by Yang at the Stony

Brook conference in 1969, I realized fit in with my parton ideas very well.

These ideas are nicely represented on a rapidity plot, where we measure the rapidity to the right of the origin for particles moving to the right, and to the left of the origin for those moving to the left. If we have an incoming particle A with momentum P to the right, its rapidity is

$$(1/2)\log(4P_A^2/\mu^2+Q^2) \ ,$$

which is approximately log $2P_A$ for large P_A. This will be the right end of the rapidity plot. If the other particle B has large momentum P_B to the left, the left end of the plot will be at -log $2P_B$. The idea of limiting fragmentation tells us that the distribution of products which fall near the ends of the rapidity plot will be determined by the nature of the incoming particle corresponding to that end, and will be independent of what goes on at the other end. Now in between, at finite values of y, if the idea of a dx/x distribution is correct, the distribution will be flat on either side of y = 0, and will match in the middle. The matching is guaranteed by the fact that a boost along the longitudinal direction would simply shift the plot by a finite amount. So we get two fragmentation regions separated by a flat plateau:



The total rapidity difference between the ends of the plot is

$$\log 2P_A + \log 2P_B = \log 4P_AP_B = \log s \ .$$

This picture of the breakup of the hadron wavefunction due to disturbance of the phase relationships of the wee partons is not an honest conclusion of the parton model itself. What is missing is some description of how, given the initial wavefunction represented in terms of partons, the Hamiltonian acts to produce the outgoing hadrons. You don't see the wavefunctions of all the hadrons in this picture. Once, when I was working on liquid Helium, I went to a meeting and Pauli was there. He asked me to come to his hotel room to tell him about my work. I tried to describe to him the character of the wavefunction when you change the atoms around--how the amplitude had to shift, and how, when they were close together, it had to go down, and so on. After forty minutes he said to me, "Feynman, you've been waving your arms

at me for forty minutes. Stop waving your arms and write an equation." Well in this particular field I haven't gotten to the condition where I can write an equation--I'm still waving my arms after four years--and I appreciate the lack!

To summarize, the best way to put the physical assumption is not to say how it works, whether by disturbing the phases or some other mystery, but to say that the wavefunction after interaction is a wavefunction which represents a large number of hadrons coming out which are distributed in momentum like the partons in the qualitative sense of limited transverse momentum, scaling in longitudinal momentum, and a dP_z/E behavior for small P_z (central plateau in y). And that the quantum numbers of the hadrons are determined locally by the quantum numbers of the partons, so that the character of the particles that come out in a certain direction depends only on the incoming particle that was going in that direction. Although I was not aware of it, every one of these features had already been suggested a decade ago by Kenneth G. Wilson (Acta Physica Austriaca 17, 37 (1963)).

I would like to present some of the arguments that led me to think that the wee partons were probably the only ones that interacted, which is an essential part of the idea. First, suppose we represent

a particle moving to the right by

$$|\overline{\Psi}_A\rangle = F_A^*|0\rangle \ ,$$

where F_A^* is some operator which contains creation operators only for partons moving to the right and, of course, as we expect, wee partons. But suppose for the sake of argument that the wavefunction actually turned out to have a vanishing number of wee partons as $P \to \infty$. Similarly, suppose

$$|\overline{\Psi}_B\rangle = F_B^*|0\rangle$$

is a similar particle moving to the left. The Hamiltonian is such that these are eigenfunctions, so

$$H|\overline{\Psi}_A\rangle = E_A|\Psi_A\rangle \ ,$$

$$H|\overline{\Psi}_B\rangle = E_B|\Psi_B\rangle \ .$$

That is, $|\overline{\Psi}_A\rangle$ is an eigenstate of the Hamiltonian involving only right-moving particles and $|\overline{\Psi}_B\rangle$ a solution containing only left. The Hamiltonian must be of such a kind as to permit such a solution. It is not rigorously necessary perhaps; but it is very suggestive that if such is the case, the product wavefunction, involving as it does only one or another of these kinds of particles--each in a distribution which is a solution of the Hamiltonian, would also be an eigenfunction of the Hamiltonian. To take the simplest example, of course, assume the Hamiltonian can be split into pieces H_A and H_B such that H_A doesn't contain any of the operators in F_B^*, and H_B

doesn't contain any of those in F_A^*. If we consider the state

$$F_A^* F_B^* |0\rangle \ ,$$

we have

$$H(F_A^* F_B^*)|0\rangle$$

$$= (H_A + H_B) F_A^* F_B^* |0\rangle$$

$$= F_B^*(H_A F_A^* |0\rangle) + F_A^*(H_B F_B^* |0\rangle).$$

This last follows since our assumptions mean that H_A and F_A^* commute with F_B^*, and H_B commutes with F_A^*. But then this is

$$(E_A + E_B) F_A^* F_B^* |0\rangle \ ,$$

so the wavefunction is an eigenfunction of H. This means that the state consisting of two particles moving in opposite directions is an eigenstate of the Hamiltonian--they don't interact, or better, the interaction falls to zero as s increases. This argument will fail if there are some partons common to both particles, and a constant cross section suggests that the number of these, necessarily wee partons, does not fall with P.

As an example to illustrate how the wee partons from the two hadrons mesh together, consider the model discussed in Lecture 2, where we had the wavefunction

$$|\Psi_A\rangle = \exp\left\{\sum_k c(\vec{k}) a^*(\vec{k})\right\} |0\rangle,$$

with

$$c(\vec{k}) = \alpha/(\omega - k_z)\omega^{3/2}$$

for a particle moving to the right. We know this gives a dx/x distribution, and wee partons, so we expect interaction in this case. A left-moving particle would have a wavefunction

$$|\bar{\Psi}_B> = \exp\left\{\sum_k c'(\vec{k})a^*(\vec{k})\right\}|0>,$$

with

$$c'(\vec{k}) = \alpha/(\omega + k_z)\omega^{3/2}.$$

Now suppose we look at the wavefunction that results from multiplying these operators--

$$|\bar{\Psi}_{AB}> = \exp\left\{\sum_k [c(\vec{k}) + c'(\vec{k})]\, a^*(\vec{k})\right\}|0> .$$

The coefficient is (note $\omega^2 - k_z^2 = \mu^2 + Q^2$)

$$c(\vec{k}) + c'(\vec{k}) = \alpha/\sqrt{\omega}\ (\mu^2 + Q^2),$$

where Q is the transverse momentum. We see that the distribution has a $dk_z/\omega \to dx/x$ behavior. It is also easy to check that this wavefunction is not an eigenfunction of the Hamiltonian. Now this initial distribution of partons is disturbed by future applications of the Hamiltonian in producing the time evolution to the final state.

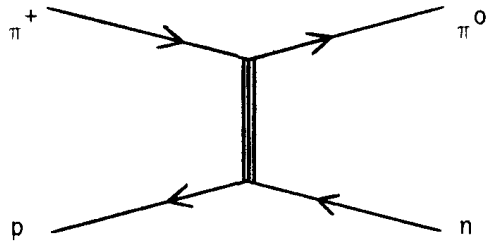
One may guess that the energy dependence, at least, of the interaction cross section is measured by the degree that these initial wavefunctions overlap; i.e., contain common partons. We may calculate the amplitude for this, $<\bar{\Psi}_B|\bar{\Psi}_A>$, which comes out as $\exp\left\{-\sum_k c'(\vec{k})c(\vec{k})\right\}$. The quantity in the exponent

becomes

$$\alpha^2 \int d^{2\vec{}}Q dk_z/(\mu^2 + Q^2)\omega^3 \ ,$$

a finite number.

To return to the subject of Regge behavior, I want to discuss some physical ideas which, although not really from the parton model, go along with it in supporting our conclusions. As we have said, the main feature we want to understand is why the cross section for processes involving the exchange of quantum numbers fall as some power of the energy. Look at a process like:

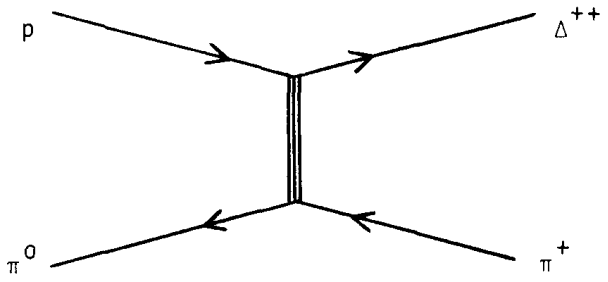


We see here that during the collision a current is suddenly reversed. The incoming neutron has 3-component of isospin -1/2, and this is flipped to give an outgoing proton with 3-component of isospin +1/2; the incoming pion has its 3-component of isospin suddenly changed from +1 to 0. We may imagine that the 3-component of isospin carried in by the π^+ is suddenly turned around during the brief time of collision. Now Gell-Mann has been emphasizing that hadronic systems are coupled to currents, and it seems very likely that the ρ, for instance, is coupled to the isospin current. So we can ask,

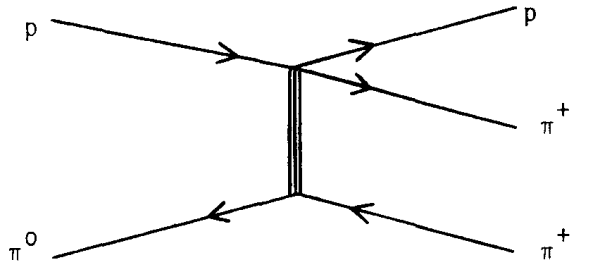
when there is such a tremendous time rate of change of a current, why don't we see radiation of the particles coupled to the current? The answer is that there is radiation--but when we are looking at a specific exclusive process such as the two-body charge-exchange one above, we are asking for that fraction of the cross section which has no such radiation. In other words, we are asking for the probability that there is a rapid acceleration of the isospin current without a corresponding radiation of the particles coupled to isospin. If you take a theory like bremsstrahlung, for example, it turns out that for a sudden reversal like this, you find more and more radiation as the energy goes up; in fact the mean number of particles radiated increases as log s. So the probability of no radiation is proportional to $\exp(-a\bar{n}) = s^{-a}$. In the theory of bremsstrahlung, this logarithmic growth of $\bar{n}$ is produced only if the particles which are going to be radiated have a dk/ω distribution. In general, we would like to say that the probability of a process which doesn't radiate any particles in a gap of rapidity Δy goes like $\exp(-a\Delta y)$. In a two-body $\rightarrow$ two-body reaction, Δy is of order log s and we get the s^{-a} behavior of Regge theory. The constant a presumably depends on the quantum numbers being exchanged, since they are related to the current that is being accelerated.

(See note at end of this lecture.)

Another point: consider the processes



and



where in the second case the $(\text{mass})^2$ of the $p\pi^+$ system might be nowhere near the resonance region. The first case is just like the one treated, and we expect a power law behavior like s^{-a}. In the second case, if we consider the p and π^+ at specific x values (so that incidentally the invariant mass squared of the outgoing $p\pi^+$ system is fixed), then in the rapidity plot we are still asking for no particles in a gap which grows as log s, and the cross section again falls as s^{-a}. The constant a will be the same in both cases, since the same quantum numbers are exchanged. More generally, the probability for any process like A + B → C + D should behave like

$$s^{-a}F(M_D^2) ,$$

for M_D^2 finite, regardless of whether D is a specific resonance or not, since finite M_D^2 means that D consists of a finite number of hadrons and there will still be a gap in the rapidity plot which grows as log s. Diffraction dissociation processes are being excluded from this discussion--when there are no quantum numbers exchanged, we have no reversal of current and no reason to expect radiation. In that case the physical situation is quite different.

We can make a multi-particle generalization of the Regge ideas. Suppose we look at a process which is partially exclusive. For example, suppose in the final state we allow any number of pions but no K's. Then by the same logic, we expect the probability of this to be $\exp(-a\bar{n}_K)$, where $\bar{n}_K$ is the mean number of K's expected to be produced in a completely inclusive collision. Physically, we expect the plateau region in rapidity to become universal as it widens, and so the mean number of any particle should grow as the width of the plateau; that is, like log s. Thus $\bar{n}_K$ goes as log s so we get a power law fall-off for the cross section in this case, too, although the powers are not the standard Regge powers.

The next case I want to consider uses the same logic but predicts something quite different. Consider the process A + B → C + "anything," and suppose

that C carries almost all the momentum of A; that is, x_C is nearly 1, but not exactly 1. In the rapidity plot, C will be at a y of order log $2P_A$. Now in the "anything" we can have some things that are also moving to the right, but they can carry a fraction of P_A which totals only $1-x_C$. So we have a gap in the rapidity plot from log $2P_A$ + log$(1-x_C)$ up to log $2P_A$, giving $\Delta y = -\log(1-x_C)$. The probability of this goes as

$$\exp(-a\Delta y) = (1-x_C)^a \ .$$

The distribution function will be the differential of this, and go as $(1-x_C)^{a-1}dx_C$. The gap carries the quantum numbers of C-A and so a can be determined from the Regge parameter α found in a two-body process with the same quantum numbers exchanged, via $a = 2(1-\alpha)$. To test this experimentally, there are some technical questions of background, particularly from diffraction dissociation, which must be taken into account, and I don't know whether it has been checked very well. Since this idea was developed, there has been a very beautiful theory by A. Mueller which makes it mathematically understandable how to connect these processes. May I refer you to an excellent report on the theory of high-energy collisions, in general, and these points in particular, by Mueller at the 1972 NAL conference (A. H. Mueller, "Production Processes at High Energy," Proceedings

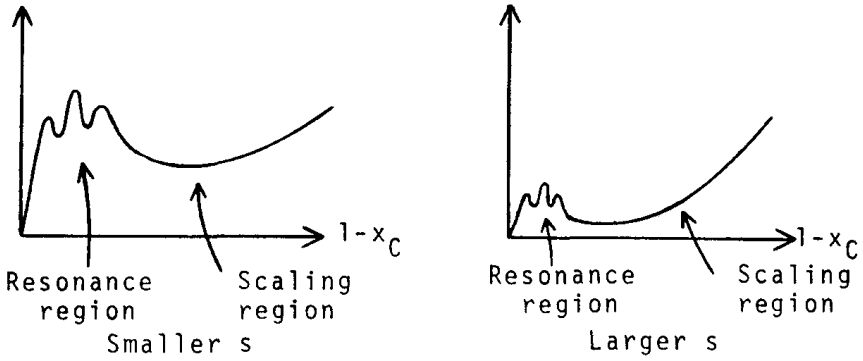
of the XVI International Conference on High-Energy Physics, Vol. 1, p. 347).

There is another interesting consequence of this, which I have developed through conversations with Bjorken, and which seems to me an important principle. (It also follows immediately in Mueller's analysis.) Consider again the reaction $A + B \to C + D$, where D now can be anything, including various resonances. Suppose the only property of D we measure is its invariant mass M_D^2. For M_D^2 finite, as we have already remarked, the probability for the reaction goes as $s^{-a}F(M_D^2)$. Now if all the momenta are large, we have

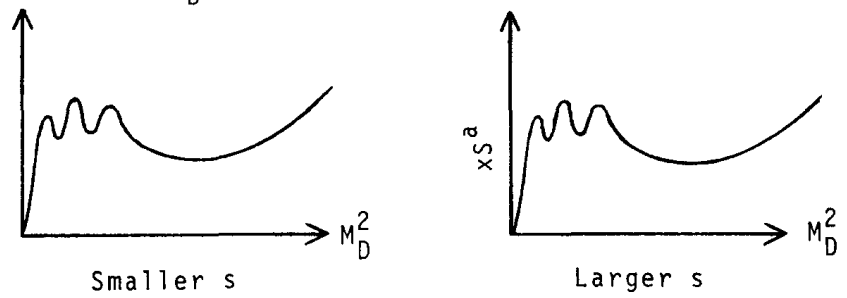
$$M_D^2 = (p_A+p_B-p_C)^2 \cong s(1-x_C) \ .$$

Thus finite M_D^2 must come from $1-x_C$ very small; that is, from x_C very close to 1. We have already argued that the probability of finding x_C near 1 goes as $(1-x_C)^a$. In order for these two distributions to fit onto one another, we see that we must have $F(M_D^2)$ go like $(M_D^2)^a$ for M_D^2 large enough to be in the transition region where the distributions join.

To see what this means, suppose we plot the invariant differential cross section, first as a function of $1-x_C$, for two different values of s. As s gets larger, the resonance region shrinks in size, because of the s^{-a} dependence, and is squeezed into a smaller region of x_C, since the transition



occurs at some $1-x_C$ proportional to M_D^2/s. The two regions will still join smoothly to one another. Now look at the picture if we plot the distribution as a function of M_D^2.



In this case the two pictures look alike, provided in the second we have scaled up the ordinate by s^a. The thing that is important about this is that you can never tell the resonances from the background. The scaling region and the resonance region join smoothly in such a way that we can't change the experimental conditions--say by going to higher energy--to isolate the resonances so that they become cleaner. There is a kind of "complementarity," if you will, between resonances and the background due

to scaling or something else. (This complementarity doesn't work in the case of diffraction dissociation, I believe, and I'll discuss it in the next lecture.) The Drell-Yan relationship comes from the same principle applied to compare the deep-inelastic scattering and the deep-elastic or deep-resonance production.

What I mean by this "complementarity" is that different physical ideas are useful in different physical regions. For low M_D^2, we have a picture in which the outgoing system is made up of resonances. As M_D^2 goes up, the picture becomes more complicated--we get more and more resonances and they begin to overlap, so the resonance picture becomes hard to use (but presumably not incorrect in principle). But in this region another physical picture can replace it--the ideas of scaling that come from a bremsstrahlung type of theory. The pictures are completely different, but nature doesn't care how you think about her. She fits together so you'll never make up your mind precisely where one picture stops and the other begins.

Note relative to Regge behavior:

We have suggested that in a two-body exchange collision, the perturbation expression for the amplitude for the exchange of a particle carrying quantum numbers may be altered by an additional

factor s^{-a} because of the requirement that the rapidly reversed current not radiate. The Regge theory coupled with the idea that trajectories seem to have a common slope α' suggests that this additional power is the difference of α at $t = 0$, $\alpha(0)$, and α at the mass of the perturbation particle exchanged, $t = \mu^2$. Thus $c = \alpha'\mu^2$. Hence for π exchange where μ^2 is very small, there is very little correction (from the perturbation theory amplitude $1/s$), but for ρ exchange the correction is nearly $1/2$. Yet the quantum number exchange, say, the 3-component of isospin, is the same. Can the radiation theory understand at least qualitatively (it does not quantitatively) that the exchange of a light mass makes a smaller correction for lack of radiation than does the exchange of a larger mass?

If you calculate by bremsstrahlung theory the high-energy behavior of the amplitude to exchange a particle of mass μ without radiating a vector particle of mass m, the result is, for a collision at impact parameter b,

$$\text{amplitude} = K_0(\mu b)\exp\left\{-g^2(\log s)K_0(mb)\right\},$$

where g^2 is proportional to the square of the coupling constant of the radiation, and $K_0(mb)$ is the Bessel function (two-dimensional projection of the Yukawa function $\exp(-mr)/r$)) which falls for large b as $\exp(-mb)$. Now consider, for example,

scattering forward, the integral over b of the above amplitude. For small μ (say $\mu << m$), large b is important where $K_0(mb)$ in the exponent is small and the effect of the radiation is small. On the other hand, if μ is not small, the range of b is limited to regions where $K_0(mb)$ is appreciable and the correction is larger. Thus we obtain the correct qualitative result.

Is there a suggestion here that data for negative t at least to be analyzed empirically, not by assuming an inverse power law in s for each momentum transfer Q, $Q^2 = -t$, but rather an inverse power of s for each impact parameter b, ultimately Fourier transformed by $\exp(i\vec{Q}\cdot\vec{b})$ to get the Q dependence?

LECTURE 5

DIFFRACTION DISSOCIATION AND SOME PROBLEMS FOR THEORISTS

In the last lecture, I discussed the inelastic hadronic collisions from the point of view of the parton model, but I pointed out that I was separating the total cross section for all events into three kinds: elastic events, those inelastic events associated with the elastic which I called diffraction dissociation, and the true inelastic events. The previous lecture dealt only with the last of these. If we picture a collision in terms of impact parameter, a target has a certain effective size for absorbing the incoming wave. At large impact parameters, the incoming wave passes by undisturbed, but in the region behind the target, it will be altered in intensity and possibly in phase--a shadow is formed. We all know that as a result of this, the shadow diffracts. If the incoming wave were a simple particle, the diffraction would just produce elastic scattering, which could be related to the inelastic effects producing the shadow through the optical theorem. Now, as Good and Walker pointed out, as is generally true in other areas of physics such as atomic physics, if the incoming wave describes a system made up of parts, then there are

two ways in which the picture is altered. First, the amplitude for absorption by the target may be different for the various parts, so that what is behind the target is a distorted wavefunction with a different proportion of these parts. And second, even if all the parts are completely absorbed, the diffusion into the shadow from the edge will differ for parts of different masses, and so on. As an example from atomic physics, the diffraction of a hydrogen atom will result in a relative distortion of the proton and electron wavefunctions which can excite or ionize the atom. If this picture is right, then when the absorption is strong, most of the distortion of the wavefunction comes from regions near the edge of the target, while elastic scattering comes from the whole disc. For this reason, I expect the elastic scattering to be the biggest fraction of the diffractive effects--I will be surprised if the diffraction dissociation turns out to be bigger than 1/2 the elastic. The diffracted wave, both the elastic and diffraction dissociation parts, carries the quantum numbers of the incoming wave.

We can describe the diffractive part of the scattering in terms of a probability of exciting a state of invariant mass M, $F(M^2)dM^2$. The elastic scattering contributes a δ-function at the mass of the incoming particle. I have studied a number of

models of systems with parts which are differentially absorbed and have always discovered, to my surprise, that the probability of excitation is small and falls off very rapidly with M^2. I would guess that in hardonic collisions, $F(M^2)$ falls essentially to zero by M of the order of a few GeV. Now, if all the views that I have regarding this are right, I can define the diffraction dissociation in an interesting way. Suppose we have the cross section for finding the products of a collision with fractions of the longitudinal momentum x_1, x_2, etc., and with transverse momenta $\vec{Q}_1$, $\vec{Q}_2$, etc., at some total center-of-mass (energy)^2 s. We call this

$$\sigma(x_1,\vec{Q}_1;x_2,\vec{Q}_2;\ldots;s) = \sigma(\{x,\vec{Q}\};s),$$

where by $\{x,\vec{Q}\}$ we mean some set of x's and $\vec{Q}$'s. We will define the limit of this as $s\to\infty$ to be the diffractive cross section for this particular set of x's and $\vec{Q}$'s, and write

$$\lim_{s\to\infty}\sigma(\{x,\vec{Q}\};s) = \sigma_D(\{x,\vec{Q}\}).$$

Now at a given value of s, the total cross section is

$$\sigma(s) = \sum_{\{x,\vec{Q}\}}\sigma(\{x,\vec{Q}\};s),$$

where the sum is over all allowed sets of x's and $\vec{Q}$'s. Let's suppose that the total cross section approaches

a constant at high energy:

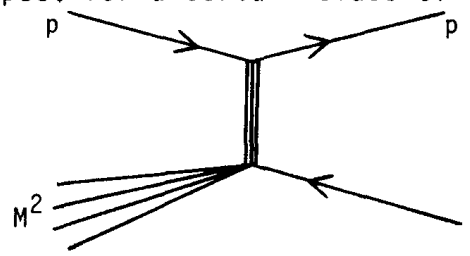
$$\lim_{s\to\infty} \sigma_{TOT}(s) = \sigma_{T\infty}.$$

We then define the diffractive part of this to be

$$\sigma_{D\infty} = \sum_{\{x,\vec{Q}\}} \sigma(\{x,\vec{Q}\}).$$

These two cross sections, $\sigma_{T\infty}$ and $\sigma_{D\infty}$, are different in general, since in the first we take the limit of a sum, and in the second, the sum of the limits. We expect that processes involving the exchange of quantum numbers will fall off as some power of 1/s, so that $\sigma_D(\{x,\vec{Q}\})$ and therefore $\sigma_{D\infty}$ will involve only processes with no quantum number exchange.

To clarify these ideas, suppose we see what we should expect for a certain class of collisions:



Here we require that one incoming proton emerge as a proton with x near 1, while the other incoming particle may break up to form a state of invariant mass M. If we considered only the true inelastic effects, we would expect a distribution that depended only on x by scaling. Since x is near 1, it is convenient to

use the variable $x' = 1-x$. The elastic scattering would contribute a $\delta(x')$, which we take out. The diffraction dissociation will add some function of M^2, which for large s and x near 1, is approximately

$$M^2 = s(1-x) = sx'.$$

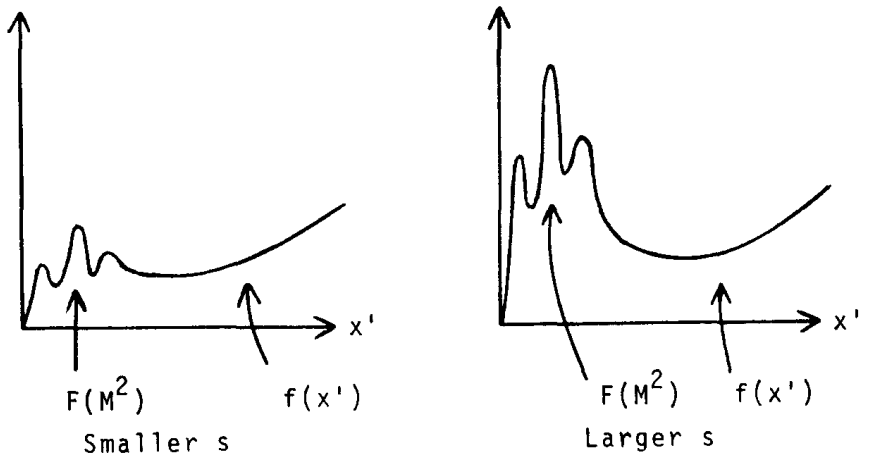
The overall distribution, ignoring the elastic part, looks like the sum

$$f(x')dx + F(M^2)dM^2.$$

If we want to plot this as a function of x', we write it as

$$[f(x') + sF(sx')]dx',$$

which gives a picture like:

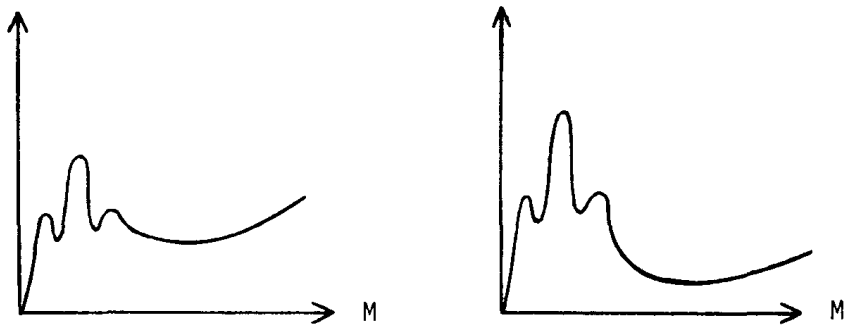


The part of this curve that is due to f(x')--the part that scales--looks the same in both pictures. The part due to $F(M^2)$ becomes higher as s increases, since we are really plotting $sF(M^2)$; and, moreover,

since a given range ΔM^2 corresponds to a range $\Delta x' = \Delta M'/s$, it is squeezed into a smaller region on the x'-axis. Now suppose we plot the situation as a function of M^2. For this we use

$$[f(M^2/s)/s + F(M^2)]dM^2,$$

and the picture looks like



Smaller s Larger s

Here we see that the background shrinks with increasing s, and so it is possible to tell diffractively produced resonances from the background of truly inelastic collisions (including pionization, for example) if my picture is right. The "complementarity" that I spoke about previously for Regge exchanges is not expected here.

If it should happen that the diffraction dissociation gives $F(M^2)dM^2$ that goes as dM^2/M^2 for large M^2, then this separation could not be made, since at high energy the distribution would become $dx/(1-x)$, which scales. But I think it is very

unlikely that the probability of excitation of high-mass states falls off so slowly. (Such a term is expected in theories which allow a triple Pomeron exchange in the Mueller analysis, but with our interpretation of "Pomeron exchange" I don't know what this could mean physically and do not expect it.)

Of course, to the diffractively produced resonances there will be a substantial diffractively produced non-resonant background, some of which is produced, for example, by the Deck effect. This effect is expected from our view for as it is usually described, say for a p-p collision, the incoming proton wavefunction does contain some amplitude to be, say, a neutron and a π^+ (two bunches of correlated partons, if you will). The π^+ is scattered by the other proton so the $n\pi^+$ system comes apart. This makes a strong non-resonant effect, which falls off rapidly enough with M^2 (like $dM^2/(M^2)^3$, I believe) that it cannot be confused with a scaling background.

I should like to summarize this part of my lecture with a list of good theoretical questions for theorists to worry about concerning hadronic collisions.

--First, how do the parton-parton interactions operate to cause the breakup of the hadronic wavefunction? I have discussed the idea that the wee partons interact and disturb the phase relationships, but

we need some mathematical example to see how the mechanism works.

--There are all sorts of questions about transverse momentum. Why are the transverse momenta limited, and how are they correlated? The wavefunction must contain this sort of information.

--Theorists should try to understand how the cross section at high energy really behaves, instead of just saying it goes to a constant because it looks constant, or goes as a logarithm because it looks that way.

--We need to understand the mechanisms which produce large transverse momentum. The work by Kislinger reported at this conference, and work of Bjorken, are suggestions in this direction. If we understood the mechanism, then the large transverse momentum data would give us more information about the parton distributions.

--How can we connect the distribution functions found in the deep-inelastic scattering experiments with the data on the products in high-energy hadronic collisions? We have data on the percentages of K's and π's produced, and so on, but as yet no theory to connect this data with what we already know about the wavefunction, or to enable us to use the data to learn more about the wavefunction.

--Finally, an interesting exercise for theorists. I

have frequently mentioned bremsstrahlung in these lectures; my analogies were with the electrical case in which the charge that is accelerated radiates only particles that have no charge. The behavior of bremsstrahlung in a Yang-Mills theory, where the radiated particles carry the charges whose acceleration gives rise to the radiation, is an interesting technical question. The paradoxical feature is that if we consider a charged object being accelerated, and its radiation carries off charge, then the charge itself may not have been accelerated, in which case there would have been no radiation...?

Now I should like to go on to another subject. We have emphasized that experiments may very shortly tell us whether partons are quarks. Let us now consider the set of questions which arise if it should turn out that these experiments do support the hypothesis that partons are quarks.

--What kinds of gluons are there? We have good reason to believe there are neutral constituents, and it is reasonable to associate these with the particles that mediate quark-quark interactions. But what are their spins--are they vector particles, for instance--and other quantum numbers? How many kinds of gluons are there? And so forth.

--How can we relate the quark-parton picture to the low energy quark model? The study of low-energy

resonances gave us a picture of baryons made of three quarks, which fits the evidence semi-quantitatively very well. Now we have a wavefunction in a rapidly moving coordinate system, and we seem to have the original quarks along with additional quark-anti-quark pairs. We can't connect the pictures very well because we have to know the Hamiltonian to transform from one to the other.

--Why do quarks, which have spin 1/2, seem to obey Bose statistics? Why are hadrons states of zero triality? These questions could be connected if there are really three colors of quarks, and the physical baryons are color singlets, but we need to fill out such a picture with some understanding of the saturation of the quark-quark forces.

--How can we explain the absence of free quarks? In particular, what is the mechanism by which the outgoing quarks in a deep-inelastic collision "decay" into hadrons in such a way that quark quantum numbers don't get out?

You won't be surprised to hear that I don't know the answers to these questions. However, I've considered a number of them, and I'd like to outline some of the ideas. They are not a solution to anything--they just represent the struggle to understand which problems are more serious and what we

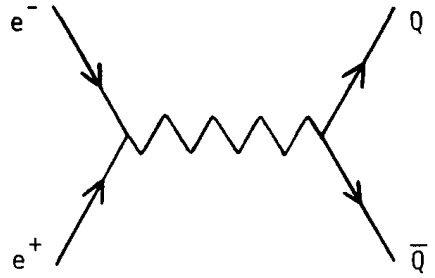
might be able to do about them.

The first consideration I want to describe relates to the harmonic forces often invoked for the low-energy quark model. Suppose we take the idea more seriously--what are the advantages, and what problems arise? One advantage would be that the quarks could not get out and we wouldn't see free quarks. This assumption works fine in the non-relativisitic model, of course, and also poses no difficulty for the arguments we have been making in the parton model. This is because the harmonic potential is soft for short distances; so when we calculate the probability of a parton going off in some direction as the result of a collision, we can neglect the binding to the other partons. Only when it gets far away, long after the collision has taken place, does it know that it is going to be pulled back. As long as the level spacing of the harmonic oscillator is small compared to the q^2 and ν of the collision, all our parton calculations go through. There will be differences in the predictions of the products of a collision if there are harmonic forces, however, since in this case there are strong couplings of the outgoing parton to the others.

Presumably, we have to have some sort of saturation of the harmonic forces so that two objects, each made of three quarks, do not have a strong long-range

interaction. This is analogous to the case of two electrically neutral atoms, but the analogy points up a problem. Even if the forces are originally zero because of saturation, a force could arise if the objects were polarized in some sense, and induced polarizations could be expected to occur, giving rise to a long-range force analogous to the van der Waals force between electrically neutral molecules. The rapid rise with distance of a truly harmonic force could make this a large effect, and this is a major disadvantage to be considered in model-building with such forces.

There is another consequence that leads to a new idea. Suppose we imagine a process like



and consider the outgoing quark and anti-quark moving farther and farther apart. As they do this, the harmonic force between them implies that a tremendous field will develop between them; and when it gets large enough, quarks and anti-quarks will be produced in pairs by the field all over the place. These will then settle down into their stable saturated systems--which are the hadrons--producing

hadron patterns like those we see. In this view, we see that the harmonic force itself doesn't look harmonic at large distances.

I do not consider the harmonic force as a particularly good idea. But I do think this idea that it has led us to may be of real utility; namely, when a single quark is moving alone away from others (say, backwards in the deep-inelastic electron scattering), it produces large numbers of pairs of quarks roughly evenly distributed in rapidity (note I have artfully changed the original space distribution for a harmonic force, to one in rapidity, guessing that this is the proper relativistic representation) and that these manifold partons there assemble themselves into hadrons of zero triality. If this is true, many of the qualitative expectations for the distributions of hadrons in deep-inelastic collisions are valid even for partons which do not carry hadron quantum numbers. (See note at end of this lecture.)

Whether the forces are harmonic or not, we still have to understand saturation and the question of triality. I have done some thinking about this in conversations with a student, Ken Kauffmann, and I am going to outline here the ideas we have discussed. I have since learned that such ideas are quite old. For example, they are discussed by Y. Nambu (Preludes in Theoretical Physics, edited by A. de-Shalit,

H. Feshbach, and L. Van Hove, North Holland, 1966, p. 133). A detailed discussion has been recently given by Lipkin ("Triality, Exotics, and the Dynamical Basis of the Quark Model," by H. J. Lipkin, Weizmann Institute preprint WIS 73/13 Ph).

In the case of electricity we get saturation--most matter is neutral. This is easy to do with only two kinds of charge, but in the case at hand we need saturation at three. We eventually saw that to get this, we had to have something analogous to SU_3. So we suppose the quark carries a quantum number of the SU_3 type, but distinct from its usual SU_3 quantum number which distinguishes the u, d, and s from one another. The new quantum number takes on three values which we call colors, and will call here simply a, b, and c. In this model we have nine quarks, since we have up quarks of type a, type b, and type c, and similarly for down and strange quarks. We assume the quarks interact by a coupling of type $\lambda_a \cdot \lambda_b$, analogous to the $\vec{\sigma}_a \cdot \vec{\sigma}_b$ coupling of spins, where the λ's are the eight color SU_3 matrices, and the dot product indicates a sum over all eight values of the SU_3 index. We can evaluate the couplings by interpreting the interaction in terms of the exchange of an octet of objects formed in the usual way. That is, if we use the notation of the ordinary meson

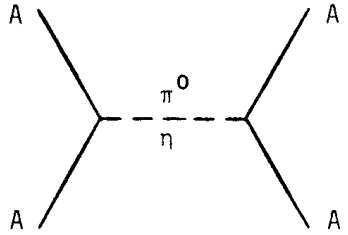
octet, we can exchange objects like

$$\eta = (1/\sqrt{6})(2\bar{C}C - \bar{A}A - \bar{B}B)$$

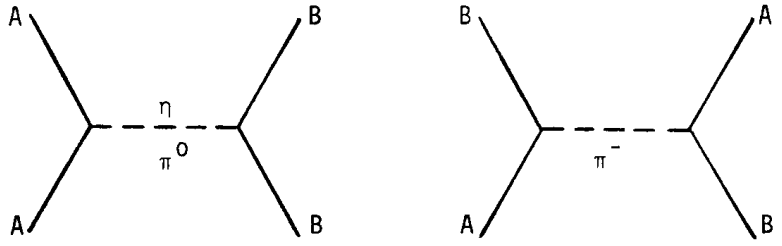
$$\pi^0 = (1/\sqrt{2})(\bar{A}A - \bar{B}B)$$

$$\pi^- = \bar{A}B$$

and so on. Consider the interaction between two A's (two quarks carrying color quantum number a--we ignore all other quantum numbers), $\langle AA|\lambda_a \cdot \lambda_b|AA\rangle$. We represent this by a diagram



The exchanged object can only be η or π^0; the others don't couple to an $\bar{A}A$ vertex. The π^0 couples to each vertex with a strength $1/\sqrt{2}$, contributing 1/2 overall. The η gives $-1/\sqrt{6}$ at each vertex, or +1/6 overall. So the net coupling is 1/2 + 1/6 = 2/3. Similarly, the diagrams



give -1/2 + 1/6 = -1/3, and 1, respectively. The coupling rules can be easily summarized by noting that

$$\lambda_a \cdot \lambda_b = -(1/3) + P_{ab},$$

where P_{ab} is the color exchange operator. (This is the analogue of Dirac's famous expression for $\vec{\sigma}_a \cdot \vec{\sigma}_b$ in terms of the spin exchange operator.) For example,

$$\lambda_a \cdot \lambda_b |AA\rangle = -(1/3)|AA\rangle + |AA\rangle ,$$

so

$$\langle AA|\lambda_a \cdot \lambda_b|AA\rangle = 2/3 .$$

Also,

$$\lambda_a \cdot \lambda_b |AB\rangle = -(1/3)|AB\rangle + |BA\rangle ,$$

so

$$\langle AB|\lambda_a \cdot \lambda_b|AB\rangle = -1/3$$

and

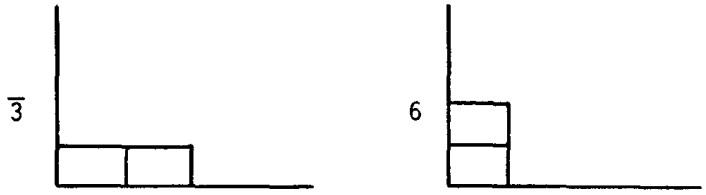
$$\langle BA|\lambda_a \cdot \lambda_b|AB\rangle = 1 ,$$

as we calculated above. This makes the calculations very easy. For any system, add one for every pair symmetric under interchange, -1 for every anti-symmetric pair, and include -1/3 for every pair. When dealing with anti-quarks, we have to reverse the sign of the couplings, so that the $\bar{A}A$ system couples with -2/3 and the $\bar{B}A$ with +2/3. The $\bar{A}A$ system has amplitude -1 to turn into $\bar{B}B$.

Now suppose we look at the coupling for a quark-anti-quark system (meson) which is a color singlet. This will have equal amplitudes to be $\bar{A}A$,

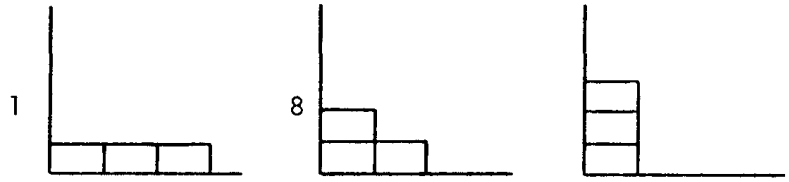
$\bar{B}B$, or $\bar{C}C$, giving a total coupling of (1/3)[3(-2/3 + 6(-1)] = -8/3. We will associate the minus sign with binding, and a positive sign with repulsive forces. Color octet states like BA have coupling +1/3 and are not bound. Thus we get the desired result that, for mesons at least, the bound states are color singlets.

Let's look at the states involving only quarks. We can classify these into color SU_3 multiplets, and use Young diagrams to represent the symmetries. In these diagrams, pairs in the same row are anti-symmetric under interchange, and those in the same column are symmetric. Consider first the two-quark system. The quarks are a color triplet, and 3 x 3 = $\bar{3}$ + 6, so the multiplets we can get are represented by



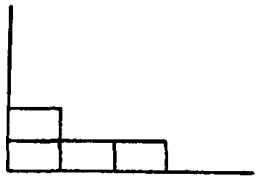
For the $\bar{3}$, we have 1 anti-symmetric pair, which contributes -1, so the net coupling is -1-(1/3) = -4/3. This is bound. The 6, however, has a symmetric pair, so the coupling is 1-(1/3) = 2/3 and it is not bound.

The three-quark states can be color singlets, octets, or decimets, with Young diagrams:



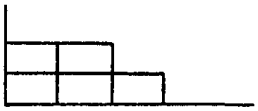
Each of these has three ways to choose a pair, so we get a constant contribution of -1 to the coupling. The singlet has all three pairs anti-symmetric, which adds -3 to give a net coupling of -4. This is strongly bound, and we see the two-quark $\bar{3}$ state is not saturated. It will tend to pick up a third quark and form a color singlet. The octet state, with one symmetric and anti-symmetric pair, has coupling -1, so the forces are attractive. But it is unstable against breakup into the two-quark $\bar{3}$ state with binding -4/3, plus a single quark, so it is not bound. The decimet has coupling +3-1 = +2 and is also not bound.

An SU_3 Young diagram has rows of maximum length three. Clearly, the maximum binding is attained when we make the rows as long as possible (to get as many anti-symmetric pairs as possible) and the columns as short as possible (to get as few symmetric pairs as possible). For example, the four-quark state with the maximum binding has Young diagram



All four-quark states have 6 pairs, contributing -2 to the coupling, and this case has 3 anti-symmetric pairs and 1 symmetric pair. The overall coupling, -2-3+1 = -4, indicates this state is neutral with respect to

Similarly, the maximally bound 5-quark state is



There are 10 pairs, of which 3+1 = 4 are anti-symmetric, and 1+1 = 2 are symmetric, so the overall coupling is -(10/3)-4+2 = -16/3. This is neutral with respect to

The state with 3k quarks has 3k(3k-1)/2 pairs. The system with the greatest binding will have a Young diagram with k rows of length three, and so will have 3k anti-symmetric pairs. Each column has k(k-1)/2 pairs, so we have 3k(k-1)/2 symmetric pairs. The net coupling is

$$-k(3k-1)/2 - 3k + 3k(k-1)/2 = -4k.$$

But this system is neutral with respect to k three-quark color singlet states (), each with coupling -4. We see that any time the anti-symmetric trio  can form it becomes neutral in its interaction

with other quarks and falls away as a baryon. If you wish all these bound states of zero triality to be near the same energy (say zero) compared to all others of very large positive energy, just add an energy of +4/3 for each quark and each anti-quark. So our scheme provides a saturation with an anti-symmetric three-quark system (which also solves the Bose quark problem) for baryons, and a color singlet state for mesons, with all other states unsaturated. We shall have to assume a large long-range force to ensure that this saturation is dynamically maintained.

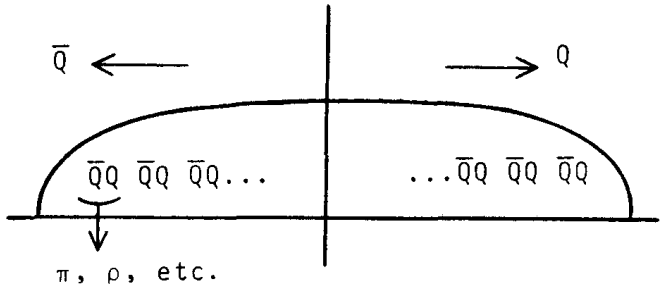
If we represent these forces in the conventional way by the exchange of vector gluons (which gives the right signs for particles and anti-particles), we have a theory with 17 particles, so it looks as if we are off again into a world of many particles, and the next generation will be working on the constituents of these! If we were any good at field theory, we could work out the consequences of the model and see if it is right or wrong. It is very amusing that we might have the correct theory and not know that we've got it--which is a funny condition to be in. However, I would expect that such an ordinary field theory will not give the long-range character to the interaction that we have had to postulate as essential to make this model keep the quarks from coming apart. Nevertheless, these leads, suggested

by a number of physicists, should be pursued more thoroughly.

Note regarding products in deep-inelastic scattering.

I had originally meant to discuss the product hadrons expected from deep-inelastic lepton-proton scattering. However, there was not enough time, and my views are all published in Photon Hadron Interactions (W. A. Benjamin Co., 1972, pp. 250 ff), with, however, one important change. On pages 260 and 261, it is argued that in a situation with a single quark moving to the right, say, the right moving hadron quantum numbers (such as charge, for example) would, on the average, be those of the quark. That this was not necessarily true was pointed out by Farrar and Rosner (Phys. Rev. D7, 2747 (1973)), and I would like to take this opportunity to state publicly that I believe they are correct and that observation of the total quantum numbers of the hadrons moving to the right need not directly give those of the quark which generated them. What is involved may be most easily seen by considering the special case e^+e^- producing a pair of quarks, say a quark Q to the right and an anti-quark Q to the left in a rapidity plot (see figure). Pairs of quarks are formed and gathered into hadrons according to our ideas. The triality (net quark

number modulo 3) of the quark to the right is +1, whereas no matter how many hadrons are formed to the right, the left over triality is still +1.



It is true that the plateau is neutral in the sense, for example, that the mean numbers of K^+ and K^- are equal (so that charge may be conserved as the plateau of length log s widens). Nevertheless, this plateau may be polarized in the sense that there is a correlation, say, for example, that whenever a K^- is found it is more likely that a nearby K^+ is to the left of the K^- than it is to the right. That such a correlation with a sign is possible is because the plateau has a sign, it is carrying +1 triality from right to left. That is, to put it invariantly, of a pair K^+K^-, the K^- is more likely to be found near the quark end. For an anti-quark to the right this would be reversed, of course, with K^+ to the right of K^-. Thus for our case, anti-quark to the left, on the left side of the plateau we have K^+ to the left of K^-. But this is just what we found for the right side of the plateau. Thus the two parts, left and

right, of the plateau fit together perfectly and the polarization continues uniformly across the entire plateau. This polarization, of course, carries quantum numbers (in our example, strangeness) so that the total quantum number of the right movers with rapidity above a certain point (say 0 in the center-of-mass) need not equal that of the quark originally moving in that direction. Isospin symmetry permits the original rule to work for isospin, but it need not work in general for strangeness, and, hence, charge (unless exact SU_3 were assumed, an unlikely hypothesis for these products). These considerations do not affect the expectations for hadron-hadron collisions, for there we have no necessary quantum number passed by the plateau and it is unpolarized as well as neutral.

Note added in proof.

Having only half a lecture on the theoretical questions, I see I didn't bring my discussion to a focus - so I would like to add these supplementary remarks. If we bring all these ideas together, we can make a "try out for size" model which seems qualitatively, at least, capable of controlling all the paradoxical aspects of the quark model. The choice of character, u, d, s, is one that any quark can have, but in addition to that we have a new perfect

SU_3 of color. Quarks are spin 1/2 fermions of three colors with perfect symmetry. The interaction is via eight vector meson gluons coupled symmetrically as in the theory of Yang and Mills, to eight currents corresponding to group generators, so that the SU_3 color character is not broken in any way. A ninth gluon coupled to "color singlet" is explicitly assumed not to be present (in this trial model). But, unlike the conventional Yang-Mills or other field theory, suppose the force is very long range, a potential energy rising without limit with distance.

How this long range can come about we do not know, and this is the most serious question in our model - but for definiteness, take the suggestions (made to me by Kenneth Kauffmann) that the propagator of the fields is $1/k^4$ instead of $1/k^2$ (and disregard "ghost" problems for now). The potential from a fixed point-charge is then r instead of 1/r.

This force between colored objects, with potential ever rising with distance, means that such objects can never come apart. The only objects which can come apart are neutral with respect to all eight currents and are therefore color singlets. These are states with the quantum numbers of three quarks of colors, a, b, c, in an anti-symmetric state in baryons, or quark-antiquark in $1/\sqrt{3}$ ($\bar{a}a + \bar{b}b + \bar{c}c$) in mesons. The gluons themselves also cannot come

far out of a system, of course, because they themselves are colored. (Murray Gell-Mann has pointed out to me that this probably removes my worry, expressed on page 82, Lecture 5, that there might be large polarization forces - the effects are second order in the interaction, an exchange of two gluons which interact may be the exchange of a massive object, and hence of limited range.)

Sometimes one attempts energetically, to separate color as in the $e^+e^- \to Q\bar{Q}$, perhaps an a to right and $\bar{a}$ to left, or in the deep-inelastic scattering in which one quark is taken from the others and sent violently backward (I am investigating whether high-energy hadron collisions may also involve this effect). In such cases, as the colored parts separate, a force develops between them as a result of gluon exchange, and in this force field pairs of new quarks are created until ultimately they are gathered into singlets which can separate indefinitely, producing the characteristic plateau in rapidity, etc., that we expect to get.

The fact that the non-relativistic quark model works roughly, and the confirming fact that there is little momentum in quark pairs in the parton distribution of a parton, indicates that forces are soft at short distances. It is not obvious that this comes out automatically from our model and something

else may have to be added. It is amusing, however, that Kauffmann's propagator $1/k^4$ is weak at large k^2 (i.e., the potential r is not as singular as 1/r for small r) and might account for the effect.

I should like to recommend a very interesting paper by Casher, Kogut and Susskind (Phys. Rev. Letters 31, 792 (1973)) that I have just seen. They do a problem of fermions in quantum electrodynamics, but in one space dimension. In one dimension the potential from a charge is $|x|$ and automatically rises with distance. It is confirmed that the phenomena of the type we expect occur. One should read this paper for a very good detailed physical description of this, following ideas of Bjorken. It is very good to have a mathematical model of these ideas, to make them clear and definite so one can advance to further ideas. By choosing the mass of the fermions to be zero, they solve the problems exactly (after Schwinger). The case that the fermion mass is not zero cannot be solved exactly, but we expect, of course, the same type of phenomena of production of fermion pairs, since the potential rises with distance. A study of such cases should be pursued to teach oneself the physics of these phenomena. A next step, after the case of non-zero fermion mass, would be an SU_3 system, Yang-Mills, in one dimension where baryons (three-quark states), and mesons would

form. Next, attempt to jump to three-space dimensions with some guess as to how the long-range force works. You may not end up with exactly the right theory, but you have a definite program to develop from which you must learn a very great deal about the real world.

Gauge Theories

R.P. Feynman

Contents

R. Balian and C.H. Llewellyn Smith, eds., Les Houches, Session XXIX, 1976 – Interactions électromagnétiques et faibles à haute énergie/Weak and electromagnetic interactions at high energy

1. Introduction

In 1954 Yang and Mills [1] wrote a paper in which they made a theory in analogy with QED for a system in which a particle could carry more than one "charge". In those days there were particles like protons and neutrons which were thought to be one object – a nucleon which appears in two guises – proton and neutron. This is the theory of isospin in which the analogue of charge is I_z, the z-component of isospin; $+\frac{1}{2}$ for proton, $-\frac{1}{2}$ for neutron. Then there was the question of what field, analogous to the photon in QED, could interact with such a charge. In the case of QED the photon is a vector field; therefore Yang and Mills tried to make a theory of a vector field interacting with "charges" that might have more than one value. There had already existed for some time a theory of (pseudo) scalar fields which could interact with particles of different "charge" but their properties weren't as interesting as those of vector fields. Since the "charge" can be flipped back and forth, the fields which are coupled to the "charges" are more interesting than that of QED. Finally the theory had very great beauty and simplicity.

We human beings see in a symmetrical theory a certain beauty; the Greeks, for example, saw in the theory of the planets that they went around in circles at a uniform speed, a phenomenon which today, we would characterise by a group theoretic property: the orbit is such that a displacement in time is equivalent to a rotation.

Today we still have this desire to see symmetrical things and therefore the Y–M theory looks very good. In contrast to QED, here we have a field with many components which couple to different "charges". This is because the field, in addition to the neutral component which couples to + and – charge, also has charged components which flip one "charge" into another. Therefore in the case of isospin we have a source of isospin $\frac{1}{2}$ (the nucleon) and a field of isospin 1. Now this theory was beautifully symmetric but it did not agree with experiment; although isospin is almost exactly conserved the theory is similar to electrodynamics in that the mass of the vector field is zero, but there is no obvious long range force between nucleons: so it's wrong. The first hope

was to put a mass in somehow, but that destroyed the symmetry and consequently the beauty. What were the vector particles anyhow? They were supposed to be ρ mesons, but there seemed to be nothing more fundamental about the ρ meson than all the other hadrons. So the idea that one of these hadrons was the fundamental field was lost.

Later people such as Goldstone [2] began to look at broken symmetries; Higgs [3,4] and Kibble [5] found that a massless vector theory with broken symmetry was in some sense equivalent to one with mass. Today the only symmetries we see in nature are isospin (a near perfect symmetry) and SU(3) (a clear, but imperfect symmetry) and we would, therefore, think that, because we like symmetries so much, the excitement of the day would be that we had an understanding of these symmetries at last. We do not. In fact, the Y–M theory with broken symmetry is assumed to apply somewhere else.

In the meantime, there was developed a weak interaction theory in which, in one interpretation, one had vector mesons with mass. Taken directly, a field theory of massive vector mesons is highly non-renormalisable. Such a theory works fine as long as we only work with first order diagrams. However, attempts to go to higher order lead to unremovable divergences. It is necessary to go to higher order for two reasons:

(1) It is not sensible to have a theory which only works to first order.

(2) Consider a process where the amplitude is calculated to first order in g. The probability of the process occuring is $O(g^2)$. Therefore the probability of non-occurrence is $\sim 1 - O(g^2)$. Hence the amplitude for non-occurrence $\sim\sqrt{1 - O(g^2)} \sim 1 - O(g^2)$. Therefore one must know something about amplitudes to order g^2. It is, therefore impossible to have a theory which only works to first order if one wants to conserve probability. People did not worry about this until recently. When they did, they found that if they started with one of these symmetric theories and used the Higgs mechanism to add mass they might be able to represent these vector mesons (intermediate vector bosons) with mass in a way that was renormalisable in the same sense as QED. This was subsequently proved in detail, and the consequences of these theories are, therefore, as calculable as those of QED.

Sunsequently it was found by Llewellyn Smith [6], that if one starts with massive vector mesons and requires renormalisability one is driven to Y–M theories with broken symmetries, provided one is prepared to introduce new particles to cancel divergences. One has a choice of such particles and a particular choice leads to a particular model. It is important to realise that there is no unique prescription for doing this. One can also use different symmetry groups and different methods of breaking the symmetry. It is therefore, not true to say that these theories make an unambiguous prediction of the exis-

tence of neutral currents, because one can always take a different theory which has no neutral current but some other new particle(s) (e.g. "heavy" lepton). It appears that there is now experimental evidence for both neutral currents and heavy leptons.

In hadron physics the quark theory was evolved and found to be paradoxical. Protons are supposed to be made out of three quarks as is the Δ^{++} which has spin $\frac{3}{2}$. Consider the case where $J_z = +\frac{3}{2}$:

u u u

↑ ↑ ↑ .

The dynamical theory says that the quarks are in a relative S-state in order to get the right order of magnitude for matrix elements and magnetic moments; then we have three particles in the same state. There is a problem in that it had been proved that we can't put three spin $\frac{1}{2}$ particles in the same state. This proof assumed that the particles could be separated from each other. We also know that quarks don't seem to appear as free particles. It is, therefore, not clear that the proof holds for quarks. Nevertheless, to be conservative, we will accept the theorem, and so a simple explanation of the problem is that the three quarks are different, i.e. we assign them a new quantum number (colour) which takes three values A, B, C:

u_A u_B u_C

↑ ↑ ↑ .

There are, at present, only two places in the experimental world where the colour hypothesis can be checked:

(1) A subtle test is connected with the anomaly in the $\pi^0 \to 2\gamma$ decay. It turns out that a theory without colour gives a decay rate a factor of three too small; a theory with three colours agrees with experiment.

(2) The ratio

$$R = \frac{\sigma(e\bar{e} \to \text{hadrons})}{\sigma(e\bar{e} \to \mu\bar{\mu})},$$

is equal to the sum of the squares of the quark charges and equals 2/3 for u, d, s without colour. In fact, the data shown symbolically* in fig. 1.1 are in disagreement with this value. u, d, s with colour give $R = 2$ in better agreement below 4 GeV.

Another possible experimental test of colour is lepton production in pp collisions by the Drell–Yan mechanism but the evidence is inconclusive. The amount of experimental evidence for the beautiful symmetry of colour is not

* See the lectures by G. Wolff and B.H. Wiik for a discussion of the data.

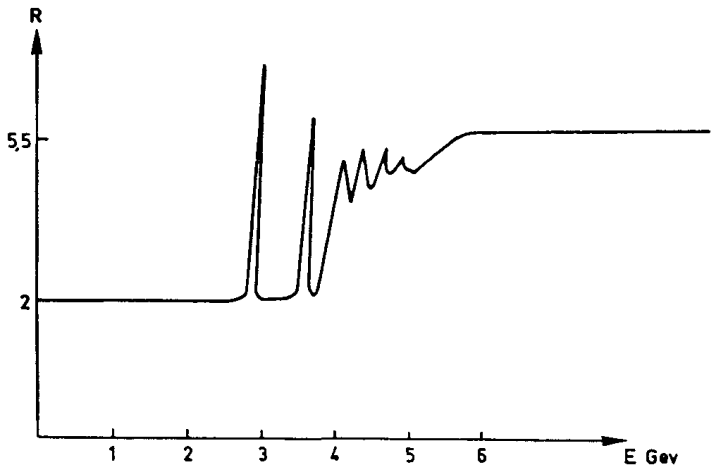


Fig. 1.1.

great but is theoretically very strong because it helps to explain a number of other features e.e. why three quarks in a baryon held together.

This theory of colour is so symmetrical that a good guess is that hadrons are made of quarks with flavours u, d, s, c, ... and three colours A, B, C with exact colour symmetry. It turns out that the colour couples to an eight component field. The reason why there are eight components is as follows:

When a field quantum (gluon) is emitted, the quark colour may or may not change. There are nine ways of coupling a gluon between an initial (three colour possibilities) and a final (three colour possibilities) quark.

$A \to B$

$A \to C$ $\quad A \to A$

$B \to A$ $\quad B \to B$

$B \to C$ $\quad C \to C$.

$C \to A$

$C \to B$

But the linear combination of the components of the vector field which couples equally to all the quarks (the singlet) need not be in the theory, all its properties are independent of the other eight components. Under a linear transformation of the colours, the eight mix together so all are necessary, but the singlet stays unchanged. Hence we are left with eight components. (Whether we add the ninth or not, and with what coupling, is up to us, but we will leave

it out as apparently unnecessary at present.) This theory with exact SU(3) colour symmetry is called quantum chromodynamics (QCD).

At first sight we have a problem, viz., massless vector mesons would imply a long range force between quarks and so we could separate them. Experimentally this does not seem to be true. One solution to the problem is that QCD is wrong. Another way out is to say that we do not understand the consequences of Y–M well enough and that at large distances the forces might become large enough to confine the quarks. That is the foremost problem of QCD. Also there are infrared divergences in QCD which are more serious that in QED and the method to handle them is not yet known. In spontaneously broken Y–M theories these infrared problems are absent.

To summarize, there are two applications of Y–M theories: with broken symmetry in weak interaction (e.g. Weinberg [7] – Salam [8] model) and in strong interactions by QCD with perfect unbroken colour symmetry.

2. Classical Yang–Mills theory

We are going to take a more or less elementary and direct view of Yang–Mills theory, rather like the authors did. We start with the example of SU(2) in which we have a 2 component spinor representing proton and neutron. We start with the Lagrangian density for the free proton and neutron fields

$$\mathcal{L}_F = i\bar{\psi}_p \partial\!\!\!/ \psi_p + i\bar{\psi}_n \partial\!\!\!/ \psi_n - m_p \bar{\psi}_p \psi_p - m_n \bar{\psi}_n \psi_n , \tag{2.1}$$

where we use the notation of Bjorken and Drell [9].

In the old days people wanted to add interactions with pseudoscalar particles. Consider 3 pseudoscalar particles (e.g. the pion) with charges $0(\phi_0)$, $+1(\phi_+)$, $-1(\phi_-)$; the simplest coupling to the nucleons is

$$\mathcal{L}_I = i\{\alpha\bar{\psi}_p\phi_0\gamma_5\psi_p + \beta\bar{\psi}_n\phi_0\gamma_5\psi_n + \gamma\bar{\psi}_p\phi_+\gamma_5\psi_n + \gamma\bar{\psi}_n\phi_-\gamma_5\psi_p\}. \tag{2.2}$$

If the nucleon–nucleon force is independent of the nucleon type (and $\gamma \neq 0$), then

$$\alpha = -\beta = \gamma/\sqrt{2}. \tag{2.3}$$

A neater way to write the Lagrangian in this case is to introduce a spinor $\psi_a (a = p, n)$ and an isovector $\phi_i (i = 1, 2, 3)$ in isospin space, where

$$\phi_1 = \frac{1}{\sqrt{2}}(\phi_+ + \phi_-), \qquad \phi_2 = \frac{i}{\sqrt{2}}(\phi_+ - \phi_-), \qquad \phi_3 = \phi_0 \tag{2.4}$$

If $m_p = m_n$ the Lagrangian $\mathcal{L}_F + \mathcal{L}_I$ must be invariant under a rotation of the ψ field in isospin space, i.e.

$$\psi_p \to \cos\theta\psi_p + \sin\theta\psi_n, \qquad \psi_n \to -\sin\theta\psi_p + \cos\theta\psi_n\,, \tag{2.5}$$

provided we also rotate the ϕ field at the same time. The Lagrangian becomes

$$\mathcal{L} = i\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi + \alpha\bar{\psi}\boldsymbol{\phi}\cdot\boldsymbol{\tau}\gamma_5\psi + \text{K.E. terms for } \phi\,, \tag{2.6}$$

where

$$2\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad 2\tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad 2\tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.7}$$

are the Pauli spin matrices.

This is the famous pseudoscalar meson theory which was supposed to be the explanation of everything. When there were summer schools, professors explained that this was the key to the whole of strong interaction theory; they were going to explain scattering and everything else and it was just a matter of calculating the next order on a machine. But that failed so they keep on trying! The question now is, can we write a similar theory for a vector particle interacting with nucleons. It is easy to guess that a vector particle A_μ could be coupled the same way i.e.

$$\mathcal{L} = i\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi + \bar{\psi}(A_\mu\cdot\boldsymbol{\tau})\gamma_\mu\psi\,, \tag{2.8}$$

where A has been rescaled to absorb the coupling constant g. The γ_μ is present to contract with the space time index on the A_μ and the $\boldsymbol{\tau}$ is present to contract with the isospin index on A. We could also add derivative couplings on the form

$$\bar{\psi}(\partial_\mu\boldsymbol{\phi})\cdot\boldsymbol{\tau}\gamma_5\gamma_\mu\psi\,, \tag{2.9}$$

in the pseudoscalar case, and similarly in the vector case. However for simplicity we exclude these and other more complicated couplings. (There is no way a priori to exclude these couplings but theoretically there may be problems with renormalisability but in the old pion theory of nuclear forces they were all tried.)

Now the problem of coupling the vector field is very easy but it is only part of the real problem which is – what are the equations for the propagation of the field? Yang and Mills treated this problem in analogy with electrodynamics. In electrodynamics, the piece of the Lagrangian connected with the propagator of the vector fields is

$$F_{\mu\nu}F_{\mu\nu}\,,$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,. \tag{2.10}$$

The piece of the Lagrangian analogous to (2.8) is

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi + \bar{\psi}A_\mu\gamma_\mu\psi\,. \tag{2.11}$$

Why do we have such a funny looking business? Why not have, for example,

$$(\partial_\nu A_\mu)(\partial_\nu A_\mu)\,. \tag{2.12}$$

Electrodynamics has a property of gauge invariance; this means firstly if $\psi \to e^{-i\alpha}\psi$ and α is a constant, nothing happens to (2.11). But now suppose that the phase of the wave function is changed by different amounts at different space-time points i.e. α is a function of x. This is usually called a local gauge transformation. Now the kinetic energy term in (2.11) will change since

$$i\bar{\psi}\partial_\mu\psi \to i\bar{\psi}\partial_\mu\psi + \bar{\psi}(\partial_\mu\alpha)\psi\,. \tag{2.13}$$

One can easily make (2.11) gauge invariant by supposing that at the same time we change

$$A_\mu \to A_\mu - \partial_\mu\alpha\,. \tag{2.14}$$

Now if the theory is to be gauge invariant we cannot use (2.12) in the Lagrangian as it changes under the gauge transformation; however $F_{\mu\nu}$ is invariant so that $F_{\mu\nu}F_{\mu\nu}$ is a possible invariant contribution to the Lagrangian.

Now we can use the same trick to try to find what kind of invariant we get in this new theory with the multiple component field (isovector). Consider the transformation

$$\psi \to \exp\{-i(\boldsymbol{\alpha}\cdot\boldsymbol{\tau})\}\psi\,, \tag{2.15}$$

applied to (2.8). The transformation has to be unitary and consequently it can be written in the form (2.15). We only consider the transformation to first order in α, viz.

$$\psi \to (1 - i\boldsymbol{\alpha}\cdot\boldsymbol{\tau})\psi\,. \tag{2.16}$$

(It can be shown that if the theory is invariant under infinitesimal transformations then it is also invariant under a finite transformation, (2.15) since this can be built up from infinitesimal transformations.) Eq. (2.8) is invariant if α is not a function of space-time provided

$$A_\mu\cdot\boldsymbol{\tau} \to \exp(-i\boldsymbol{\alpha}\cdot\boldsymbol{\tau})(A_\mu\cdot\boldsymbol{\tau})\exp(i\boldsymbol{\alpha}\cdot\boldsymbol{\tau})$$

i.e.

$$A'_\mu\cdot\boldsymbol{\tau} = A_\mu\cdot\boldsymbol{\tau} - i[\boldsymbol{\alpha}\cdot\boldsymbol{\tau}, A_\mu\cdot\boldsymbol{\tau}] = A_\mu\cdot\boldsymbol{\tau} + (\boldsymbol{\alpha}\times A_\mu)\cdot\boldsymbol{\tau}$$

i.e.

$$A'_\mu = A_\mu + (\alpha \times A_\mu)\,, \tag{2.17}$$

which is an infinitesimal rotation of the vector A_μ in isospace.

So far we have considered only SU(2). In another group one should have a similar sort of thing except that the τ's would be a different set of matrices, with another set of commutation relations. In general

$$[\tau_i, \tau_j] = \mathrm{i} f_{ijk} \tau_k, \qquad (\text{e.g. for SU(2)}, f_{ijk} = \epsilon_{ijk})\,. \tag{2.18}$$

We can generalise the notion of a vector to any group $a_i = a$ where $c = a \times b$ means $c_i = f_{ijk} a_j b_k$.

Then all the equations we write are completely general, they apply not just for SU(2). Some properties of the cross product are

$$a \cdot (b \times c) = (a \times b) \cdot c, \qquad a \times a = 0, \qquad a \times b = -b \times a,$$

$$a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0 \ (\text{Jacobi identity}) \tag{2.19}$$

but note $a \times (b \times c) = b(a \cdot c) - c(a \cdot b)$ holds only for SU(2). For any group SU(n), the number of components of A_μ is $n^2 - 1$ because the transformation parameter, α can have $n^2 - 1$ components (this being the dimension of the group).

Now consider α in (2.16) to be a function of space-time. The idea is that we should be able to change the phase of the wave function arbitrarily at each space-time point and transform the vector field such that the physics does not depend on this choice. We must choose the transformation of A_μ so that (2.8) is invariant under the transformation (2.16). α is called the gauge parameter. Therefore under

$$\psi \to (1 - \mathrm{i}\alpha(x) \cdot \tau)\psi, \qquad \mathrm{i}\bar{\psi}\gamma_\mu \partial_\mu \psi \to \mathrm{i}\bar{\psi}\gamma_\mu \partial_\mu \psi + \bar{\psi}\gamma_\mu (\partial_\mu \alpha) \cdot \tau\, \psi\,. \tag{2.20}$$

So A must transform like

$$A_\mu \to A_\mu + \alpha \times A_\mu - \partial_\mu \alpha\,. \tag{2.21}$$

The next problem is to find a Lagrangian term for A_μ which is invariant when A_μ is changed in this manner. There are very beautiful and elegant ways of getting these things these days; but suppose that you were inventing it, what would you do to find an invariant form? You fiddle around. All the elegant stuff is found later; the way to learn is not to learn elegant things, it's to fiddle around blind and stupid. Later you see how it works; polish it up; remove the scaffolding and publish the result for other students to be amazed at your ingenuity.

For the moment forget the $\alpha \times A_\mu$ term in (2.21), and try to find some ex-

pression like the square of the electromagnetic field tensor, i.e.

$$(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 .$$

This transforms under (2.21) as

$$\partial_\mu A'_\nu - \partial_\nu A'_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu + \boldsymbol{\alpha} \times (\partial_\mu A_\nu - \partial_\nu A_\mu)$$

$$+ (\partial_\mu \boldsymbol{\alpha}) \times A_\nu - (\partial_\nu \boldsymbol{\alpha}) \times A_\mu . \tag{2.22}$$

If the last 2 terms were absent we would be O.K. because then

$$\partial_\mu A_\nu - \partial_\nu A_\mu ,$$

would transform as an isovector (i.e. $V \to V + \boldsymbol{\alpha} \times V$) and therefore its square would be invariant. So we must try to get rid of the last 2 terms. Notice that the gradient of $\boldsymbol{\alpha}$ is coming from the transformation of A_μ (2.21) so that if we had a term like $A_\mu \times A_\nu$ then when we transformed it we would pick up a $A_\nu \times \partial_\mu \boldsymbol{\alpha}$. Try

$$A'_\mu \times A'_\nu = A_\mu \times A_\nu + A_\mu \times (\boldsymbol{\alpha} \times A_\nu) - A_\mu \times \partial_\nu \boldsymbol{\alpha}$$

$$+ (\boldsymbol{\alpha} \times A_\mu) \times A_\nu - (\partial_\mu \boldsymbol{\alpha}) \times A_\nu .$$

But

$$(\boldsymbol{\alpha} \times A_\mu) \times A_\nu = A_\nu \times (A_\mu \times \boldsymbol{\alpha}) .$$

Add and subtract $\boldsymbol{\alpha} \times (A_\nu \times A_\mu)$ and use the Jacobi identity (eq. (2.19)) to get

$$A'_\mu \times A'_\nu = A_\mu \times A_\nu - (\partial_\mu \boldsymbol{\alpha}) \times A_\nu - A_\mu \times (\partial_\nu \boldsymbol{\alpha}) + \boldsymbol{\alpha} \times (A_\mu \times A_\nu) . \tag{2.23}$$

We can now get rid of the debris between (2.22) and (2.23) by defining

$$E_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu \times A_\nu ,$$

and so

$$E_{\mu\nu} \to E_{\mu\nu} + \boldsymbol{\alpha} \times E_{\mu\nu} . \tag{2.24}$$

Hence we can make an invariant quantity $E_{\mu\nu} \cdot E_{\mu\nu}$. Therefore we may write the Lagrangian density (in analogy with QED) as

$$\mathcal{L} = -\frac{1}{4g^2} E_{\mu\nu} \cdot E_{\mu\nu} + \mathrm{i}\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi + \bar{\psi}(A_\mu \cdot \boldsymbol{\tau})\gamma_\mu \psi . \tag{2.25}$$

Note if we rescale $A_\mu \to gA_\mu$, $\mathcal{L}$ becomes

$$\mathcal{L} = -\tfrac{1}{4} E_{\mu\nu} \cdot E_{\mu\nu} + \mathrm{i}\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi + g\bar{\psi}(A_\mu \cdot \boldsymbol{\tau})\gamma_\mu \psi ,$$

where $E_{\mu\nu}$ is now

$$E_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + gA_\mu \times A_\nu .$$

This is the form given by Abers and Lee [10].

You may ask why use $E_{\mu\nu} E_{\mu\nu}$ in the Lagrangian instead of another invariant piece? For example $\tilde{E}_{\mu\nu}$ called the dual of $E_{\mu\nu}$, defined by $\tilde{E}_{\mu\nu} = \epsilon_{\mu\nu\rho\sigma} E_{\rho\sigma}$ is such that the quantities $\tilde{E} \cdot E$ and $\tilde{E} \cdot \tilde{E}$ are also invariant. The former is a pseudoscalar and one could consider it as a possible additional term in $\mathcal{L}$. It can be shown that this term has no consequences for the equations of motion provided that when A is varied to obtain them one assumes that as usual there is no variation of A at ∞. Possible consequences of a violation of this condition will be discussed later.

Now consider the action S derived from the Lagrangian (2.25)

$$S = \int \{ -\frac{1}{4g^2} E_{\mu\nu} \cdot E_{\mu\nu} + \bar{\psi}(i\partial\!\!\!/ + A \cdot \tau)\psi - \bar{\psi} m \psi \} \, \mathrm{d}^4 x . \tag{2.26}$$

We find the equations of motion by varying the action. Varying with respect to $\bar{\psi}$ gives

$$(i\partial\!\!\!/ + A \cdot \tau)\psi = m\psi , \tag{2.27}$$

and a similar equation for $\bar{\psi}$ is obtained by varying with respect to ψ. Varying with respect to A gives

$$-\frac{1}{g^2}(\partial_\mu E_{\mu\nu} + A_\mu \times E_{\mu\nu}) = J_\nu , \tag{2.28}$$

where

$$J_\mu = \bar{\psi} \gamma_\mu \tau \cdot \psi . \tag{2.29}$$

Any vector will transform as

$$\boldsymbol{\phi} \to \boldsymbol{\phi} + \boldsymbol{\alpha} \times \boldsymbol{\phi} . \tag{2.30}$$

The derivative of a vector transforms in a more complicated way:

$$\partial_\mu \boldsymbol{\phi} \to \partial_\mu \boldsymbol{\phi} + \boldsymbol{\alpha} \times \partial_\mu \boldsymbol{\phi} + \partial_\mu \boldsymbol{\alpha} \times \boldsymbol{\phi} . \tag{2.31}$$

Therefore if $\boldsymbol{\phi}$ is a vector, its derivative is not. However we notice that

$$\partial_\mu \boldsymbol{\phi} + A \times \boldsymbol{\phi}, \tag{2.32}$$

transforms as a vector. Hence we define a covariant derivative

$$D_\mu \equiv (\partial_\mu + A_\mu \times) \tag{2.33}$$

which when operating on a vector produces a vector. We can now re-write (2.28) as

$$-\frac{1}{g^2} D_\mu E_{\mu\nu} = J_\nu \,. \tag{2.34}$$

Consider the action of the commutator $[D_\mu, D_\nu]$ on $\boldsymbol{\phi}$. We easily see that

$$D_\mu D_\nu \boldsymbol{\phi} - D_\nu D_\mu \boldsymbol{\phi} = E_{\mu\nu} \times \boldsymbol{\phi} \,. \tag{2.35}$$

This is the first time we have got this combination of A's (viz. $E_{\mu\nu}$) out in a logical way.

Since $E_{\mu\nu}$ is an isovector, we can substitute it for $\boldsymbol{\phi}$ in (2.35). Then we get, using the antisymmetry of $E_{\mu\nu}$

$$D_\mu D_\nu E_{\mu\nu} = 0 \,. \tag{2.36}$$

Comparing this with (2.34) we see that for consistency we must have

$$D_\mu J_\mu = 0 \,. \tag{2.37}$$

In other words these field equations are meaningless equations unless the current is conserved in the sense that its covariant divergence is zero. This causes a lot of complications when we go to the quantum theory. The reason is that in the quantum theory, when we calculate a diagram and so forth, some particles in the theory interact and provide a contribution to the current which is then a source which generates a new field propagating to the next interaction. We figure out how the vector fields propagate by solving the differential equations (2.34). It may not be that our source automatically satisfies (2.37), and hence eq. (2.37) does not make sense, it has no solution and we do not know how the field should propagate. Eq. (2.34) is the analogue of the Maxwell equations*

$$\partial_\mu F_{\mu\nu} = J_\nu \,. \tag{2.38}$$

The current J_μ is produced by the matter field and it is a consequence of the Dirac equation (2.27) (and its conjugate) that this current is indeed conserved and satisfies (2.37). If matter is a Dirac spinor as we have assumed, J_μ does not explicitly depend on A_μ. This is not true in general, if we define J_μ as the

* The analogue of the other 2 Maxwell equations, which are an algebraic consequence of $F_{\mu\nu}$ being a curl, is

$$D_\mu E_{\rho\sigma} + D_\rho E_{\sigma\mu} + D_\sigma E_{\mu\rho} = 0 \,.$$

It is easy to check that this is an identity satisfied by $E_{\mu\nu}$ since it is of the form $\partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu \times A_\nu$.

first variation of the matter term (in the action) with respect to A_μ. The current would have a different form in, for example, scalar electrodynamics where the quadratic term is the Lagrangian has the form

$$(\partial_\mu + A_\mu)\phi^+(\partial_\mu + A_\mu)\phi + m^2\phi^+\phi \,. \tag{2.39}$$

The current is found by differentating this with respect to A and is

$$J_\mu = \phi^+ \overleftrightarrow{\partial}_\mu \phi + 2\phi^+ A_\mu \phi \,, \tag{2.40}$$

and so in this case there is an extra non-matter term in the current. Let us write

$$E_{\mu\nu} = F_{\mu\nu} + A_\mu \times A_\nu \,, \tag{2.41}$$

where $F_{\mu\nu}$ looks like the field tensor in electrodynamics being just the curl of A_μ. Then eq. (2.34) can be written as

$$-\frac{1}{g^2}\partial_\mu F_{\mu\nu} = J_\nu + \frac{1}{g^2}\{\partial_\mu(A_\mu \times A_\nu) + A_\mu \times E_{\mu\nu}\} \,. \tag{2.42}$$

This is just like electrodynamics. Each field is produced by a source; the source is isospin density (analogous to electric charge in electrodynamics) which here is a sum of the contributions from the matter and the field itself. The disadvantage of looking at it in this way is that we have lost the gauge invariance. It's strange but it's true that the amount of isospin density in the field depends upon the gauge – it's not a gauge invariant quantity. This is analogous to the way some people like to do gravitation.

The gravitational field equations (cf. (2.34)) are

$$G_{\mu\nu} = T_{\mu\nu} \,, \tag{2.43}$$

where $T_{\mu\nu}$ is the energy-momentum tensor of the matter and $G_{\mu\nu}$ is the Einstein tensor. Algebraically

$$D_\mu G_{\mu\nu} \equiv 0 \,, \tag{2.44}$$

where D_μ is some covariant derivative, and hence we have

$$D_\mu T_{\mu\nu} = 0 \,, \tag{2.45}$$

analogous to (2.37). However, as in (2.42) we can rewrite (2.43) as

$$G'_{\mu\nu} = T_{\mu\nu} + K_{\mu\nu} \,, \tag{2.46}$$

where $G'_{\mu\nu}$ is linear in $g_{\mu\nu}$ and $K_{\mu\nu}$ contains only terms quadratic and higher in $g_{\mu\nu}$. This is now the equation for a spin 2 particle where $K_{\mu\nu}$ is the energy-momentum density in the gravitational field, and we say that the gravitational field is produced by all energy, the energy of matter and the energy of the field

itself; that's why it's non-linear. However if we make a generalised co-ordinate transformation, $T_{\mu\nu} + K_{\mu\nu}$ is not a real tensor; this is a famous problem, there is no real way to define the total energy-momentum tensor of the universe.

3. A geometrical look at gauge invariance

At each point in space-time imagine a frame defining axes in SU(n) in some sense continuous (nearby frames nearly the same), but otherwise arbitrary. Then physically we might hope to define what we mean by the frames at x and x' are in the same direction. That is, imagine that we take an up particle (e.g. a proton in SU(2)) at x and send it over to x' so the guy at x' can see what we call up. If there were no external influence, which could rotate isospin, acting in the space between x and x', we might hope to define and check that everyone is using the same frame and can expect that one choice of frames at all x to be best in the sense of making the physics equations simplest (e.g. all parallel).

But under an influence, we must correct for the influence. If the influence is not universal (e.g. acts on protons but not on pions), we can compare frames using a particle which is affected least, or by looking at the different ways in which different particles are affected. If a universal influence acts which rotates the axis of isospin of every particle to the same degree, it clearly has no effect locally; but if the rotation varies from point to point and from time to time, than under some circumstances there may be an effect.

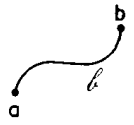
Now suppose there is such a universal influence and let us try to compare a frame at b to a frame at a by sending a particle from a to b. Then the frame at a "carried" to point b might find (the frame as it is carried is of course turned by the universal influence), in general, that it is not lined up with the frame originally chosen at b, but requires an additional rotation $R(b \leftarrow a)$. Thus $R(b \leftarrow a)$ tells us how much the frame b differs from the frame a when a is carried over, through space-time, to make the comparison. We could, of course, get a set of "best" frames so that all $R = 1$ by choosing at b the frame we get by carrying our a frame to each space-time point.

$$R(c \leftarrow a) = R(c \leftarrow b)\, R(b \leftarrow a)\ , \tag{3.1}$$

i.e. unless $R(c \leftarrow a)$ is independent of the route by which a is carried, in which case we do not have an interesting physical theory at all: R can be made equal to the identity by a proper choice of a universal set of frames.

The interesting theory arises if this is not the case, i.e. if the rotation $R(\overset{b \leftarrow a}{\ell})$ depends on the path ℓ in space connecting the points. We now study this gen-

eral case – the geometry of a field of frames with a method of parallel displacement, or comparison along geomatrical paths. (If the frames were Lorentz frames arbitrarily displaced, the theory is differential geometry and the physical theory is Einstein's General Relativity.)



Of course $R(^{b\leftarrow a}_{\mathscr{C}})$ which compares frames at b and a does depend on the original frame choice at each point. If the frames at each point were rotated by $P(a)$ then the rotation between b and a would now be

$$R'\left(^{b\leftarrow a}_{\mathscr{C}}\right) = P(b)R\left(^{b\leftarrow a}_{\mathscr{C}}\right)P^{-1}(a)\,. \tag{3.2}$$

Physics should not depend on this choice, so we look for invariant properties of R, this is most easily done geometrically.

For a closed path $\mathscr{C}_0$, $a \leftarrow a$, we might get a resulting rotation $R(^{a\leftarrow a}_{\mathscr{C}_0})$ dependent on $\mathscr{C}_0$. This determines a "strain" or physical effect independent of the choice of frames and thus invariant. To analyse these things most easily, we work with infinitesimal displacements and closed circuits (as any finite closed circuit can be represented as an area integral of infinitesimal closed circuits, in the manner familiar in the usual demonstration of Stokes's theorem).

Thus consider b to be separated from a by an infinitesimal coordinate displacement Δx_μ. Then R is nearly 1, the difference being of order Δx_μ. Hence we can write, to first order in Δx_μ

$$R(a + \Delta x \leftarrow a) = 1 - \mathrm{i}\eta_\mu(x)\Delta x_\mu\,, \tag{3.3}$$

where η_μ is a vector field, in Minkowski space, depending on x, the location of point a, and is an operator in isospace. The transformation property of η_μ under a rotation P of the frames is given (using (3.2)) by

$$1 - \mathrm{i}\eta'_\mu \Delta x_\mu = P(x + \Delta x)(1 - \mathrm{i}\eta_\mu \Delta x_\mu)P^{-1}(x)\,.$$

Putting

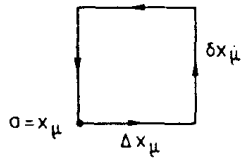
$$P(x + \Delta x) = P(x) + \frac{\partial P(x)}{\partial x_\mu}\Delta x_\mu$$

we get

$$\eta'_\mu = P(x)\eta_\mu P^{-1}(x) + \mathrm{i}\frac{\partial P(x)}{\partial x_\mu}P^{-1}(x) \tag{3.4}$$

(we call this a gauge transformation of η).

What happens if we go round a small square?



(3.5)

Calculating to second order we obtain (to be correct to 2nd order we should expand each R to second order beyond (3.3), but it is readily seen that these terms will cancel in this order in going up and down the sides of the square, i.e. such terms in the first bracket will cancel with terms in the third bracket, since to second order they are opposite)

$$R = \left[1 + i\eta_\mu\left(x + \frac{\delta x}{2}\right)\delta x_\mu + \ldots\right]\left[1 + i\eta_\nu\left(x + \frac{\Delta x}{2} + \delta x\right)\Delta x_\nu + \ldots\right]$$
$$\times \left[1 - i\eta_\sigma\left(x + \Delta x + \frac{\delta x}{2}\right)\delta x_\sigma + \ldots\right]\left[1 - i\eta_\tau\left(x + \frac{\Delta x}{2}\right)\Delta x_\tau + \ldots\right]$$
$$= 1 + i\{\partial_\mu\eta_\nu - \partial_\nu\eta_\mu + i[\eta_\mu, \eta_\nu]\}\delta x_\mu \Delta x_\nu \tag{3.6}$$
$$= 1 + i\mathcal{M}_{\mu\nu}\delta x_\mu \Delta x_\nu , \tag{3.7}$$

where we have defined

$$\mathcal{M}_{\mu\nu} = \partial_\mu\eta_\nu - \partial_\nu\eta_\mu + i[\eta_\mu, \eta_\nu] , \tag{3.8}$$

which is associated with an area $\delta x_\mu \Delta x_\nu$ and is an antisymmetric tensor operator of the second rank. $\mathcal{M}_{\mu\nu}$ is the physically interesting thing associated with the connection η_μ which takes us from place to place.

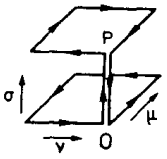
Suppose $\mathcal{B}$ is any tensor operator. We wish to know how it changes from place to place. It will not do simply to take $\mathcal{B}(x + \Delta x) - \mathcal{B}(x)$ since we cannot compare objects at a distance because of the effects of the universal influence. We must take account of the rotation of the frame by transporting $\mathcal{B}(x + \Delta x)$ back to x before making a comparison. Hence the total change in $\mathcal{B}$ is

$$[1 + i\eta_\mu \Delta x_\mu]\mathcal{B}(x + \Delta x)[1 - i\eta_\nu \Delta x_\nu] - \mathcal{B}(x)$$
$$= \left[1 + i\eta_\mu \Delta x_\mu\right]\left(\mathcal{B}(x) + \frac{\partial}{\partial x_\tau}\Delta x_\tau\right)\left[1 - i\eta_\nu \Delta x_\nu\right] - \mathcal{B}(x)$$
$$= \left\{\frac{\partial \mathcal{B}(x)}{\partial x_\mu} + i[\eta_\mu, \mathcal{B}(x)]\right\}\Delta x_\mu . \tag{3.9}$$

This enables us to define a covariant derivative on any tensor $\mathcal{B}$ by

$$D_\mu \mathcal{B} = \partial_\mu \mathcal{B} + i[\eta_\mu, \mathcal{B}] . \tag{3.10}$$

We will now deduce an interesting geometrical identity satisfied by $\mathcal{M}_{\mu\nu}$.



We wish to calculate the difference in circulation between the top and bottom faces of the cube. To get this difference it will not do simply to take $\mathcal{M}_{\mu\nu}(x + dx_\sigma) - \mathcal{M}_{\mu\nu}(x)$, because to get back to O we must go from O to P (a factor $1 - i\eta_\tau dx_\tau$), then around the top (a factor $1 + i\mathcal{M}_{\mu\nu}(x + dx)\delta x_\mu \Delta x_\nu$), and then back down to O (a factor $1 + i\eta_\sigma dx_\sigma$); hence we want to compare

$$[1 + i\eta_\sigma dx_\sigma][1 + i\mathcal{M}_{\mu\nu}(x + dx_\sigma)\delta x_\mu \Delta x_\nu][1 - i\eta_\tau dx_\tau]$$

with

$$1 + i\mathcal{M}_{\mu\nu}(x)\delta x_\mu \Delta x_\nu .$$

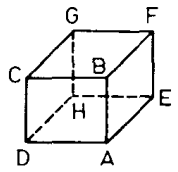
The difference between these two terms is

$$\left\{\frac{\partial \mathcal{M}_{\mu\nu}}{\delta x_\sigma} + i[\eta_\sigma \mathcal{M}_{\mu\nu}]\right\} dV_{\mu\nu\sigma} , \tag{3.11}$$

where

$$dV_{\mu\nu\sigma} = \delta x_\mu \Delta x_\nu dx_\sigma .$$

Notice that the term in the bracket is just the covariant derivative of $\mathcal{M}_{\mu\nu}$; this is not surprising as all we have done is to compare $\mathcal{M}_{\mu\nu}$ (cf. (3.5)) at O and P.



By comparing $\mathcal{M}_{\mu\nu}$ on opposite pairs of faces, and noting that the net effect of going round the sum of the following 3 paths is zero,

$$(A \to B \to C \to G \to F \to B \to A \to E \to H \to D \to A)$$

$$+ (A \to D \to C \to B \to A \to E \to F \to G \to H \to E \to A)$$

$$+ (A \to B \to F \to E \to A \to D \to H \to G \to C \to D \to A) ,$$

we get

$$D_\sigma \mathcal{M}_{\mu\nu} + D_\mu \mathcal{M}_{\nu\sigma} + D_\nu \mathcal{M}_{\sigma\mu} = 0 . \tag{3.12}$$

By taking B round a small closed loop, or by using (3.10), we get

$$D_\mu D_\nu B - D_\nu D_\mu B = \mathrm{i}[\mathcal{M}_{\mu\nu}, B] , \tag{3.13}$$

and hence

$$D_\mu D_\nu \mathcal{M}_{\mu\nu} = 0 . \tag{3.14}$$

Now we will associate this with what we did before with Yang–Mills theory. Here each R is a rotation in isospace and so can be written in the form

$$R = \exp(-\mathrm{i}\boldsymbol{\alpha} \cdot T) \tag{3.15}$$

where the T are the generators of some representation of the SU(2) Lie algebra, and $\boldsymbol{\alpha}$ is a vector depending on the rotation it is desired to represent.

For the infinitesimal rotation (3.3), $\boldsymbol{\alpha}$ is also an infinitesimal of first order in Δx_μ, say $\boldsymbol{\alpha} = -A_\mu \Delta x_\mu$ so

$$R = 1 + \mathrm{i}(A_\mu \cdot T)\Delta x_\mu . \tag{3.16}$$

Comparing this with (3.3) we see that

$$\eta_\mu = -A_\mu \cdot T . \tag{3.17}$$

Substituting this into (3.8) we obtain

$$\mathcal{M}_{\mu\nu} = -E_{\mu\nu} \cdot T \tag{3.18}$$

where

$$E_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu \times A_\nu , \tag{3.19}$$

and we see that is the same as the $E_{\mu\nu}$ which we obtained in (2.24). The question now is – can we understand any of the Yang–Mills field equations (2.26)–(2.37) geometrically? The answer is for most of the equations, no. In particular, why did we put the combination $-(1/4g^2)E_{\mu\nu} \cdot E_{\mu\nu}$ in the Lagrangian? Physics tells us, not geometry. All we have discussed are the qualitative features of a classical field, and the response of particles to it (they rotate the axes as they move through it), but how the field itself has energy and behaves

dynamically we haven't determined. But there are a few things we can work out. Using (3.12) and (3.18) we have

$$D_\mu E_{\nu\sigma} + D_\nu E_{\sigma\mu} + D_\sigma E_{\nu\mu} = 0 \,. \tag{3.20}$$

And using (3.13) and (3.14), we have

$$D_\mu D_\nu \bar{\boldsymbol{\phi}} - D_\nu D_\mu \bar{\boldsymbol{\phi}} = E_{\mu\nu} \times \bar{\boldsymbol{\phi}} \,, \tag{3.21}$$

where $\bar{\boldsymbol{\phi}}$ is any isovector and

$$D_\mu D_\nu E_{\mu\nu} = 0 \tag{3.22}$$

(cf. (2.35) and (2.36)).

The analogue of (2.34) would be

$$D_\mu {}^{*}\mathcal{M}_{\mu\nu} = g_\nu \,. \tag{3.23}$$

What is $D_\mu {}^{*}\mathcal{M}_{\mu\nu}$ geometrically? If I knew an easy way to describe this geometrically then we could state this equation as: $D_\mu {}^{*}\mathcal{M}_{\mu\nu}$ is the total isospin in a small volume. Unfortunately I haven't worked this out, and therefore I cannot describe the full Yang–Mills classical theory in an elementary way. (I am looking for a law like that in gravity, which says that the excess of the proper radius of a small 3 dimensional sphere over the radius calculated from the area, $\sqrt{(\text{area}/4\pi)}$, is proportional to the mass inside the sphere – which is, assuming Lorentz invariance, a complete statement of the Einstein law $R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = T_{\mu\nu}$.)

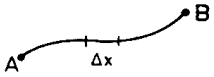
We have said that the A field represents a universal turning of the axes of a diffusing particle, and now we should check that the equation of motion of the matter $(\mathrm{i}\partial\!\!\!/ - A\!\!\!/ \cdot \boldsymbol{\tau})\psi = m\psi$, for example, implies that indeed it does. To make it is easy, we first do it with the Schroedinger equation and electrodynamics but you will see that the method of proof is readily extendible to other cases. The free particle Schroedinger equation is

$$-\frac{\nabla^2 \psi}{2m} = \mathrm{i}\frac{\partial \psi}{\partial t} \,. \tag{3.24}$$

Solving this equation, we find that a particle propagates as follows. Suppose it is confined at x_1 at time t_1 (the wave-function is a delta-function). Then the wave-function at time t_2 and position x_2 is

$$\left[\frac{2\pi \mathrm{i}(t_2 - t_1)}{m}\right]^{-3/2} \exp\left[\frac{\mathrm{i}m(x_2 - x_1)^2}{2(t_2 - t_1)}\right], \tag{3.25}$$

which we call $K_0(2,1)$, this being the function which describes how the particle diffuses outwards in time. Now consider a possible trajectory for the particle



The amplitude for the particle to go from A to B can be found by multiplying the amplitudes for it to go along all the infinitesimal sections of the path.

Now add an external field to the Schroedinger equation

$$\frac{1}{2m}(\mathrm{i}\underline{\nabla} - \underline{A})^2\psi + V\psi = \mathrm{i}\frac{\partial\psi}{\partial t} \tag{3.26}$$

Consider an infinitesimal distance Δx. Over this distance, A and V may be treated as constants and are therefore expressible as the gradient of a potential χ

$$\underline{A} = \underline{\nabla}\chi, \qquad V = \partial\chi/\partial t\ , \tag{3.27}$$

where

$$\chi = \underline{A}\cdot\underline{x} + Vt\ .$$

By a gauge transformation we see that the wave-function $\psi' = \mathrm{e}^{-\mathrm{i}\chi}\psi$ is a solution of (3.26) if ψ is a solution of (3.24), and hence we may write the amplitude for propagation over a short distance and time Δx in the presence of the field as

$$K_A(x + \Delta x, x) = \exp\{\mathrm{i}\chi(x + \Delta x)\}\exp\{-\mathrm{i}\chi(x)\}K_0(x + \Delta x, x)$$

$$= \exp(\mathrm{i}A_\mu \Delta x_\mu)K_0(x + \Delta x, x)\ , \tag{3.28}$$

where

$$A_0 = -V, \qquad A_i = \underline{A}\ .$$

Iterating this over a continuous path, we get

$$\exp(\mathrm{i}\int A_\mu \mathrm{d}x_\mu) \times (\text{Amplitude for process without the } A \text{ field})\ . \tag{3.29}$$

In the case of Yang–Mills theory, this generalises to

$$\exp(\mathrm{i}\int A_\mu \cdot T\mathrm{d}x_\mu) \times (\text{Amplitude for process without Y–M field})\ , \tag{3.30}$$

where we must order the operator T along the path

For infintesimal paths this reduces to

$$[1 + i(A_\mu \cdot T)\Delta x_\mu] \times (\text{Amplitude for process without Y–M field}) , \quad (3.31)$$

and thus we see that the expression (3.16) has emerged naturally.

4. A qualitative critique of QCD

One of the possible applications of Yang–Mills theory is to Quantum Chromodynamics. The question to which we wish to address ourselves in this chapter is whether or not this theory has a real chance of being right. Ordinarily, when the right theory is found, it isn't long before we can calculate consequences and check that it agrees with everything relevant that is known. For example, the whole subject of electrostatics was in complete confusion until the Coulomb law was discovered; before then people were rushing around in complete chaos and then suddenly they were calculating the capacity of elliptical condensors etc. .... Similarly, before Schroedinger's equation, there was a lot of pulling and hauling on ideas which were inconsistent, and suddenly, as soon as the equation was discovered, there was a tremendous tumbling out of results which showed how everything worked. Therefore it's expected (at least by an old fogey like myself) than when the correct theory is found, lots of results will tumble out which will agree with experiment. Now QCD is proposed as a theory which is supposed to be the correct theory of strong interactions; it's been around for a few years now and we don't have any quantitative results. At the moment we cannot look at the theory quantitatively (due possibly to technical difficulties in its interpretation) so we will look at it qualitatively to try and decide whether it is useful or not.

4.1. Forces between the quarks

The most characteristic thing about the quark bound state which have been seen is that they are colour singlets. Coloured states have not been seen and we must conclude that either their mass is infinite or is out of the reach of present experiments.

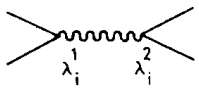
In this theory the three quarks in a baryon form an antisymmetric state with respect to colour:

$$\frac{1}{\sqrt{6}}\{|A\rangle|B\rangle|C\rangle + |B\rangle|C\rangle|A\rangle + |C\rangle|A\rangle|B\rangle - |B\rangle|A\rangle|C\rangle - |A\rangle|C\rangle|B\rangle - |C\rangle|B\rangle|A\rangle\}. \quad (4.1)$$

The first thing to look at is when we put three quarks together, can their energy be lower in some other state than in the singlet state? Similarly putting a quark and an anti-quark together, is the colourless (singlet) state lower than any other state? At the level we will work, we will not attempt to calculate the energies correctly, we merely wish to see whether we can get their order right. In QED we know that the force between two static particles is

$$\sim \frac{e^2}{r} \tag{4.2}$$

Similarly the force between two static quarks due to single gluon is



$$\sim \frac{\Sigma(\lambda_i^1)\ (\lambda_i^2)}{r} \tag{4.3}$$

where λ_i are the SU(3) matrices.

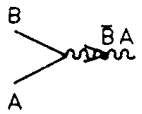
One can see this as follows. For a static source $J_x = J_y = J_z = 0$, and $J_t = \rho$, the density of colour charge, we find a solution of the form $A_x = A_y = A_z = 0$, $A_t \neq 0$ where

$$\nabla^2 A_t = \rho_t \tag{4.4}$$

(from 2.34). This is just as in electrostatics except that here we have three isospin components. We could look at this problem mathematically (by fiddling around with the λ's); however we wish to take a simpler view. The gluons which couple to the colours of the quarks, can be represented as

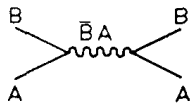
$$\bar{A}B,\ \bar{B}A,\ \bar{A}C,\ \bar{C}A,\ \bar{B}C,\ \bar{C}B,\ \frac{1}{\sqrt{2}}(\bar{B}B-\bar{C}C),\ \frac{1}{\sqrt{6}}(2\bar{A}A-\bar{B}B-\bar{C}C). \tag{4.5}$$

The last two gluons are the non colour changing states orthogonal to the singlet $(1/\sqrt{3})(\bar{A}A + \bar{B}B + \bar{C}C)$ which is omitted for reasons discussed earlier. This symbolism really tells us what happens to the colour of a quark when it emits or absorbs a gluon. That is, an $\bar{A}B$ gluon can be absorbed by an A quark turning it into B, with amplitude 1. (Annihilate the A and create a B instead.) We now calculate the relative strengths with which various quarks couple. Consider this vertex; ar A quark emits a $\bar{B}A$ gluon changing to B.



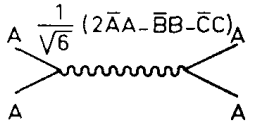
(4.6)

This $\bar{B}A$ gluon can be absorbed by a B quark (turning it to A) but cannot be absorbed by an A or C (it can by an $\bar{A}$ antiquark however). Therefore, if we wish to look at the A–A quark force, this vertex cannot contribute because this gluon cannot be absorbed by an A quark, i.e.



does not go.

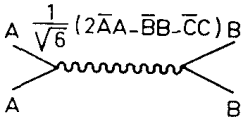
The only contribution to the A–A force is that due to the colourless gluon exchange, viz.



(4.7)

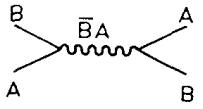
The interaction energy for this process is $+2/\sqrt{6}\,(+2/\sqrt{6}) = +\frac{2}{3}$, which is positive, so they repel just as in electrostatics (like charges repel).

Now what about the A–B force? Here there are two allowed diagrams.



$$\text{interaction energy} = +\frac{2}{\sqrt{6}}\left(-\frac{1}{\sqrt{6}}\right) = -\frac{1}{3}, \quad (4.8)$$

and the exchange diagram



$$\text{interaction energy} = (+1)(+1) = +1\ . \quad (4.9)$$

4.8 and 4.9 must be combined to give the total interaction energy which gives $+\frac{2}{3}$.

We could do the same for all the other colour combinations, but we'd be wasting our time as we know that the interaction is symmetric with respect to colour. The only cases we have to worry about are when the colours of the quarks are the same or different. We can summarize the results as follows. Introduce a colour exchange operator P which interchanges a pair of quarks, it

has eigenvalues +1 or −1 corresponding to whether the quarks are in a symmetric or antisymmetric colour state. In the case when the quarks are different (cf. (4.8), (4.9), the interaction energy can be written

$$(p - \tfrac{1}{3}) , \tag{4.10}$$

where p is the eigenvalue of P. The lovely thing about this formula is that, if we apply it to the case where the colours are the same (cf. (4.7), $p = 1$ in this case), we get the right answer $1 - \frac{1}{3} = \frac{2}{3}$. Thus we have shown that $\Sigma(\lambda_i)^1(\lambda_i)^2 = P - \frac{1}{3}$. This is analogous to the Dirac formula for the interaction between two spins (e.g. two electrons in an atom):

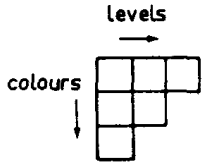
$$p_{\text{exch}} - \tfrac{1}{2} = \sum_i (\sigma_i)^1(\sigma_i)^2 , \tag{4.11}$$

(the $-\frac{1}{2}$ becomes $-1/n$ for SU(n).

In order to generalise formula (4.10) to states of more than two quarks, we merely sum it over all possible pairs of quarks.

We shall now calculate the energy of various quark states using (4.10). Each quark will have some self-energy, we don't know what this is, but to make it easier to see what's going on, we will take $+\frac{4}{3}$ for each quark – the qualitative results do not depend on this choice (because we will always compare the energies of states of the same total number of quarks).

We will symbolise the quark states as follows: draw a series of boxes with one box for each quark in the state, for example a possible 6 quark state is



(4.12)

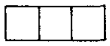
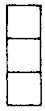
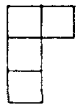
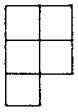
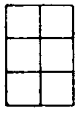
This represents a quark configuration in which the wave-function is symmetric with respect to the interchange of any pair of quarks in the same row, and antisymmetric with respect to the interchange of any pair of quarks in the same column, e.g. for a two quark state,  represents the symmetric state and  represents the antisymmetric state. We can use these (Young) diagrams to calculate the energy of a bound state using (4.10). We will work out one of these diagrams in detail for a three quark state which is symmetric under exchange of one pair of quarks and antisymmetric under exchange of another pair:

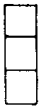
(4.13)

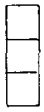
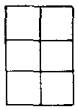
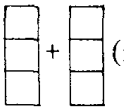
The contribution from the exchange is

$$\underset{\text{row}}{(+1)} + \underset{\text{column}}{(-1)}$$

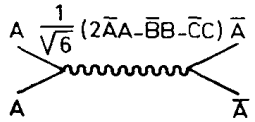
Therefore the total interaction energy is $(+1) + (-1) - 3(\frac{1}{3}) = -1$. When we add $+\frac{4}{3}$ for each quark we obtain $+3$. The following table summarises the results for various states.

No. of quarks	Diagram	Interaction energy	State energy ($+\frac{4}{3}$ added per quark)
1		0	$+\frac{4}{3}$
2		$+1 - \frac{1}{3} = +\frac{2}{3}$	$+\frac{10}{3}$
		$-1 - \frac{1}{3} = -\frac{4}{3}$	$+\frac{4}{3}$
3		$3(+1) - 3(\frac{1}{3}) = +2$	+6
		$+1 + (-1) - 3(\frac{1}{3}) = -1$	+3
		$3(-1) - 3(\frac{1}{3}) = -4$	0
4		$+1 + 3(-1) - 6(\frac{1}{3}) = -4$	$+\frac{4}{3}$
5		$2(+1) + 4(-1) - 10(\frac{1}{3}) = -\frac{16}{3}$	$+\frac{4}{3}$
6		$3(+1) + 6(-1) - 15(\frac{1}{3}) = -8$	0

Notice that  , which is totally antisymmetric with respect to colour and which we identify with a baryon, has the lowest energy. Further  has the

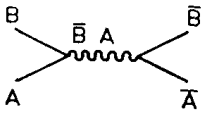
same energy as  i.e. a single quark should be only weakly bound to a proton. Similarly  has the same energy as  (in our approximation, zero) and it is impossible to say, in this poor approximation, whether or not such a bound state of two baryons will exist.

Now let us try the same thing for the mesons. Consider



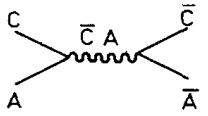
$$\text{interaction energy} = +\frac{2}{\sqrt{6}}\left(-\frac{2}{\sqrt{6}}\right) = -\frac{2}{3}. \qquad (4.14)$$

The minus sign appears at the antiparticle vertex because the theory is a vector theory; this is just like electricity where the antiparticle has the opposite charge to the particle. Another possible term in the A–$\bar{A}$ interaction is that in which the quarks change their states



$$\text{interaction energy} = +1\,(-1) = -1 \qquad (4.15)$$

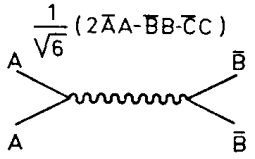
Similarly



has the same energy as (4.15), -1 . (4.16)

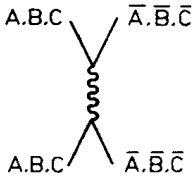
Now we can figure out the energy of the meson state $(1/\sqrt{3})(|A\rangle|\bar{A}\rangle + |B\rangle|\bar{B}\rangle + |C\rangle|\bar{C}\rangle)$ (singlet) i.e. we need to consider the coupling of this state to itself. Since this state is symmetric in A, B, C we get $3 \times 1/\sqrt{3}|A\rangle|\bar{A}\rangle \times 1/\sqrt{3}(|A\rangle|\bar{A}\rangle + |B\rangle|\bar{B}\rangle + |C\rangle|\bar{C}\rangle)$ which is simply the sum of (4.14)–(4.16). Therefore the interaction energy is $-\frac{8}{3}$. Adding $+\frac{4}{3}$ for each quark, we find the meson energy to be 0. (The mesons have the same energy scale as baryons, with this $+\frac{4}{3}$ choice for each quark.)

Finally, we examine the state $|A\rangle|\bar{B}\rangle$, a coloured meson; there is only one possible diagram



$$\text{interaction energy} = +\frac{2}{\sqrt{6}}\left(+\frac{1}{\sqrt{6}}\right) = +\frac{1}{3}. \qquad (4.17)$$

Adding the quark self-energy we get a coloured meson energy of +3, which is much higher than for a colourless meson. We could ask why



does not contribute to the energy of the colourless meson. The answer is that the s channel gluon must be colourless; no such gluon exists – it was eliminated earlier.

4.2. Infra-red behaviour

We have indicated that a r^{-1} potential exists between two quarks in this theory. We know that this cannot really be correct, since the force between two quarks is known not to be long range – it is not easy to knock the quarks apart inside a proton (c.f. the case with which we can knock electrons out of atoms). This problem cannot be argued away by saying that the charges in QCD are stronger than in QED – with enough energy, we should be able to pull them apart. The long distance (infra-red) forces have to be modified in order to agree with experiment. Those who believe in QCD believe that this will happen when the theory is worked out. I think that the central problem in QCD is to see if, qualitatively, the forces are so modified. In some models, e.g. the lattice theory (see [12]), they are, but we are not sure if this an artifact of the models or not. It is true that when we go to higher order perturbations, we find the force increasing with distance (there are logarithmic corrections) relative to $1/r^2$, but it is unknown if this change is sufficient.

4.3. Dependence of masses on flavours

4.3.1. Isospin dependence

The proton and the Δ are both supposed to be made out of 3 quarks in a totally antisymmetric colour state.

	Mass (MeV)	Isospin	Spin
P	938	$\frac{1}{2}$	$\frac{1}{2}$
Δ	1236	$\frac{3}{2}$	$\frac{3}{2}$

If QCD has forces depending only colour, how could there be a difference between these two masses depending on the isospin? In addition to the isospin being different, the spins are also different and we know, for example, that the forces are spin dependent and different for different spin states in QED; so there is no reason why this cannot be the case in QCD. The difference in the P and Δ masses, therefore, may simply be due to the fact that their spins are different. Then we say – look through the Rosenfeld tables to find particles with different isospins and masses, but the same spins and for which we expect the same space. Then QCD could be in trouble. However we cannot find any for the following reason.

The wave-function is antisymmetric with respect to interchange of two quarks (Fermi statistics). This total exchange is equivalent to an exchange of space, spin, flavour and colour. Since the state is antisymmetric with respect to colour exchange (colour singlet), the flavour symmetry must equal the space symmetry $\times$ the spin symmetry; hence the flavour symmetry properties of a state are completely determined by its space and spin symmetry properties. But colour forces depend on the space and spin configurations and therefore can apparently depend on the isospin symmetry. In other words, we can find no way to verify the proposition that the forces are independent of flavour. This, by the way, should be noted because in the early days of hadron theory the forces had explicit isospin dependence, e.g. there were interaction terms like $\bar{\psi}(\phi \cdot \gamma)\psi$, for a proton coupled to a pion. In QCD we cannot have any such directly isospin dependent terms, but it doesn't matter as we have seen that the masses can depend indirectly on isospin.

A very rough empirical formula exists [13] which summarises the mass splitting between baryon multiplets. It says that there is a contribution of -0.53 (GeV)^2 to the (mass)^2 for every pair of quarks which are both symmetric in space and antisymmetric in spin. Mathematically this is

$$\Delta M^2 = -0.53\ (\mathrm{GeV})^2 \sum_{\text{pairs}} \left(\frac{1 - P_{\text{ex spin}}}{2}\right)\left(\frac{1 + P_{\text{ex space}}}{2}\right). \tag{4.18}$$

Notice that the masses are lower if the spacial state is symmetric as opposed to antisymmetric, in an antisymmetric state, it is impossible for the quarks to be at the same point; but in the symmetric state this is allowed, so that the qualitative features of (4.18) may be summarized by saying that the correction force is short range. The spin part says that antiparallel spin states are lower which is the right sign for the spin interaction of attracting particles due to a vector potential.

4.3.2. Dependence on the quark masses

We know that the K mass does not equal the π mass, although the K and the π both have the same spin-parity. How can we explain this? If all the forces are independent of flavour, then the masses should be exactly the same. It is therefore a failure of the simplest possible picture that SU(3) is broken. But we do not need to destroy QCD, we need only complicate it by supposing that the s quark has a higher mass than the u and d quarks (but keeping the interaction independent of flavour). No-one knows where this extra mass comes from, we have left for the people of the future the problem of why the masses are different. Some people think that the masses of the u and d quarks are equal and that all the mass differences in isospin multiplets are due to electrodynamics (e.g. proton–neutron), there are some technical difficulties in signs and magnitudes in this approach and it would help if we could say that the d quark is slightly heavier than the u quark for the same intrinsic (unknown) reason that the s quark is heavier than the u and d quarks. Another flavour now seems to have been found and maybe there are more; this new quark (usually called charm) must have a higher mass than the other three. We mustn't forget that it is strange that we have to put in mass differences which are of the same order of magnitude as the bound state masses we are trying to explain. If we go to very high momentum transfers where masses are irrelevant, the SU(3) (flavour) should become better, I don't know of any direct demonstration that this is, or is not, the case – it would be nice to think of some experiment in which this could be tested.

4.4. Zweig rule

Consider the following groups of mesons

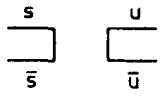
	π	η	η'	0^-
Mass(GeV)	0.14	0.55	0.96	
	ρ	ω	ϕ	1^-
Mass(GeV)	0.77	0.78	1.02	
	A_2	f	f'	2^+
Mass(GeV)	1.31	1.27	1.52	

These are the non-strange mesons from the 0^-, 1^- and 2^+ nonets and one might expect some similarity between these groups. Now it is known that the ϕ (e.g. in its decay, $K\bar{K}$ dominates) is an $|s\bar{s}\rangle$ state. Similarly the ω (3π decay mode dominates) is $1/\sqrt{2}(|u\bar{u}\rangle + |d\bar{d}\rangle)$ and the ρ is $1/\sqrt{2}(|u\bar{u}\rangle - |d\bar{d}\rangle)$ and they are very nearly degenerate. But the η and the η' do not have this pattern Approximately,

$$\eta = \tfrac{1}{2}(|u\bar{u}\rangle + |d\bar{d}\rangle - \sqrt{2}|s\bar{s}\rangle) \quad \text{and} \quad \eta' = \tfrac{1}{2}(|u\bar{u}\rangle + |d\bar{d}\rangle + \sqrt{2}|s\bar{s}\rangle)$$

agrees with experiment. Note that this combination is very different from the ω, ϕ case. Why? We would think that, in a state which contains $|s\bar{s}\rangle$, the s and $\bar{s}$ could annihilate and then turn back into s and $\bar{s}$, but also into u and $\bar{u}$ or d and $\bar{d}$; therefore an $|s\bar{s}\rangle$ state should become a mixture of all 3 quark types, and this is presumbly what is happening in the η, η' system. Why does this not happen in the ω, ϕ case (and also does not happen in the f, f' case)? This is a mystery. This mystery is summed up in the Zweig rule, an ad hoc rule which says that this annihilation process is inhibited.

One possible attitude is as follows



With an 0^- state, we need 2 gluons to connect both sides of this diagram. At first sight, we would think that with a 1^- state, we could connect with 1 gluon since the gluon has quantum numbers 1^- (cf. photon); but this would require a colour singlet gluon and we have no such object; so we need a minimum of 3 gluons. If we could assume that the gluons were weakly coupled to the quarks, we might try to argue that the mixing in the ω, ϕ case is less than in the η, η' case, but it's a little hard to get g^3 less than g^2 unless g (the coupling constant) itself is very small. Furthermore, if we look at the 2^+ mesons,

we could make the connection in the diagram with only 2 gluons, so it should look rather like η, η' system; but it doesn't (here the suggestion was made that the higher angular momentum 2^+ makes it harder for the gluons to get together to annihilate).

There is also another attitude, which is to suppose that, for some reason, when momentum transfers and energies become large, the coupling becomes small, then we can say the mixing is less in the 2^+ case than in the 0^- case because the energy is higher. In order for this work, the rate of change of coupling with mass scale must be large.

This is not a satisfactory situation – it looks suspicious, as there is some feature of QCD that we do not understand. If we have a way of calculating with QCD (e.g. on a lattice) and if we wish to concentrate on something which will tell us if the theory is wrong, this seems to me to be an ideal place.

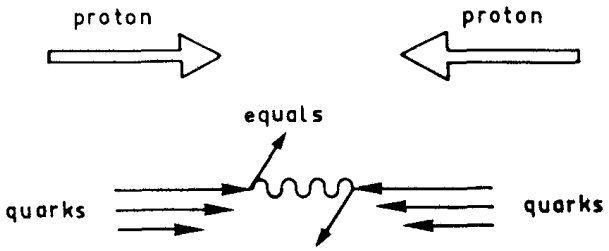
4.5. Large transverse momentum in hadronic collisions

Consider the process

$$\mathrm{p} + \mathrm{p} \to \text{hadron (large } p_\perp) + \mathrm{X} \,. \tag{4.19}$$

The average $p_\perp$ of a produced hadron at high energies is of order 350 MeV. In a field theory, it is not easy to understand why there isn't a reasonable amount of larger $p_\perp$. Let us look specifically at particles produced at very large $p_\perp$.

In our picture the proton is made of a bunch of quarks, so we might expect that a large $p_\perp$ particle would be produced by quarks scattering via gluon exchange.



(4.20)

If the ratio $x_\perp = p_\perp/p$ is kept constant, the cross-section at large $p_\perp$ should go as $p_\perp^{-4}$. (This follows merely by dimensional arguments and this result would be obtained for any gluon diagram.) Experimentally, it is more like $p_\perp^{-8.2}$. This is very disturbing. Is this process operating or not? A possibility is that it really does occur, but when we put in the correct couplings and allow for

the fact that the coupling constant will fall as $p_\perp$ increases, the process (4.20) is masked by some other process which falls like $p_\perp^{-8.2}$. We must then assume that this other process hasn't yet fallen enough to let us see the mechanism (4.20). However the cross-section is already pretty small at 400 GeV and still seems to be falling like $p_\perp^{-8}$; it is therefore up to the people who propose this explanation to give some energy above which they believe the process (4.20) will take over; and to explain what mechanism is responsible for the present trend. (One possible mechanism is described by the constituent interchange model, which gives a cross-section falling like $p_\perp^{-8}$.) But in any event there is a challenge to see what process QCD predicts that is so large as to dominate over all the experimental range so far investigated and which behaves like $p_\perp^{-8}$. Any theory of this process has to explain much data such as charge ratios correlations, etc... . Details such as the fact that as $x_\perp$ rises (toward 0.6) the ratio π^+/π^- rises to more than 2, have to be explained.

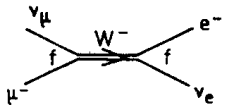
4.6. Partially conserved axial current and vector meson dominance

Certain hadrons have special properties and it is not clear where these properties come from. One of these hadrons is the pion, which has the property of PCAC associated with it; another is the ρ meson which participates in VMD.

Both PCAC and VMD were discovered when it was thought that the particles were themselves fundamental fields. Why bound states of quarks (as QCD assumes the π and ρ are) should behave as if they were fundamental fields is a puzzle. I don't know whether this is a serious problem for QCD or not.

5. Spontaneously broken symmetries and the Higgs mechanism

We will try to show why it is necessary to use a Yang–Mills theory in weak interactions by considering the difficulties which appear in a more phenomenological approach. We will discuss some relevent points from weak interaction theory although we do not propose to give a review of it. Consider the decay of a μ^-; one interpretation of this decay is the following diagram.



(5.1)

where the weak interaction is mediated by a massive charged spin one boson W^-. (The W^- must be massive as the weak interaction is short range.) The

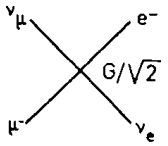
couplings at the two vertices are assumed to be V–A and equal (coupling constant f). The amplitude for this decay is

$$\alpha = \frac{f^2}{q^2 - M_W^2}, \tag{5.2}$$

where q is the momentum transfer between the μ^- and the e^-. For $q \ll M_W$, this reduces to

$$f^2/M_W^2 . \tag{5.3}$$

In this limit the interaction looks like the four point Fermi interaction



Amplitude = $(G/\sqrt{2})$ (5.4)

i.e. comparing (5.3) with (5.4) we get

$$f^2/M_W^2 = G/\sqrt{2} . \tag{5.5}$$

Let us write a Lagrangian for the W^- and its anti-particle, the W^+, with a mass; we do this by analogy with QED for a massive photon:

$$\mathcal{L} = -\tfrac{1}{4}(\partial_\mu W_\nu - \partial_\nu W_\mu)^2 + \tfrac{1}{2}M^2 W_\mu W_\mu + \text{matter terms} . \tag{5.6}$$

Varying the action with respect to W_μ to obtain the equations of motion, we get

$$\partial_\mu(\partial_\mu W_\nu - \partial_\nu W_\mu) + M^2 W_\nu = S_\nu , \tag{5.7}$$

where S_μ is the current generated by the matter fields in the Lagrangian. S_μ is not conserved: one contribution to S_μ (the $\mu\nu W$ vertex in (5.1)) is

$$\bar{\psi}(\nu_\mu)\gamma_\mu(1 - \gamma_5)\psi(\mu) . \tag{5.8}$$

This has non-zero divergence. It would have zero divergence if the γ_5 were absent and the masses of μ^- and ν_μ were equal, but this non-zero divergence is all right since the divergence of (5.7) is

$$M^2 \partial_\mu W_\mu = \partial_\mu S_\mu . \tag{5.9}$$

So we see that the presence of the mass term saves us from a possible inconsistency.

We derive the propagator for the W_μ as follows. Eq. (5.7) can be re-written using (5.9) as

$$\Box W_\mu + M^2 W_\mu = S_\mu + \frac{1}{M^2}\partial_\mu(\partial_\nu S_\nu) . \tag{5.10}$$

In momentum space, this is

$$W_\mu = \frac{1}{k^2 - M^2}\left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{M^2}\right) S_\nu . \tag{5.11}$$

Therefore, given the sources S_ν, we can calculate the field W_μ using the propagator

$$\frac{\delta_{\mu\nu} - (k_\mu k_\nu / M^2)}{k^2 - M^2} . \tag{5.12}$$

This tells us how a field W_μ propagates from one interaction to the next. Using this, we can construct diagrams for the theory. There are indices on the propagator because W is a vector and therefore has various polarisation states. We can see that a free W_μ has three polarisation states, as required of a vector particle, and not four as we might guess at first sight: take (5.7) with $S_\mu = 0$; this describes the propagation of non interacting W's. Substitute a free particle solution

$$W_\mu = e_\mu \, e^{ik\cdot x} , \tag{5.13}$$

where e_μ is the W polarisation vector. This gives

$$-k_\mu(k_\mu e_\nu - k_\nu e_\mu) + M^2 e_\nu = 0 . \tag{5.14}$$

Multiplying by k_ν we get

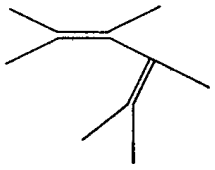
$$-k^2(k\cdot e) + k^2(k\cdot e) + M^2(k\cdot e) = 0 . \tag{5.15}$$

Hence

$$k\cdot e = 0 \quad (M \neq 0) , \tag{5.16}$$

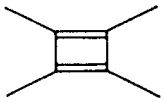
i.e. the polarisation vector e_μ is orthogonal to k_μ and so has only three degrees of freedom. Of course, when the W's are off mass shell, there are four polarisation states. Similarly, in massless QED, the photon has two polarisation states when it is on its mass shell, and three when it's off (the extra degree of freedom represents the freedom to make gauge transformations in QED which is lost if a mass term is added as in (5.6)).

We attempt to calculate with this massive W theory. Everything works well as long as we do not have any diagrams with closed loops i.e. as long as we only calculate tree diagrams:



(5.17)

Another way to say this is that the theory works perfectly at the classical level. At the quantum level, we run into difficulties with closed loops



(5.18)

The reason is that we must integrate over the momentum k running round the closed loop. At large k, the propagator, 5.12 $\sim k_\mu k_\nu / k^2$ so the propagator does not assist the convergence of the integral; the diagram diverges. Unfortunately we cannot remove these divergences as in QED, because as we go to more loops the divergences become more and more severe: (i.e. it is not renormalizable), e.g.

(5.19)

The theory is therefore a disaster quantum mechanically and in order to construct a workable renormalisable theory of weak interactions we go to a Yang–Mills theory with broken symmetry. To do this, we first discuss symmetry breaking in simpler cases. I shall take a rather physical view of symmetry breaking and leave the more abstract mathematics to other people. I do this because I think that it is useful to have more than one way of looking at a problem; the way I shall present it may be unfamiliar, but people may benefit from this physical approach.

We will start with a simple model and gradually increase its complexity. First take a real scalar field only; the Lagrangian is

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu \phi)^2 - \frac{\mu^2}{2}\phi^2 - \frac{\lambda}{4}\phi^4 , \tag{5.20}$$

where we have included a ϕ^4 interaction term. The theory has discrete symmetry $\phi \to -\phi$. Normally the vacuum expectation value of $\phi(\langle\phi\rangle)$ is zero; however

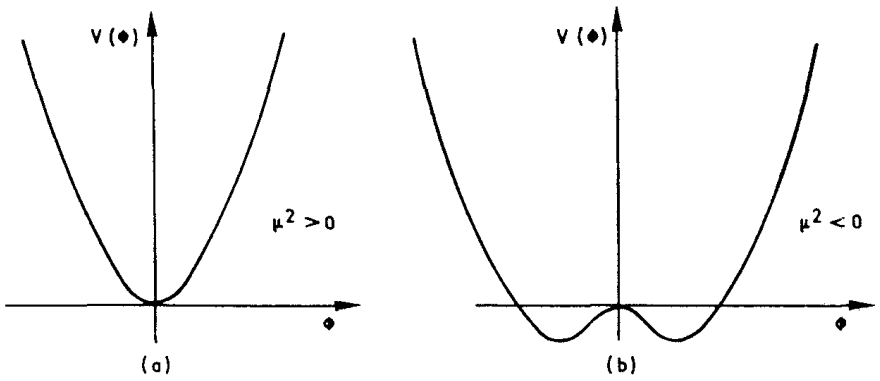


Fig. 5.1.

this is not always the case as we must choose $\langle\phi\rangle$ to minimise the energy.

In fig. 5.1, we have plotted the classical potential

$$V(\phi) = \frac{\mu^2}{2}\phi^2 + \frac{\lambda}{4}\phi^4 \tag{5.21}$$

for the two cases (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$. We look for the state of minimum energy (the vacuum) which we get by minimising $V(\phi)$:

$$\partial V(\phi)/\partial\phi = 0\,, \tag{5.22}$$

i.e. $\phi = 0$ or $\phi = \pm\sqrt{-\mu^2/\lambda}$.

From fig. 5.1, we see that

$$\text{for} \quad \mu^2 > 0, \quad \langle\phi\rangle = 0$$

$$\text{for} \quad \mu^2 < 0, \quad \langle\phi\rangle = \pm\sqrt{-\mu^2/\lambda} = \pm\, v\,. \tag{5.23}$$

Take $\langle\phi\rangle$ to be $+v$, although the choice of sign is arbitrary; however notice that once we have chosen $\langle\phi\rangle = +v$, we have broken the symmetry $\phi \to -\phi$.

In case (b), we now perturb about v

$$\phi(x) = v + \eta(x) \tag{5.24}$$

Substituting this in (5.20) we get

$$\mathcal{L} = -\tfrac{1}{2}(\partial_\mu\eta)^2 - \lambda v^2\eta^2 - \lambda v\eta^3 - \frac{\lambda}{4}\eta^4 + \frac{\lambda}{4}v^4\,. \tag{5.25}$$

Suppose η is small*; to a first approximation only the first two terms matter; these represent a scalar meson with a real mass of $\sqrt{2\lambda v^2}$. (Note that in 5.20 the ϕ meson had an imaginary mass in the case $\mu^2 < 0$.) The other two η terms in the Lagrangian represent η self couplings viz.

(5.26)

These couplings would appear when we do perturbation theory with (5.25).

Why does the whole world have $\langle\phi\rangle = +v$? Why doesn't it have $\langle\phi\rangle = -v$ somewhere? Suppose that God created the universe in the state $\langle\phi\rangle = 0$ and then the universe discovered that it could lower its energy; where it puts its energy is none of my business, but it gets rid of it – gives it back to God or something; then under some disturbance the vacuum tries to fall down with some parts going to $+v$ and other parts to $-v$. But what happens in between? It just changes suddenly, but not too suddenly, because to get low energy the $(\partial_\mu\phi)^2$ term must not be too large. This extra $(\partial_\mu\phi)^2$ energy, plus the energy due to the fact that ϕ is above its minimum potential energy, is stored in the boundary between the two regions, so that if we have some region in which $\langle\phi\rangle = -v$ surrounded by a region where $\langle\phi\rangle = +v$, we can lower the energy by shrinking the region where $\langle\phi\rangle = -v$ to zero decreasing the surface area and hence the energy so that the whole universe is in the same state.

Secondly, we look at a complex scalar field i.e. ϕ is a two component field with the components Re ϕ and Im ϕ.

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu\phi)^*(\partial_\mu\phi) - \frac{\mu^2}{2}\phi^*\phi - \frac{\lambda}{4}(\phi^*\phi)^2 , \tag{5.27}$$

with $\mu^2 < 0$.

We minimise the potential

$$V(\phi,\phi^*) = \frac{\mu^2}{2}\phi^*\phi + \frac{\lambda}{4}(\phi^*\phi)^2, \qquad \frac{\partial V}{\partial\phi} = \frac{\mu^2}{2}\phi^* + \frac{\lambda}{2}\phi^*(\phi^*\phi) = 0 \tag{5.28}$$

i.e. $\phi^* = 0$ (a local maximum) or $|\phi|^2 = -\mu^2/\lambda \equiv v^2$.

* Large perturbations about the minimum will not interest us. They can only occur is systems at very high energy density e.g. in very dense material such as neutron stars. Otherwise there are only minor effects due to barrier penetration.

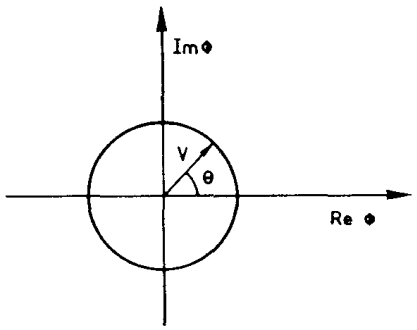


Fig. 5.2.

The minimum fixes $|\phi|^2$ to be v^2, i.e. the minimum is a circle in the ϕ plane.

Since θ (the phase angle of the complex field) is undetermined, we can fix the vacuum to be at any θ we wish. We will take $\theta = 0$. In order to make perturbations about this vacuum we now make the substitution

$$\phi = e^{i\xi/v}(v + \eta) , \tag{5.29}$$

where ξ represents perturbations in the θ direction and η represents perturbations in the radial direction. For small ξ, η this reduces to

$$\phi = v + \eta + i\xi , \tag{5.30}$$

which is equivalent to doing perturbations about $\mathrm{Re}\,\phi = v$ and $\mathrm{Im}\,\phi = 0$. Substituting this into (5.27) we get

$$\mathcal{L} = \tfrac{1}{2}[(\partial_\mu \xi)^2 + (\partial_\mu \eta)^2] - \frac{\lambda v^2}{4}\eta^2 + \text{cubic and higher order terms} . \tag{5.31}$$

This tells us that we have a particle η with mass $\lambda v^2/2$. This mass is a consequence of trying to displace the η against the restoring forces of the potential (the potential is like fig. 5.1(b) in the η direction). The ξ particle has no mass -- it is known as a Goldstone boson; it corresponds to displacements around the minimum surface (i.e. around the circle in fig. 5.2) where there is no restoring force, since the potential is flat.

We can ask in this case, what happens if $\langle\phi\rangle$ takes different phase values in different places A and B. In between, we would like to keep the gradient of ϕ as small as possible in order to keep the energy as small as possible; but if A and B are far apart we can manage this because $\langle\phi\rangle$ can continuously vary as we go from A and B. As A and B get infinitely far apart, the gradients tend

to zero and so the energy stored is zero. We can consider this changing of ϕ in the vacuum as a long wavelength excitation; as the wavelength tends to infinity (A and B tend to infinite separation) the energy tends to zero, so this excitation in the vacuum will correspond to a zero mass particle.

The need for such a massless particle appears to be general, and not dependent on our specific example. If the original Lagrangian has symmetry, this symmetry may be broken by the solution of minimum energy in the real world (the physical vacuum). But if the symmetry is represented by a continuous variable (like a rotation of phase) the "direction" of breaking can be slightly different in different places and waves due to perturbations in this direction must always be possible. In general, little energy is associated with long wave disturbances and we have (after quantising these perturbation waves) particles of necessarily zero mass – called Goldstone bosons. There are, however, cases where the variation of direction generates a current or charge density of some kind with which there are long range forces associated (i.e. r^{-1} potentials). Then the large contribution of large volumes to the energy of interaction in the long wavelength waves, increases the energy ω of the long waves to a finite value; the quantised excitations are now of finite energy as the wavenumber $k \to 0$ and hence of finite mass. (This "mass" generation by long range force is familiar in solid state physics where density variations of neutral molecules give rise to phonons with a dispersion $\omega = C_s k$, but compressional oscillations of charges like electrons give rise to plasma waves with a dispersion $\omega = \sqrt{\omega p^2 + k^2}$ (ωp = a constant) due to the long range Coulomb interaction between the charge densities.)

Since later we shall want to use symmetry breaking to explain how mass terms arise and since zero mass Goldstone bosons are not found, we whall have to add long range interactions (of zero mass Yang–Mills fields) to give them mass by this mechanism, called the Higgs–Kibble mechanism [4,5,14]. If there is more than one way to vary ϕ, while keeping the energy a minimum then there will be more than one Goldstone boson. The following example illustrates this. In SU(2), take an isovector $\boldsymbol{\phi}$; the minimum of the potential $V(\phi^+\phi)$ will occur for some non zero magnitude of ϕ, but its direction is undetermined and we can take $\langle\phi\rangle = \begin{pmatrix} 0 \\ 0 \\ v \end{pmatrix}$. In order to look at small perturbations about the minimum we consider (by analogy with the previous example)

$$\boldsymbol{\phi} = \exp\left(\frac{i\xi_1 L_1}{v}\right)\exp\left(\frac{i\xi_2 L_2}{v}\right)\begin{pmatrix} 0 \\ 0 \\ v+\eta \end{pmatrix}, \tag{5.32}$$

where L_1 and L_2 are two of the three SU(2) generators. Now the minimum is on the surface of a sphere (cf. the circle in fig. 5.2) there are now two inde-

pendent directions in which the potential is flat viz. the ξ_1 and ξ_2. The η meson has a mass corresponding to radial displacements off the sphere, in which direction $V(\phi^+\phi)$ increases.

In general if the Lagrangian is invariant under a group G but the vacuum has a lower symmetry i.e. it is invariant under a group G′ where $G' < G$. Then there will be massless bosons whose number n = dim G − dim G′. In the SU(2) example above, the vacuum is invariant under U(1) so that the number of massless mesons is $3-1=2$.

We give another example of this phenomenon which has more relevance to the physical world. Suppose the u and d quarks have zero mass. Then the Lagrangian

$$\mathcal{L} = \bar{\psi}_u i\partial\!\!\!/ \psi_u + \bar{\psi}_d i\partial\!\!\!/ \psi_d + \text{chirally invariant interaction terms}, \tag{5.33}$$

is invariant under the transformation

$$\psi \to e^{-ib\gamma_5}\psi, \qquad \bar{\psi} \to \bar{\psi} e^{-ib\gamma_5} \tag{5.34}$$

which is called a "chiral transformation" for arbitrary constant b. We can see this easily, as γ_5 anticommutes with the other γ matrices. Note however that a mass term would not be invariant

$$\bar{\psi} m \psi \to \bar{\psi} m e^{-2ib\gamma_5}\psi, \tag{5.35}$$

but a coupling to a vector potential via γ_μ would be invariant. Thus unless the symmetry is broken all the solutions of the Lagrangian must be chirally invariant and so, for example, the proton would have to be massless or parity doubled (because the chiral transformation changes the parity of a wavefunction). Experimentally the physical world does not have this symmetry. Therefore this symmetry must be broken and a Goldstone boson must exist if (5.33) is valid. No such massless particle exists but it would have to be a pseudoscalar meson, and therefore there is a temptation to associate it with the pion. Unfortunately the pion is not massless – but it is very light and people have therefore concluded the following. It might be that the u and d quarks have a small mass and therefore the chiral symmetry is only approximate; this would then possibly give a small mass to the pion. If this is the case, we can deduce a number of things about the pion couplings; these relations, such as the Goldberger–Treiman relation [e.g. 15], known as PCAC, are approximately verified experimentally. (We could expect another low mass boson coupled to $u\bar{u} + d\bar{d}$ i.e. a pseudoscalar but with isospin equal to zero. This is not found, people feel that the η' is too heavy, and how this difficulty can be resolved has always been a fascinating problem (the problem of the ninth pseudoscalar boson in SU(3)). I am sorry to find that the length of this course is too short to permit

me to discuss it along with many other interesting things I had hoped to discuss).

Let us return to vector fields. As we have so far discussed them, only the massless ones seem to make sense quantum mechanically. These have long range forces and we might therefore expect that the presence of these fields would eliminate the Goldstone boson when we break the symmetry. This is indeed the case as we will now show. Take a Lagrangian describing charged scalar particles interacting with photons

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu + iA_\mu)\phi^*(\partial_\mu - iA_\mu)\phi - \frac{\mu^2}{2}\phi^*\phi - \frac{\lambda}{4}(\phi^*\phi)^2 - \tfrac{1}{4}F_{\mu\nu}F_{\mu\nu}, \quad (5.36)$$

with $\mu^2 < 0$.

This is just the same as the Lagrangian (5.27) with electromagnetism added. We make the same substitution as before viz. $\phi = e^{i\xi/v}(v + \eta)$. Again we take the minimum of the potential to be v, giving

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu \xi)^2 + \tfrac{1}{2}(\partial_\mu \eta)^2 - \frac{\lambda v^2}{4}\eta^2 + \tfrac{1}{2}v^2A^2 - vA_\mu \partial_\mu \xi - \tfrac{1}{4}F_{\mu\nu}F_{\mu\nu}. \quad (5.37)$$

It looks as if the A_μ field has acquired a mass (see the term $\frac{1}{2}v^2A^2$) but there is a peculiar term where A_μ is coupled to $\partial_\mu \xi$ but the coupling is really $\frac{1}{2}v^2(A_\mu - (\partial_\mu \xi / v))^2$ so we can remove this term by doing a gauge transformation

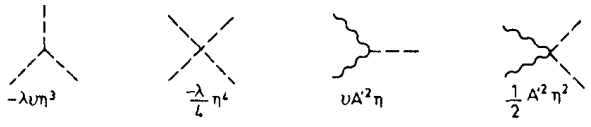
$$\phi' = e^{-i\xi/v}\phi, \qquad A'_\mu = A_\mu - (\partial_\mu \xi / v). \quad (5.38)$$

The Lagrangian then becomes

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu \eta)^2 - \frac{\lambda v^2}{4}\eta^2 + \tfrac{1}{2}v^2A'^2 - \lambda v\eta^3 - \frac{\lambda}{4}\eta^4$$

$$+ \tfrac{1}{2}A'^2\eta^2 + vA'^2\eta - \tfrac{1}{4}F_{\mu\nu}F_{\mu\nu} \quad (5.39)$$

($F_{\mu\nu}$ is invariant under this gauge transformation). In this Lagrangian we again have a massive field A'_μ but now the field ξ has disappeared altogether and all the interactions are cubic or higher order. This phenomenon is called the Higgs mechanism. (There is no term like $A_\mu \partial_\mu \xi$.) These interactions may be represented by the following diagrams



(5.40)

where ------ is an η propagator; ~~~~~ is an A'_μ propagator.

One could say: haven't we lost a degree of freedom (ξ) between Lagrangians (5.36) and (5.39)? No, because the theory is no longer gauge invariant electrodynamics – the free vector particle A'_μ now has three polarisation states since it has a mass, whereas the massless A_μ only had two polarisation states. There are four dynamical degrees of freedom in each case. We lost the gauge invariance when we made the explicit gauge transformation to eliminate ξ and a gauge transformation on (5.39) will not leave it invariant; in particular, it will bring back a term of the form $A_\mu \partial_\mu \xi$ where ξ is the gauge parameter.

We will now give a physical example of this phenomenon. The example is superconductivity and it is discussed in detail in [16]; here we shall merely give an outline. Superconductivity has the same properties as described above* except that it is non-relativistic. In the case of a metal at low temperatures, the electrons form (Cooper) pairs of opposite spin in such a way that these pairs act as bosons. Let ψ be the wave function for one of these bosons; it satisfies the Schroedinger equation. We consider the case of ψ interacting with an electromagnetic field

$$\frac{1}{2m}(\mathrm{i}\underline{\nabla} - \mathrm{e}\underline{A})^2 \psi = E\psi \,, \tag{5.41}$$

where E is a constant and e is the charge of a pair of electrons. Since the ψ is a boson, many pairs of electrons can be in the same state. The electromagnetic current is

$$j = \frac{1}{2m}[[(\mathrm{i}\underline{\nabla} - \mathrm{e}\underline{A})\psi]^* \psi + [\psi^* (\mathrm{i}\underline{\nabla} - \mathrm{e}\underline{A})\psi]\,] \,. \tag{5.42}$$

Suppose that in the absence of $\underline{A}$ there is no current but that all particles are in the same state because of the Bose condensation. This j represents not only the probability current of one particle, but when multiplied by e and $N/2$ (the number of particles) represents the physical electric current of the bosons. When very many bosons are in the same state, the wave function acquires a real physical significance, just as the photon wave function becomes the physically real (gauge transformations excepted) $\underline{A}\,(x, t)$ of classical mechanics and electromagnetism (Maxwell theory) when there are sufficient numbers of photons in the same state to make a "real" light wave. Turn on a very small field $\underline{A}$. To a first approximation, in many cases (those which yield superconductivity) the wavefunction is unchanged because so many interacting bosons are in the same state; so that the current we get is

* The Higgs Lagrangian (5.36) in the static case is identical to the Ginzburg–Landau free energy in the theory of type II superconductors [17].

$$j = \underline{A}\psi^*\psi/m\,. \tag{5.43}$$

But $\psi^*\psi$ is simply the density of bosons; call this $N/2$ (N is the density of the electrons). The current is

$$\underline{j}_{\text{super}} = \frac{N}{2} - \frac{e^2\underline{A}}{m}\,, \tag{5.44}$$

i.e. the current in a superconductor is proportional to the vector potential. The Maxwell equation

$$\nabla^2\underline{A} = \underline{j}, \tag{5.45}$$

becomes

$$\nabla^2\underline{A} = \underline{j}_{\text{external}} + \underline{j}_{\text{superconducting}}\,, \tag{5.46}$$

i.e.

$$(\nabla^2 - \Lambda)\underline{A} = \underline{j}_{\text{external}}\,, \tag{5.47}$$

where $\Lambda = Ne^2/2m$ and is called the London constant. This is the equation which describes the behaviour of a vector particle with a mass. The theory has spontaneously acquired a mass exactly as in the Higgs case. Notice that we have also lost the gauge invariance as we did in the Higgs case; we have chosen the gauge where the equations describing the physical properties of the theory are in simplest form.

Recall that when we considered the geometrical significance of gauge invariance in sect. 3 we saw how the gauge information could be carried from place to place by a particle. In this problem we have merely chosen the gauge which is carried by the electrons and it is clear that this is the natural gauge for this problem; this is why the equations appear simpler in this gauge.

We shall now look at spontaneous symmetry breakdown in Yang–Mills theories. We look at two cases

(i) Isospinor scalar particles. We construct the same $V(\phi)$ as we had earlier

$$V(\boldsymbol{\phi}) = \frac{\mu^2}{2}\phi_i^+\phi_i + \frac{\lambda}{4}(\phi_i^+\phi_i)^2\,, \tag{5.48}$$

where $i = 1, 2$. Let us assume that at some point in space

$$\boldsymbol{\phi} = v\begin{pmatrix}1\\0\end{pmatrix}. \tag{5.49}$$

If ϕ is pointing in the three direction of the three dimensional space of SU(2), there will no longer be gauge independence in the 1 and 2 directions since if

we rotate about the 1 or 2 directions ϕ changes, so we expect that two components of the A_μ field pick up mass. But what about the 3 direction? When we rotate about the 3 direction, a spinor gets multiplied by a phase i.e.

$$\begin{pmatrix}1\\0\end{pmatrix} \to \begin{pmatrix}e^{i\theta/2}\\0\end{pmatrix}, \tag{5.50}$$

for a rotation through θ. So we do not have any gauge freedom left and therefore all three components of A_μ pick up a mass.

Mathematically

$$\mathcal{L} = -\frac{1}{4g^2}E_{\mu\nu}E_{\mu\nu} + \tfrac{1}{2}(\partial_\mu + i\tau\cdot A)\boldsymbol{\phi}^+(\partial_\mu - i\tau\cdot A)\boldsymbol{\phi} - V(\boldsymbol{\phi}). \tag{5.51}$$

Put $\phi = \binom{1}{0} v$ and look at the mass term for the A's. This is

$$\tfrac{1}{2}v^2(1 \quad 0)(\tau\cdot A)(\tau\cdot A)\begin{pmatrix}1\\0\end{pmatrix} = \frac{v^2}{8}[A_1^2 + A_2^2 + A_3^2]. \tag{5.52}$$

So we see that all three components of A have acquired the same mass.

(ii) Isovector scalar particles

$$\boldsymbol{\phi} = \begin{pmatrix}\phi_1\\\phi_2\\\phi_3\end{pmatrix}.$$

The potential has the same form as (5.48) with $i = 1, 2, 3$. Let the vacuum expectation value of ϕ, $\langle\phi\rangle = \begin{pmatrix}0\\0\\1\end{pmatrix} v$, define the 3 direction in isospace. Carrying this ϕ particle around tells us where the 3 directions is everywhere. Hence, we lose the freedom to rotate the ϕ field about the 1 and 2 directions and therefore A^1 and A^2 will become massive – we have broken the gauge invariance in the 1 and 2 directions. But if we rotate by θ about the 3 direction the 1 component is multiplied by $e^{i\theta}$, the 2 component is multiplied by $e^{-i\theta}$, and the 3 component is left unchanged i.e. we still have a freedom of gauge rotation about the 3 direction. A^3 will still have zero mass.

Mathematically,

$$\mathcal{L} = \frac{1}{4g^2}E_{\mu\nu}E_{\mu\nu} + \tfrac{1}{2}[i(\partial_\mu + A_\mu X)\boldsymbol{\phi}]^+[i(\partial_\mu + A_\mu X)\boldsymbol{\phi}] - V(\phi) \tag{5.53}$$

Substituting $\phi = \begin{pmatrix}0\\0\\1\end{pmatrix} v$ we get the mass term for A:

$$\frac{v^2}{2}[(A^1)^2 + (A^2)^2] \tag{5.54}$$

i.e. A^1 and A^2 have the same mass and A^3 is massless, as expected. One might ask – why do we use scalar particles to break the symmetry? They are of course the simplest to write down. But this "superconductivity" effect can be generated in many ways (as indeed in real electrodynamics where pairs of fermions do it), and it is possible that the symmetry breaking if it occurs, in say, weak interactions or other places in physics, may have a more complex mechanism than the Higgs scalar method. But to-day we have a severe restriction on the meaningful theories we can write down i.e. that they are relativistic, quantum mechanical, and renormalisable. If all these restrictions are imposed it looks as if only the Higgs method can be formulated at present.

6. Quantisation

6.1. Philosophy

Before we commence a detailed study of the quantisation of a Yang–Mills fields, we shall describe the various approaches to quantising a classical theory. Fig. 6.1 illustrate a general schema for quantising particle mechanics.

We begin with a Lagrangian from which we can deduce a classical Hamiltonian, written in terms of $q(t)$ and its conjugate momentum $p(t)$, defined by $p(t) = \mathrm{d}\mathcal{L}/\mathrm{d}\dot{q}(t)$. There are two routes by which we can quantise the theory.

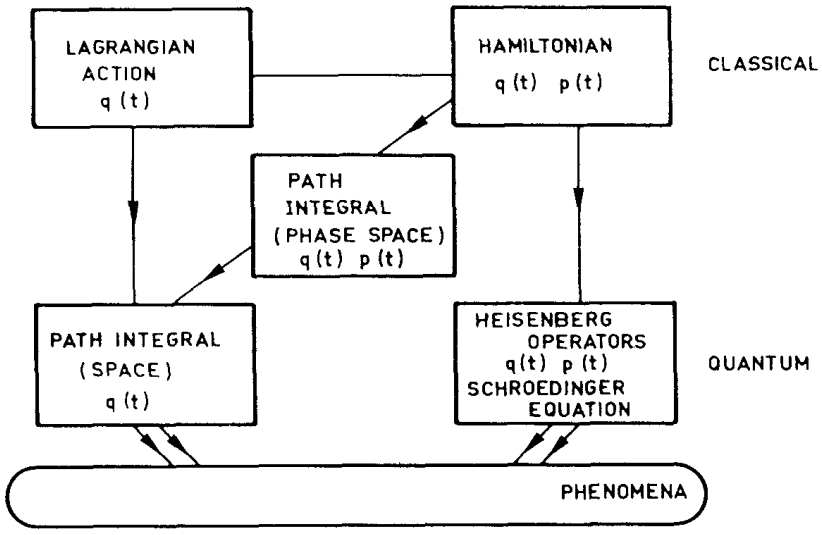


Fig. 6.1.

From the Hamiltonian we can define operators associated with $p(t)$ and $q(t)$ and hence obtain either the Heisenberg or Schroedinger pictures; from these we can deduce results which predict the behaviour of phenomena in the physical world. There may be a problem with operator ordering when we go from the classical to the quantum theory. The alternative approach is to start with the Lagrangian and introduce a path integral which associates a certain amplitude with each trajectory in space [18]; this enables us to proceed directly to calculate the consequences of the quantum theory. (There exists an alternative path integral method, not much used these days, devised by DeWitt–Morette [19] and Garrod [20] in which one can construct a path integral in phase space directly from the Hamiltonian. This resolves many questions of the order of operators in the Hamiltonian. From it the ordinary path integral is easily obtained.) All methods lead to the same consequence physically, but each method has its advantages and disadvantages and we choose which ever is the more convenient for the problem at hand e.g. spin $\frac{1}{2}$ is very awkward to handle in path integral approach.

Consider now the situation when we try to quantise a field theory (fig. 6.2).

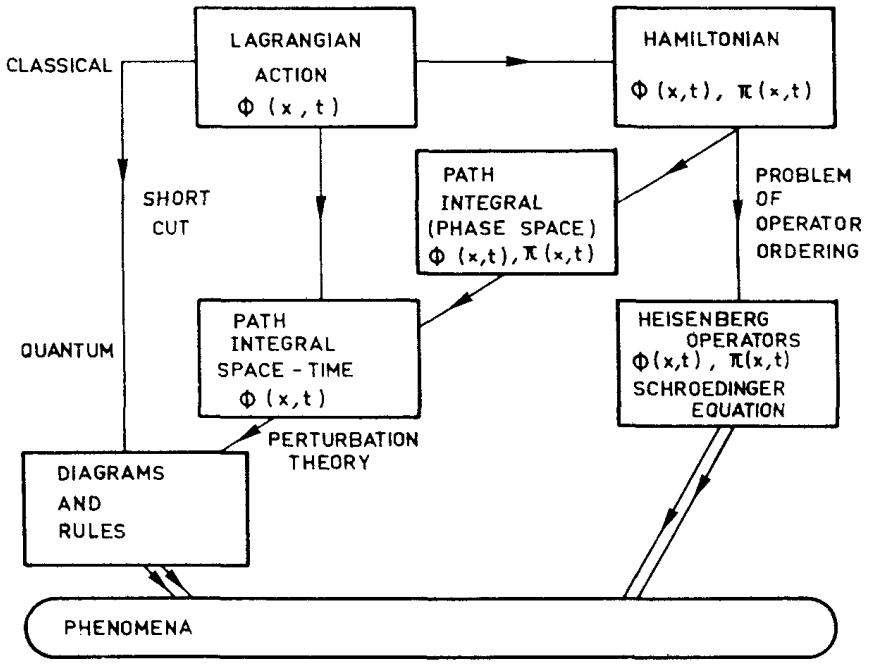


Fig. 6.2.

The classical variables $q(t)$ and $p(t)$ are replaced by $\phi(x, t)$ and its conjugate momentum $\pi(x, t)$ defined by $\pi(x, t) = \partial\mathcal{L}/\partial\dot{\phi}(x, t)$. The situation is similar to that in quantising a particle theory and historically people like Heisenberg and Dirac [e.g. 21] followed the Hamiltonian approach. This leads to $\phi(x, t)$ and $\pi(x, t)$ being operators and again we have a problem with operator ordering in the more complicated field theories. There is also a Schroedinger approach which people don't use much these days in which the vacuum has a wave function in terms of the co-ordinates i.e. an explicit functional of the field function $\phi(x)$. Alternatively we can proceed by the path integral method; the path is now a function of $\phi(x, t)$. Because of difficulties in calculating with the theory we usually proceed by perturbative methods. To do this we normally use diagrams and rules, all of which can easily be deduced from the path integral formulation. We can also get these rules from a Heisenberg approach, but it is more difficult. Nowadays we know of an obvious and simple-minded short cut to get straight from the Lagrangian to the diagram rules. Some people who are not sufficiently acquainted with the theory, think that the rules are all there is, and then say that the theory is only defined by a perturbation expansion. It is true that we can only calculate things by perturbation theory, but this may be only a limitation of the era and in any case there are certain things that we can deduce from the Lagrangian without using perturbation theory. (For non-relativistic field theories, such as those that arise in solid state physics, we are not at all limited to perturbation theory, and many methods and solutions for large or intermediate couplings are known.)

In a field theory there is a problem of renormalisability, because, when we calculate diagrams with closed loops we get infinite answers. When we go through the Heisenberg approach, the theory is not manifestly Lorentz invariant; but in the path integral approach it is, so for this reason the latter approach is to be preferred when we attempt to renormalise the theory. The difficulty is to keep the renormalisation process Lorentz invariant when the form of the equations is not manifestly invariant. There is difficulty in the path integral approach, however, and that is concerned with the inclusion of fermions. The path integral method doesn't work in this case: but when people went through the operator approach they found that there were only differences in signs when fermions were introduced, and so minor were these differences that they forced the path integral formalism to work by introducing Grassmann algebras.

In attempting to quantise Yang–Mills theory, Schwinger [22], after a lot of hard work, found the Hamiltonian from the Lagrangian, but an attempt to proceed with the Hamiltonian approach ran into serious difficulties with the ordering of operators and progress in this direction ceased. I took the short cut

(in fig. 6.2) [23] and found that there are certain complications in the diagrams at the one loop level; but I got round these by introducing a contribution from a fictitious particle. The correct rules for this particle were first worked out by de Witt [24] and subsequently understood in a more general way by Fadde'ev and Popov [26]. However I could only do it for one closed loop; if there were two or more closed loops I didn't know what to do – I sort of half understood it. Fadde'ev and Popov straightened this problem out by going via the path integral method and discovered that to all orders we have to add a closed loop of scalar particle with Fermi statistics for every closed loop with a Yang–Mills vector meson, they also went via the Morett path integral approach using the Hamiltonian derived by Schwinger. The problem of renormalisability was solved by 't Hooft [25] using his method of dimensional regularisation.

Yang–Mills theory is often presented as being complicated, but now that all this work is done and proofs proved, it is really not much more complicated than QED apart from the more complex algebra which is involved; and a bit more care is needed with the gauge invariance. We just add the contribution of the ghost and regulate by the dimensional regularisation scheme.

It would be sensible now to give the correct and complete theory, say as outlined in Abers and Lee [10] and many students might prefer this. But to get a clear feeling for the need for the Fadde'ev–Popov ghost, we can contrast the theory with QED by asking what happens if we just plough along and make diagram rules by direct analogy with QED. This is a sort of "damn the torpedoes, full speed ahead" approach. We will discover that we are hit by a torpedo, but let's try it anyhow.

We will attempt to calculate with the theory by first deriving the diagram rules directly from the Lagrangian.

6.2. Derivation of the rules for diagrams

We shall assume that the reader is familiar with the methods of deriving the rules from the Lagrangian and so we will merely give an outline of the derivation as we go along. Consider the following model Lagrangian for spin $\frac{1}{2}$ isospinors ψ and spin 0 isovectors $\boldsymbol{\phi}$ interacting with Yang–Mills fields A_μ

$$\mathcal{L} = -\tfrac{1}{4}E_{\mu\nu}E_{\mu\nu} + \bar{\psi}(\mathrm{i}\partial\!\!\!/ + gA\!\!\!/\cdot\boldsymbol{\tau})\psi - m\bar{\psi}\psi$$
$$+ \tfrac{1}{2}[\mathrm{i}(\partial_\mu + gA_\mu \times)\boldsymbol{\phi}]^+[\mathrm{i}(\partial_\mu + gA_\mu \times)\boldsymbol{\phi}] - \frac{M^2}{2}\boldsymbol{\phi}^+\boldsymbol{\phi}, \tag{6.1}$$

where we have rescaled A_μ to gA_μ and so now

$$E_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + gA_\mu \times A_\nu . \tag{6.2}$$

We separate the Lagrangian into terms of second order and terms of third and higher orders in the fields. The second order terms give the propagators and the higher order terms give the interactions:

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}\cdot F_{\mu\nu} + \bar{\psi}(i\gamma_\mu\partial_\mu - m)\psi + \tfrac{1}{2}[(\partial_\mu\boldsymbol{\phi})^+(\partial_\mu\boldsymbol{\phi}) - M^2\boldsymbol{\phi}^+\boldsymbol{\phi}] \qquad \text{2nd order}$$

$$+ g\bar{\psi}\gamma_\mu A_\mu\cdot\boldsymbol{\tau}\psi - g(A_\mu\times A_\nu)\cdot\partial_\mu A_\nu + \tfrac{1}{2}g[\partial_\mu\boldsymbol{\phi}^+\cdot(A_\mu\times\boldsymbol{\phi}) + \partial_\mu\boldsymbol{\phi}\cdot(A_\mu\times\boldsymbol{\phi}^+)] \qquad \text{3rd order}$$

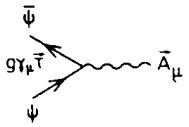
$$+ \tfrac{1}{2}g^2(A_\mu\times\boldsymbol{\phi}^+)\cdot(A_\mu\times\boldsymbol{\phi}) - \tfrac{1}{4}g^2(A_\mu\times A_\nu)\cdot(A_\mu\times A_\nu) \qquad \text{4th order} \tag{6.3}$$

where we have retained $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ to pull out second order terms in A_μ which will give the vector field propagator.

As there is some problem in deriving the propagator for the A_μ field, we will first derive the diagram rules for the interaction terms. Consider the term

$$g\bar{\psi}\gamma_\mu A_\mu\cdot\boldsymbol{\tau}\psi\,, \tag{6.4}$$

this corresponds to the following interaction — ψ coming in, absorbing or emitting an A_μ and going out as a $\bar{\psi}$, viz.

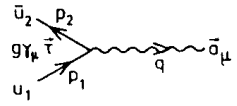


(6.5)

In co-ordinate space we take

$$\psi = u_1 e^{-ip_1 x}, \qquad \bar{\psi} = \bar{u}_2 e^{+ip_2 x}, \qquad A_\mu = a_\mu e^{+iqx}\,, \tag{6.6}$$

i.e. free waves. Going to momentum space the diagram becomes



(6.7)

where the momenta carried by the particles are indicated on the diagram. We can now read off the vertex; it is

$$g\bar{u}_2(\gamma_\mu\boldsymbol{\tau}\cdot a_\mu)u_1\,. \tag{6.8}$$

This, except for the presence of the $\boldsymbol{\tau}$ matrix, is identical to QED. In fact, as in QED, we do not need actually to write the u_1, $\bar{u}_2$ spinors for virtual fer-

mions, but just string the Dirac (and τ) matrices together in a product in order, as we come to them following a fermion line.

We now consider, in a similar fashion, the third order (in the fields) term for the scalar field ϕ interacting with the Yang–Mills field A_μ

$$+\tfrac{1}{2}g\{\partial_\mu\boldsymbol{\phi}^+\cdot(A_\mu\times\boldsymbol{\phi})+\partial_\mu\boldsymbol{\phi}\cdot(A_\mu\times\boldsymbol{\phi}^+)\}. \tag{6.9}$$

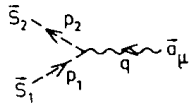
Putting in, as before, the free fields

$$\boldsymbol{\phi}=S_1\,e^{-ip_1x},\qquad \boldsymbol{\phi}^+=S_2\,e^{+ip_2x},\qquad A_\mu=a_\mu\,e^{-iqx}, \tag{6.10}$$

we get

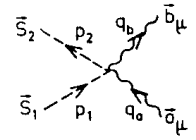
$$+\tfrac{1}{2}g\{+ip_{2\mu}S_2\cdot(a_\mu\times S_1)-ip_{1\mu}S_1\cdot(a_\mu\times S_2)\}\exp\{-i(p_1-p_2+q)x\}, \tag{6.11}$$

then going to momentum space we have



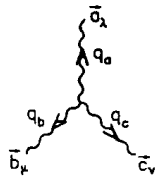
$$ig(p_1+p_2)_\mu[S_2\cdot(a_\mu\times S_1)] \tag{6.12}$$

Similarly, the term $\frac{1}{2}g^2(A_\mu\times\boldsymbol{\phi}^+)\cdot(A_\mu\times\boldsymbol{\phi})$ gives



$$g^2[(b_\mu\times S_2)\cdot(a_\mu\times S_1)+(a_\mu\times S_2)\cdot(b_\mu\times S_1)]. \tag{6.13}$$

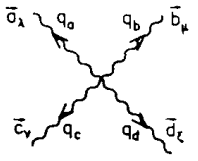
So far, these interaction are similar to (apart from the isospin labels) those in QED. But in Yang–Mills theory there are two extra vertices which have no QED analogue (they arise because the A_μ field itself carries "charge"). The term $-g(A_\mu\times A_\nu)\cdot\partial_\mu A_\nu$ gives, in momentum space, $-g(a_\mu\times b_\nu)\cdot c_\nu q_{c_\mu}$ when we substitute $a\,e^{iq_ax}$ for the first, $b_\mu e^{iq_bx}$ for the second, and $c_\nu e^{iq_cx}$ for the third with momenta q_a, q_b, q_c. Hence altogether it gives, taking account of permutations and with a bit of re-arranging of the dots and crosses:



$$-g\{(q_a-q_c)_\mu b_\mu\cdot(c_\nu\times a_\nu)+(q_c-q_b)_\mu a_\mu\cdot(b_\nu\times c_\nu)$$
$$+(q_b-q_a)_\mu c_\mu\cdot(a_\nu\times b_\nu)\}. \tag{6.14}$$

There are six terms altogether since each of the three A_μ's can be either a_μ, b_μ or c_μ in every possibly way.

The term $-\frac{1}{4}g^2(A_\mu \times A_\nu)\cdot(A_\mu \times A_\nu)$ gives



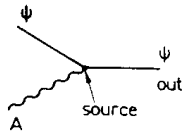
$$-\frac{g^2}{4}\{(a_\mu \times b_\nu)\cdot(c_\mu \times d_\nu) + \text{three other symmetric combinations}\}, \tag{6.15}$$

which comes to a total of $-g^2(a_\mu \times b_\nu)\cdot(c_\mu \times d_\nu)$ by a bit of rearranging.

We must finally derive the propagators for the theory from the quadratic terms in (6.3). The propagators for ψ and ϕ are easy to get. The equation of motion for ψ is (obtained from the Lagrangian by varying with respect to $\bar{\psi}$)

$$(i\partial\!\!\!/ - m)\psi = J\!\!\!/, \tag{6.16}$$

where J is some source for ψ (whose exact form is irrelevant, e.g. some non-linear combination of other fields like $A\cdot\tau\psi$, that in the diagrams generate a source away from which a particle ψ is to propagate)



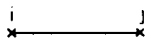
So symbolically

$$\psi = \frac{1}{i\partial\!\!\!/ - m} J\!\!\!/. \tag{6.17}$$

In momentum space this is

$$\psi_p = \frac{1}{p\!\!\!/ - m} J\!\!\!/_p \tag{6.18}$$

This gives the ψ which enters into a source term for another interaction, and so each virtual ψ line brings in a factor $(p\!\!\!/ - m)^{-1}$. This then defines the propagator



$$\frac{\delta_{ij}}{p\!\!\!/ - m}. \tag{6.19}$$

The i and j are isospin indices. The δ_{ij} is present since ψ must couple to the same isospin at both ends.

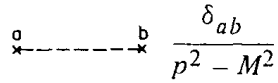
For the ϕ field, the equation of motion is

$$(\Box + M^2)\phi = -J' . \tag{6.20}$$

Therefore

$$\phi_p = \frac{1}{p^2 - M^2} J'_p , \tag{6.21}$$

where we have again transformed to momentum space. This gives the ϕ propagator



$$\frac{\delta_{ab}}{p^2 - M^2} . \tag{6.22}$$

where the a and b are isospin indices.

Finally consider the A_μ propagator; the equation of motion is

$$\partial_\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) = -J'_\nu , \tag{6.23}$$

i.e.

$$\Box A_\nu - \partial_\nu(\partial_\mu A_\mu) = -J'_\nu \tag{6.24}$$

where J'_ν is the total current – matter plus contributions from the A_μ itself due to the third and higher order terms in the Lagrangian. We need to solve this equation to get the propagator. However in general we cannot solve it because by itself it is meaningless unless the divergence of J'_ν is zero, for the divergence of the left side is identically zero. We conclude therefore that

$$\partial_\nu J'_\nu = 0 ,$$

i.e.

$$\partial_\nu\{J_\nu + \partial_\mu(A_\mu \times A_\nu) + A_\mu \times E_{\mu\nu}\} = 0 . \tag{6.25}$$

This equation can indeed be verified using the relation (2.42).

When we attempt to obtain the propagator in QED, we do so by choosing

$$\partial_\mu A_\mu = 0 , \tag{6.26}$$

as we may do due to gauge invariance. The equation of motion then becomes simply

$$\Box A_\mu = -J_\mu , \tag{6.28}$$

and we can solve this to obtain the propagator viz.

$$A_\mu = \frac{\delta_{\mu\nu}}{p^2} J_\nu \,, \tag{6.29}$$

so the propagator is

$$\delta_{\mu\nu}/p^2 \,. \tag{6.30}$$

It turns out that when we use this and calculate diagrams in QED, the current we get from the diagrams is automatically conserved so that everything is alright, as it must be since

$$\partial_\mu \Box A_\mu = -\partial_\mu J_\mu \,, \tag{6.31}$$

from (6.28), and both sides are separately zero; so the gauge choice is self consistent. This is equivalent to replacing the action $-\frac{1}{4}\int F_\mu F_{\mu\nu} \mathrm{d}^4x$ by $-\frac{1}{2}\int(\partial_\nu A_\mu)(\partial_\nu A_\mu)\mathrm{d}^4x$, or in other words to adding an extra term $+\frac{1}{2}\int(\partial_\mu A_\mu)^2 \mathrm{d}^4x$ to the action, so the Lagrangian becomes simply $-\frac{1}{4}F_{\mu\nu}F_{\mu\nu} + \frac{1}{2}(\partial_\mu A_\mu)(\partial_\nu A_\nu)$. Take a flying guess and try to do the same thing for Yang–Mills – an excellent method of doing physics. The purpose of physics is to find out what's true, not to find out what you can prove. If you allow yourself the liberty of knowing things with different degrees of certainty, then you can know a lot more physics than if you have to prove everything. So in (6.24) we suppose that we can choose the gauge

$$\partial_\mu A_\mu = 0 \,, \tag{6.32}$$

and we then get the following for the equation of motion

$$\Box A_\nu = -J'_\nu \,. \tag{6.33}$$

We can invert this equation to get

$$A_\nu = \frac{\delta_{\nu\mu}}{\Box}(-J'_\mu) \tag{6.34}$$

which in momentum space is

$$(A_p)_\nu = \frac{\delta_{\nu\mu}}{p^2}(J'_p)_\mu \,. \tag{6.35}$$

This defines the propagator

$$\underset{a}{\mu} \sim\sim\sim \underset{b}{\nu} \quad \frac{\delta_{\mu\nu}\delta_{ab}}{p^2} \,. \tag{6.36}$$

Notice that if we take the divergence of both sides of (6.33), both sides are

separately zero, the left side by the gauge condition (6.32) and the right side by (6.25) so we appear to be self consistent. We have all the rules and can calculate processes and see whether we run up against any problems, such as gauge non invariance of a physical process. We will, indeed, run into problems when we consider Yang–Mills loops and shall see that we need to introduce something extra (the ghost particle) to cure them.

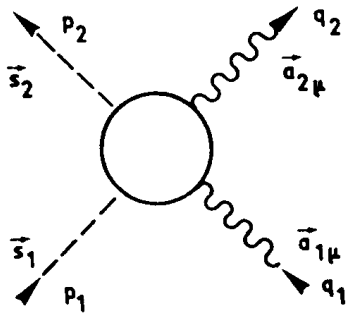
6.3. Explicit calculations of physical processes in Yang–Mills theory

As an application of the rules that we have derived, we shall now calculate several physical processes. We begin with Compton scattering for $I = 1$ scalon.

The process under consideration here is Compton scattering of a gluon A_μ and an isovector scalar particle ϕ(scalon) using the Lagrangian (6.1)

$$\text{gluon} + \text{scalon} \rightarrow \text{gluon} + \text{scalon} . \tag{6.37}$$

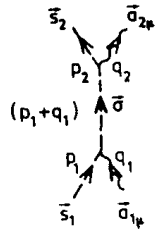
We define the momenta, polarisation vectors and isospin labels as follows



(6.38)

The momenta are p_1, p_2 and q_1, q_2, the isospin labels of the scalons are S_1 and S_2 and the polarisation vectors and isospin labels of the gluons are $a_{1\mu}$ and $a_{2\mu}$. To second order in the coupling constant there are four diagrams which contribute to this process, and we will calculate them in turn

(i)



(6.39)

where we have temporarily put an isospin index σ on the intermediate scalon to make it easier to apply our rules. The amplitude for this process is (we must sum over all isospin directions of σ as all will contribute to the diagram)

$$\sum_{\substack{\text{isospin}\\ \text{directions}}} g^2(2p_2+q_2)_\nu S_2\cdot(a_{2\nu}\times\sigma)\frac{1}{(p_1+q_1)^2-m^2}$$

$$\times\sigma\cdot(a_{1\mu}\times S_1)(2p_1+q_1)_\mu\,, \tag{6.40}$$

Re-writing $S_2\cdot(a_{2\nu}\times\sigma)$ as $(S_2\times a_{2\nu})\cdot\sigma$ and using the formula

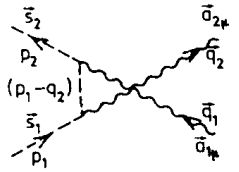
$$\sum_\sigma(A\cdot\sigma)(\sigma\cdot B)=A\cdot B\,, \tag{6.41}$$

we get for the contribution from this diagram

$$g^2\frac{(2p_2+q_2)_\nu(S_2\times a_{2\nu})\cdot(a_{1\mu}\times S_1)(2p_1+q_1)_\mu}{(p_1+q_1)^2-m^2}. \tag{6.42}$$

The second diagram is the same as (i) with the two gluons interchanged.

(ii)



(6.43)

The amplitude for this diagram can be obtained immediately from (i) if we notice that making the substitutions

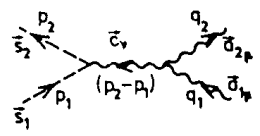
$$q_1\leftrightarrow -q_2,\qquad a_{1\mu}\leftrightarrow a_{2\mu}\,, \tag{6.44}$$

in (i) gives (ii). Therefore the contribution from (ii) is

$$g^2\frac{(2p_2-q_1)_\nu(S_2\times a_{1\nu})\cdot(a_{2\mu}\times S_1)(2p_1-q_2)_\mu}{(p_1-q_2)^2-m^2}. \tag{6.45}$$

The third diagram is

(iii)



(6.46)

$$\text{amplitude} = -\frac{(S_2 \times S_1)\cdot c_\lambda (p_1+p_2)_\lambda}{(p_1-p_2)^2}\{(q_1-p_1+p_2)_\nu a_{2\nu}\cdot(a_{1\mu}\times c_\mu)$$

$$+(-q_2-q_1)_\nu c_\nu \cdot (a_{2\mu}\times a_{1\mu}) + (p_1-p_2+q_2)_\nu a_{1\nu}\cdot(c_\mu \times a_{2\mu})\},$$

where c_ν is the polarisation vector of the intermediate gluon, and therefore we must sum over its directions in space and isospace using the formula

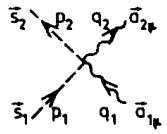
$$\sum_{c_\nu}(A\cdot c_\nu)(c_\mu \cdot B) = A\cdot B\delta_{\mu\nu}\,. \tag{6.47}$$

The amplitude for this diagram is then

$$-g^2\frac{(S_2\times S_1)}{(p_1-p_2)^2}\{(a_{2\nu}\times a_{1\mu})[(p_1+p_2)_\mu(2q_1-q_2)_\nu$$

$$+(p_1+p_2)_\nu(2q_2-q_1)_\mu] - (a_{2\mu}\times a_{1\mu})(p_1+p_2)_\lambda(q_1+q_2)_\lambda\}. \tag{6.48}$$

Finally we have the diagram



(6.49)

The amplitude is

$$g^2\{(S_2\times a_{2\mu})\cdot(S_1\times a_{1\mu}) + (S_2\times a_{1\mu})\cdot(S_1\times a_{2\mu})\}. \tag{6.50}$$

Adding the contributions from each of the diagrams, rearranging dot and cross vector products, putting the scalons on mass shell and using momentum conservation (but keeping the gluons off mass shell for the moment) we get the total amplitude for the process to be

$$(S_2\times a_{2\nu})\cdot(S_1\times a_{1\mu})\left\{\frac{(2p_2+q_2)_\nu(2p_1+q_1)_\mu}{2p_1q_1+q_1^2}\right.$$

$$-\frac{(2q_1-q_2)_\nu(p_1+p_2)_\mu + (p_1+p_2)_\nu(2q_2-q_1)_\mu}{(q_1-q_2)^2}$$

$$\left.+\frac{(p_1+p_2)\cdot(q_1+q_2)+(q_1-q_2)^2}{(q_1-q_2)^2}\delta_{\mu\nu}\right\} \tag{6.51}$$

equation continued on next page.

$$+(S_1 \times a_{2\nu}) \cdot (S_2 \times a_{1\mu}) \left\{ -\frac{(2p_1 - q_2)_\nu (2p_2 - q_1)_\mu}{-2p_2 q_2 + q_2^2} \right.$$

$$+\frac{(2q_1 - q_2)_\nu (p_1 + p_2)_\mu + (p_1 + p_2)_\nu (2q_2 - q_1)_\mu}{(q_1 - q_2)^2}$$

$$\left. +\frac{-(p_1 + p_2) \cdot (q_1 + q_2) + (q_1 - q_2)^2}{(q_1 - q_2)^2} \delta_{\mu\nu} \right\}. \qquad (6.51, \text{cont'd})$$

So what? What do we do with the answer? There are two possibilities – we can use it to calculate the Compton scattering – maybe we are interested in this; or we can use it to calculate the annihilation of a pair of scalons into a pair of gluons (see below).

In order to get the Compton scattering of real gluons we must put them on their mass shell i.e. $q_1^2 = q_2^2 = 0$ but this alone is not sufficient, as we can see as follows. Consider the equation of motion for the gluons in the absence of any sources (free gluons):

$$\partial_\mu (\partial_\mu A_\nu - \partial_\nu A_\mu) = 0 \qquad (6.52)$$

In momentum space, this becomes

$$q_\mu (q_\mu a_\nu - q_\nu a_\mu) = 0 \qquad (6.53)$$

i.e.

$$q^2 a_\nu - (q_\mu a_\mu) q_\nu = 0 \,. \qquad (6.54)$$

There are two solutions to this equation: either (a) $q^2 = 0$ and we must have $q_\mu a_\mu = 0$. I.e. for a Yang–Mills particle on its mass shell, the polarisation vector is perpendicular to the momentum; or (b) $q^2 \neq 0$; then $a_\nu = (q_\mu a_\mu / q^2) q_\nu$, i.e.

$$a_\nu = \alpha q_\nu \,. \qquad (6.55)$$

This corresponds in co-ordinate space to $A_\nu = \partial_\nu \chi$ which is a pure gradient. So we would perhaps expect that this field could be removed by an infinitesimal gauge transformation. Let us try this; to gauge it away we would have to have

$$0 = \partial_\nu \chi + \partial_\nu \alpha + (\partial_\nu \chi \times \alpha) \,.$$

Since $\partial_\nu \chi \times \alpha$ is higher order in the field, it can be neglected since we are working with a free wave and hence a_ν must be small (otherwise there would be self coupling terms on the right hand side of 6.52 which we neglected in order to get the free wave solution, since they are of second and higher order in the fields). We can therefore take α to be $-\chi$ and gauge the field away.

Case (b) cannot produce any physics. If we calculate an amplitude for a physical process first with a_μ, and then with a'_μ, where a'_μ is related to a_μ by an infinitesimal gauge transformation $a'_\mu = a_\mu + \partial_\mu \alpha$, there should be no difference in the answers since the theory is gauge invariant. The effect of a_μ and a'_μ is linear and this means that if we were to calculate with $\partial_\mu \alpha$ the amplitude must be zero.

So we can test for the correctness of the solution to any physical process by putting the polarisation vector a_μ of an external gluon equal to αq_μ where q_μ is its momentum, and checking that this gives zero. Note that we must carry out this test for a whole physical process; we should not expect that a single diagram by itself is gauge invariant as it is only part of an answer and by itself represents no physics.

In the Compton effect above, we gauge the incoming gluon, putting

$$a_{1\mu} = \alpha q_{1\mu} \tag{6.56}$$

Do we get zero? We make the substitution, and of course we don't get zero – because we made a couple of mistakes; however fiddling around a bit more and correcting the mistakes, the Compton amplitude (6.51) becomes

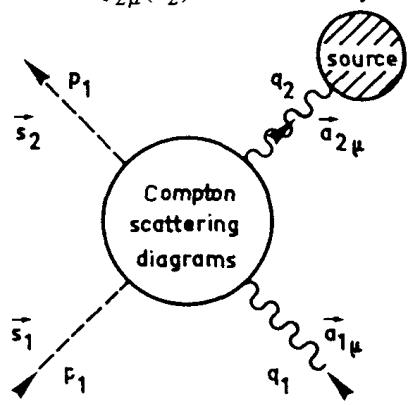
$$-g^2(S_1 \times \alpha) \cdot (S_2 \times j_{2\mu}(a_2))\left\{\frac{(2p_1 + q_1)_\mu}{(q_1 - q_2)^2}\right\}$$

$$-g^2(S_2 \times \alpha) \cdot (S_1 \times j_{2\mu}(a_2))\left\{\frac{(-2p_2 + q_1)_\mu}{(q_1 - q_2)^2}\right\}$$

where

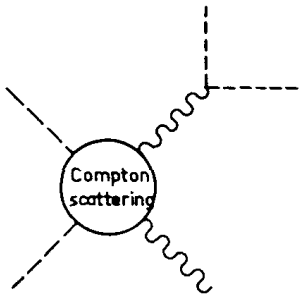
$$j_{2\mu}(a_2) = q_2^2 a_{2\mu} - q_{2\mu}(q_{2\nu} a_{1\nu}) \,. \tag{6.57}$$

Is this zero? Yes, because $a_{2\mu}$ is supposed to represent a free sourceless gluon, so the term $j_{2\mu}(a_2)$ must vanish by 6.54. If $a_{2\mu}$ is not sourceless



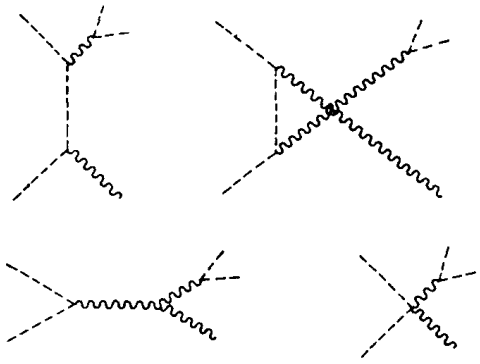
(6.58)

$j_{2\mu}(a_2)$ is not zero, so the process we have calculated does not seem to be gauge invariant. Is there a fault in the theory? Need the theory be gauge invariant if the $a_{2\mu}$ has come from a source? Yes, the theory must be – the fault lies in the fact that we are not now calculating a complete physical process. Consider as an example that the source of $a_{2\mu}$ is another scalon, i.e.



(6.59)

Then in addition to the diagrams



(6.60)

which contribute to (6.58), there are two additional diagrams which we must consider for the whole physical process, viz.

(6.61)

When the sum of all diagrams in (6.60) and (6.61) is considered and we put $a_{1\mu} = \alpha q_{1\mu}$, we will get zero. It is not surprising that we do not get zero for the diagrams (6.60) as they only constitute part of an answer to a physical process, and there is no reason why this partial answer should be gauge invariant. The diagrams (6.61) correspond to a modification of the source in

(6.58). Therefore, if we always only apply a gauge transformation to a physical process we will find that the amplitude is invariant (provided we stick with tree diagrams). Hence the diagram rules which we have constructed will work as long as we do not calculate any diagrams with closed loops. Before considering the difficulties with closed loops, we will consider the Compton scattering process further.

The formula for the real Compton effect is obtained from (6.51) by putting both gluons on their mass shells i.e.

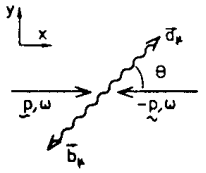
$$q_1^2 = q_2^2 = 0; \qquad q_{1\mu} a_{1\mu} = q_{2\mu} a_{2\mu} = 0 , \tag{6.62}$$

in accordance with the situation (a) earlier. We get

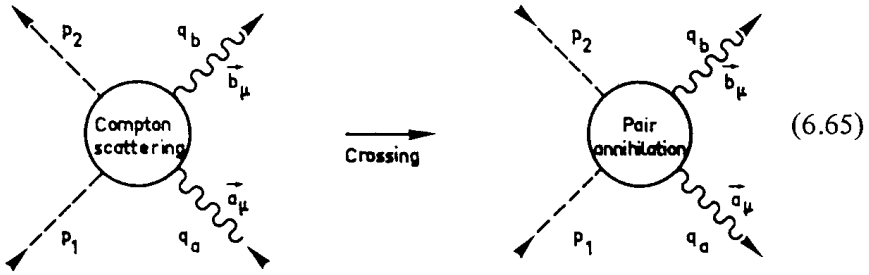
$$\frac{2g^2(S_2 \times a_{2\nu}) \cdot (S_1 \times a_{1\mu})}{q_1 \cdot q_2} \left\{ - \frac{(p_{1\mu} p_{2\nu})(p_{1\lambda} q_{2\lambda})}{(p_1 \cdot q_1)} \right.$$

$$\left. + (p_{2\mu} p_{1\nu}) - \delta_{\mu\nu}(p_{1\lambda} q_{1\lambda}) \right\} + \begin{Bmatrix} S_1 \leftrightarrow S_2 \\ p_1 \leftrightarrow -p_2 \end{Bmatrix} . \tag{6.63}$$

By taking the crossed diagrams of the Compton scattering process we obtain the process for scalon annihilation into gluons



in the center of mass system (6.64)



(6.65)

The amplitude for this process can be deduced immediately from (6.63); it is

$$\frac{2}{q_a \cdot q_b} \left[\frac{p_{1\mu} p_{2\nu}}{p_1 \cdot q_a} + \frac{p_{1\nu} p_{2\mu}}{p_1 \cdot q_b} + \delta_{\mu\nu} \right] \left[(p_1 \cdot q_b) A_{\mu\nu} + (p_1 \cdot q_a) B_{\mu\nu} \right] , \tag{6.66}$$

where

$$A_{\mu\nu} = -(S_1 \times a_\mu) \cdot (S_2 \times b_\nu), \qquad B_{\mu\nu} = -(S_1 \times b_\nu) \cdot (S_2 \times a_\mu). \tag{6.67}$$

In the centre of mass system, the incident scalons have three momenta p_i and $-p_i$ respectively and velocity v_i and $-v_i$ where

$$v_i = p_i/\omega = p_i/\sqrt{p^2 + M^2}\,. \tag{6.68}$$

The amplitude for this process is then

$$\left[\frac{2v_i v_j}{1 - v^2\cos^2\theta} - \delta_{ij}\right]\left[(1 + v\cos\theta)A_{ij} + (1 - v\cos\theta)B_{ij}\right] \tag{6.69}$$

We consider the possible polarisation states for the gluon: state (a) in the plane of the reaction viz.,

$$a_\mu = (\overset{t}{0}, -\overset{x}{\sin\theta}, \overset{y}{\cos\theta}, \overset{z}{0})\,\alpha\,, \tag{6.70}$$

state (b) perpensicular to the plane of the reaction viz.,

$$a_\mu = (\overset{t}{0}, \overset{x}{0}, \overset{y}{0}, \overset{z}{1})\,\alpha\,. \tag{6.71}$$

Apply (6.69) to the following polarisation combinations:

(i) both gluons in state (b)

$$\text{Amplitude} = [(1 + v\cos\theta)A_{ii} + (1 - v\cos\theta)B_{ii}] \equiv \beta\,, \tag{6.72}$$

(ii) both gluons in state (a)

$$\text{Amplitude} = \beta t(\cos\theta)\,, \tag{6.73}$$

where

$$t(\cos\theta) = 1 - \frac{2\sin^2\theta\, v^2}{1 - v^2\cos^2\theta}\,, \tag{6.74}$$

(iii) one gluon in state (a) and one in state (b)

$$\text{Amplitude} = 0\,. \tag{6.75}$$

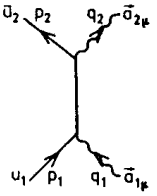
We can calculate β given the isospin states of the scalons. The results are tabulated below, where I is the total isospin of the scalons in the initial state

I	0	1	2
β	4	$2v\cos\theta$	2

(6.76)

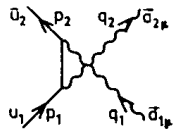
Finally we outline the calculation of the Compton scattering of gluons from isospin $\frac{1}{2}$ spinors using the Lagrangian (6.1) and the rules which we derived from it. The diagrams and the amplitudes are

(i)



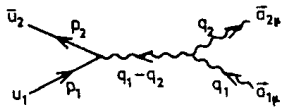
$$g^2\bar{u}_2(a_{2\mu}\cdot\tau\gamma_\mu)\frac{1}{\not p_1+\not q_1-m}(\gamma_\nu a_{1\nu}\cdot\tau)u_1\ , \tag{6.77}$$

(ii)



$$g^2\bar{u}_2(a_{1\nu}\cdot\tau\gamma_\nu)\frac{1}{\not p_1-\not q_2-m}(\gamma_\mu a_{2\mu}\cdot\tau)u_1 \tag{6.78}$$

(iii)

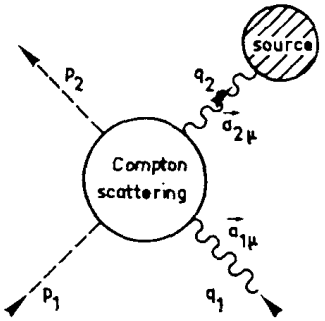


$$\begin{aligned}&\frac{g^2\bar{u}_2\gamma_\mu\tau\cdot u_1}{(q_1-q_2)^2}[(a_{2\nu}\times a_{1\mu})(2q_1-q_2)_\nu\\ &\quad+(a_{2\mu}\times a_{1\nu})(2q_2-q_1)_\nu\\ &\quad-(a_{2\nu}\times a_{1\nu})(q_1+q_2)_\mu]\ .\end{aligned} \tag{6.79}$$

We do not have the fourth diagram in this case. As in the scalar case, we can gauge one vector boson, $a_{1\mu}\to\alpha q_{1\mu}$, where upon we get, adding the results (6.77)–(6.79)

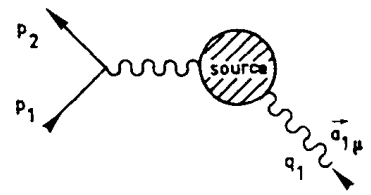
$$\frac{g^2\bar{u}_2\tau\cdot\gamma_\mu u_1}{(q_1-q_2)^2}[j_{2\mu}(a_2)\times\alpha] \tag{6.80}$$

As before (6.57), $j_{2\mu}(a_2)=q_2^2a_{2\mu}-q_{2\mu}(q_{2\nu}a_{2\nu})$ and again this is zero if $a_{2\mu}$ is free wave, or if it is a pure gradient (c.f. (6.54). However $j_{2\mu}(a_2)$ is not zero if $a_{2\mu}$ has a source ($j_{2\mu}(a_2)=\text{source}_{2\mu}$), but in this case, viz.,



(6.81)

We have again another type of diagram where the $a_{1\mu}$ acts on the source:

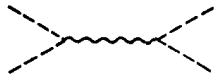


(6.82)

All the diagrams are taken into account, we again find that the physical process is gauge invariant. Hence there is no difficulty here as we expect, since these are tree diagrams.

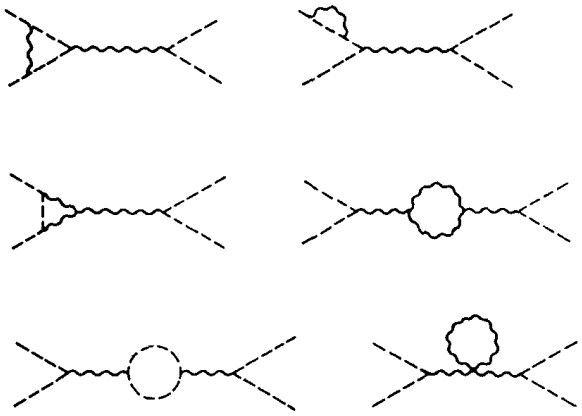
7. Quantisation continued – Loops

So far we have been successful in our attempt to quantise Yang–Mills theory. Now we want to try the procedure that we have adopted for higher order diagrams; this will entail diagrams with closed loops (the higher order tree diagrams do not present any difficulties). Why not therefore try to calculate the polarisation of the vacuum? But wait, if you want to discover a difficulty with a theory, you've got to look at a physical problem because some of your difficulties might come from not asking for a complete physical process (see the previous section). So the only way to discover whether something is right or wrong is not to pick up some arbitrary thing like the vacuum polarisation or the vacuum expectation value of a pair of operators, but to ask a physical question. But what should we ask? Correction to the Compton effect. This would be a fine problem, as good as any, but I decided to look at the following example – the scattering of two scalons by gluon exchange. In lowest order this is



(7.1)

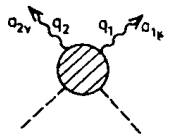
The correction diagrams to this coming from next order are of the type



(7.2)

We are not actually going to evaluate these diagrams, but only indicate what happens. When we calculate the diagrams (7.2) we, of course, have problems with renormalisation; but when we've straightened these out we would still like to check the result in some way. We can do this by checking whether or not we satisfy unitarity.

Consider the following diagram, where a pair of scalons annihilate into two gluons



(7.3)

The amplitude for this process is

$$\mathcal{M}_{\mu\nu} a_{1\mu} a_{2\nu} \,, \tag{7.4}$$

where $\mathcal{M}_{\mu\nu}$ is some tensor function of the momenta and isospin labels of the scalons. (We calculated this process to first order in the previous chapter.)

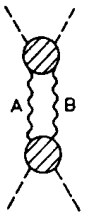
If we assume that the gluons are physical particles, there are constraints which their momenta and polarisation vectors must satisfy.

$$q_1^2 = q_2^2 = 0, \qquad q_1 \cdot a_1 = q_2 \cdot a_2 = 0\,, \tag{7.5}$$

i.e. there are only two polarisation states for each physical gluon. The probability for process (7.3) to take place is $\propto |\text{Amplitude}|^2$ i.e.

$$\text{Probability} \propto (\mathcal{M}_{\mu\nu} a_{1\mu} a_{2\nu})(\mathcal{M}^{\dagger}_{\mu'\nu'} a_{1\mu'} a_{2\nu'}) \tag{7.6}$$

where we have taken the a's to be real. In calculating this we see that it is closely related to the amplitude for the following process



(7.7)

The amplitude for this process is

$$\sum_{\substack{\text{all } \mu \\ \text{all } \nu}} \mathcal{M}_{\mu\nu} \mathcal{M}^{\dagger}_{\mu\nu} \times (\text{propagator terms for gluons A and B})\,. \tag{7.8}$$

If we take the imaginary part of (7.8) we get

$$\sum_{\substack{\text{all } \mu \\ \text{all } \nu}} \mathcal{M}_{\mu\nu} \mathcal{M}^{\dagger}_{\mu\nu}\,, \tag{7.9}$$

where the propagators have been removed by this process and the corresponding gluons are now on mass shell, and irrelevant constants have been ignored. Unitarity tells us that this must be the same as (7.6) summed over physical polarisation states:

$$\sum_{\substack{\text{physical} \\ \text{polarisation} \\ \text{states}}} (\mathcal{M}_{\mu\nu} a_{1\mu} a_{2\nu})(\mathcal{M}_{\mu'\nu'} a_{1\mu'} a_{2\nu'})\,. \tag{7.10}$$

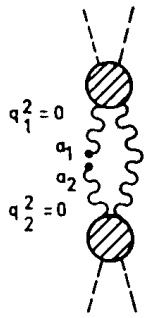
However note that in evaluating (7.9) we have summed over more polarisation states than we would have done had the intermediate gluons been physically real particles, so it is not obvious that (7.9) and (7.10) are equal.

In making the comparison we will only consider the μ index. We lose no

generality by doing this since if we can sort this out then we can also sort out the ν index. In making this assumption, we really compare

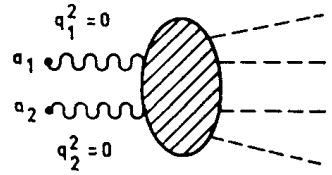
$$\sum_{\substack{\text{2 poln. states} \\ \text{of } a_{1\mu}\text{; all} \\ \nu \text{ states}}} \mathcal{M}_{\mu\nu} a_{1\mu} \mathcal{M}^{\dagger}_{\mu'\nu} a_{1\mu'} , \tag{7.11}$$

with (7.9). We can think of (7.11) as follows



(7.12)

The gluons a_1 and a_2 are on mass shell since they are physical gluons. This diagram, when the gluon line is not connected, is similar to, but more complicated than the Compton scattering process; it is



(7.13)

When we calculated the Compton effect we noticed that when $a_{1\mu} = \alpha q_{1\mu}$, then the amplitude was of the form

$$g(\alpha \times c_\lambda) \cdot j_\lambda(a_2) , \tag{7.14}$$

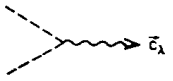
where

$$j_\lambda(a_2) = q_2^2 a_{2\lambda} - (q_{2\nu} a_{2\nu}) q_{2\lambda} . \tag{7.15}$$

In the case of the Compton scattering proper

$$c_\lambda = \frac{(S_1 \times S_2)(p_1 + p_2)_\lambda}{(p_1 - p_2)^2} \tag{7.16}$$

but in general we may think of c_λ as being the gluon field produced by some source



(7.17)

Since we are really only interested in the indices, define

$$\tilde{\mathcal{M}}_{\mu\mu'} = \sum_{\text{all } \nu} \mathcal{M}_{\mu\nu} \mathcal{M}^{\dagger}_{\mu'\nu} , \tag{7.18}$$

so that (7.11) becomes

$$\sum_{\substack{\text{2 poln. states} \\ \text{of } a_{1\mu}}} \tilde{\mathcal{M}}_{\mu\mu'} a_{1\mu} a_{1\mu'} , \tag{7.19}$$

and (7.9) becomes

$$\sum_{\text{all } \mu} \tilde{\mathcal{M}}_{\mu\mu} \tag{7.20}$$

Since $q_1^2 = 0$ we can choose the axes so that

$$q_1 = \omega(\overset{t}{1}, \overset{z}{1}, \overset{x}{0}, \overset{y}{0}) . \tag{7.21}$$

a_1 then has components only in the x and y direction and (7.19) becomes

$$\tilde{\mathcal{M}}_{xx} - \tilde{\mathcal{M}}_{yy} . \tag{7.22}$$

Similarly (7.20) becomes

$$\tilde{\mathcal{M}}_{tt} - \tilde{\mathcal{M}}_{xx} - \tilde{\mathcal{M}}_{yy} - \tilde{\mathcal{M}}_{zz} . \tag{7.23}$$

Eqs. (7.22) and (7.23) differ by

$$\tilde{\mathcal{M}}_{zz} - \tilde{\mathcal{M}}_{tt} . \tag{7.24}$$

So when we took the imaginary part of (7.8) we obtained the extra piece (7.24), which we do not want; the theory will be unitary if we can show that this extra piece is zero. Rotating the axes we can re-write (7.24) as

$$\tfrac{1}{2}\{\tilde{\mathcal{M}}_{(z-t),(z+t)} + \tilde{\mathcal{M}}_{(z+t),(z-t)}\} . \tag{7.25}$$

Consider the first term in (7.25): the second term is similar. We would get such a term if we considered a process with two gluons which have polarisa-

tions in the directions

$$\frac{1}{\sqrt{2}}\overset{t\ z\ x\ y}{(1,\ 1,\ 0,\ 0)} \quad \text{viz.} \quad a_{1\mu} = q_{1\mu} \ , \tag{7.26}$$

and

$$\frac{1}{\sqrt{2}}(1, -1, 0, 0) \quad \text{viz.} \quad a_{2\mu} \equiv \frac{N_\mu}{q_2 \cdot N} \ , \tag{7.27}$$

We get zero for $\tilde{\mathcal{M}}_{(z-t),(z+t)}$ only if $j_\lambda(a_2)$ is zero (from (7.14)). However it is not zero although q_2^2 is, since $q_2 \cdot a_2$ is not zero (the particle is not free).

In QED, this problem does not occur, because when we have a process with two external photons, if we gauge one photon (i.e. put the polarisation vector parallel to the momentum) the amplitude vanishes irrespective of whether or not the other photon is physical. That is, the quantity analogous to $j_\lambda(a_2)$ is identically zero independently of the state of the photon.

There is nothing we can do about this problem in Yang–Mills theory. If we wish to make the theory unitary we must subtract something to get rid of these extra pieces. By taking the simple Compton case, we can get some idea of the form of the thing we have to subtract. The Compton amplitude (7.14) becomes

$$g(\boldsymbol{\alpha} \times c_\lambda) \cdot \boldsymbol{\alpha} q_{2\lambda} \ , \tag{7.28}$$

when we use (7.27) and the fact that $q_2^2 = 0$. The polarisation $\boldsymbol{\alpha}$ is summed over when we make a closed loop. This is the extra piece that we have to get rid of. We can see that the most direct way to do this is to add an isovector scalar particle $\boldsymbol{\alpha}$ which is self coupled and also coupled to the vector according to (7.28). (This is the ghost particle.) In co-ordinate space this coupling becomes

$$g(\boldsymbol{P} \times A_\lambda) \cdot \partial_\lambda \boldsymbol{P} \ . \tag{7.29}$$

Clearly since the particle has to cancel a piece coming from the A_μ, its propagator must have the same form as that of the A_μ viz. $1/k^2$. We can see from the form of its coupling that this ghost appears in closed loops in diagrams in the same way as A's. It is not allowed to appear in the initial or final states since it is not a physical particle. It was not obvious to me from the type of analysis presented above, what to do in a diagram with more than one closed loop. (Unitarity is not a sufficient constraint in this case.) In fact the solution to the problem for larger numbers of loops is to add the ghost as if it were a Fermi particle. (At the one loop level, adding a Fermi particle is equivalent to subtracting the contribution from a particle coupled in the same way but

having Bose statistics, which is what we did above.) The contribution from this ghost could come from a Lagrangian of the form

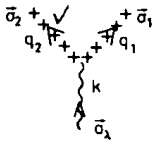
$$\mathcal{L}_g = \tfrac{1}{2}\partial_\mu P^+ \cdot \partial_\mu P + g\partial_\mu P^+ \cdot (A_\mu \times P) \tag{7.30}$$

This Lagrangian leads to the following diagram rules.

ghost propagator:

$$\delta_{ab}/k^2 \tag{7.31}$$

ghost vertex:



$$ig q_{2\lambda} \boldsymbol{\sigma}_2 \cdot (a_\lambda \times \boldsymbol{\sigma}_1)\,, \tag{7.32}$$

where the ghost only enters in closed loops. Topologically for every diagram with an A_μ closed loop there is one with a ghost loop in the same place. Note that (7.32) is not symmetric looking and this asymmetry is fundamental, so we add a check mark ($\surd$) to one of the ghost legs. We must keep track of these $\surd$'s as we go round a closed loop so as to ensure that they are always on the same side of a vertex. (It does not matter on which leg we put the $\surd$ provided that we always put it on the same one, since it can be shown that both choices lead to the same results.) In order to get the right factors it is necessary to assume that the ghost is a complex field.

All these details were proved by Fadde'ev and Popov [26]. We haven't proved them; we have merely indicated how if we proceed in a straight forward and naive way, we come into difficulties; and we got a pretty good smell of what to do. But for the full glory of the theory we must follow the approach of Fadde'ev – it requires more machinery and we will discuss it later.

Finally we write the total Lagrangian as

$$\begin{aligned}\mathcal{L} = &-\tfrac{1}{4}E_{\mu\nu}\cdot E_{\mu\nu} - \tfrac{1}{2}(\partial_\mu A_\mu)^2 + \tfrac{1}{2}\partial_\mu P^+ \cdot \partial_\mu P \\ &+ \tfrac{1}{2}g\partial_\mu P^+ \cdot (A_\mu \times P) + \text{Matter}\,(\phi, \psi, \text{A})\end{aligned} \tag{7.33}$$

Why is the second term in (7.33) present? If we expand the quadratic parts of $-\frac{1}{4}E_{\mu\nu}\cdot E_{\mu\nu}$ which will give us the propagator we get:

$$-\tfrac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)\cdot(\partial_\mu A_\nu - \partial_\nu A_\mu) = -\tfrac{1}{2}(\partial_\mu A_\nu)^2 + \tfrac{1}{2}(\partial_\mu A_\mu)^2\,. \tag{7.34}$$

We have changed the form of the Lagrangian in this by integrating the action by parts, throwing away (as usual) the surface terms at infinity. These changes

can produce no physics as they do not change the action. We now see that the $-\frac{1}{2}(\partial_\mu A_\mu)^2$ term in (7.33) cancels the same term in (7.34) to leave the quadratic terms in the Lagrangian as

$$-\tfrac{1}{2}(\partial_\mu A_\nu)^2 . \tag{7.35}$$

This is exactly what we want to give the propagator $\delta_{ab}\delta_{\mu\nu}/k^2$ which we have been using. (This is equivalent to choosing the gauge $\partial_\mu A_\mu = 0$.) That the Lagrangian (7.33) is equivalent to the Lagrangian

$$\mathcal{L} = -\tfrac{1}{4} E_{\mu\nu} \cdot E_{\mu\nu} + \text{Matter}\,(\phi, \psi, A) \tag{7.36}$$

(i.e. that the ghost Lagrangian (7.30) compensates for the extra term) is a more difficult problem which was solved by Fadde'ev and Popov [26], who also showed what form the extra terms in the Lagrangian would take in different gauges.

8. Quantisation via the Hamiltonian formalism. A ghostless gauge

We will now quantise Yang–Mills theory by going via the Hamiltonian formalism, rather than by path integrals; a similar procedure has been carried out in QED and one disadvantage is that we lose manifest Lorentz invariance, although we know that the final results must be relativistically invariant. We must choose a gauge before constructing the Hamiltonian since, as we know from our experience with QED, otherwise there are too many variables. The Hamiltonian is a machine for telling us how operators change with time; thus the important operator is

$$\partial/\partial t . \tag{8.1}$$

But we have seen that in Yang–Mills theory the physically interesting operator is D_μ which has a time component

$$\left(\frac{\partial}{\partial t} + A_0 \times\right). \tag{8.2}$$

Hence to recover the operator (8.1) which is appropriate for the Hamiltonian formalism, we want to put

$$A_0 = 0 . \tag{8.3}$$

Since we have the freedom of gauge transformations we can make this choice. If we do not do this the Hamiltonian formalism becomes extremely cumbersome and very difficult. By the way, why don't we make this choice in QED?

If we do, then a static charge has a vector potential associated with it which rises linearly with time. This is an inconvenience. However, in QCD, perhaps there are no static charges, i.e. we imagine that hadrons cannot separate into quarks or, when doing a problem, we start initially with no matter present and then create the matter and anti-matter (so that by doing this we ensure that the total colour charge of the world stays zero) which exists over a finite time. There seems to be a danger in this non-linear theory, since if the vector potential rises linearly with time, certain cross terms may produce physical effects (e.g. by quantum fluctuations); of course that's a dream, since if it were true, we could say that to avoid these difficulties, the vector potential is not allowed to rise linearly with time for ever, and hence we have proved colour confinement.

We shall write the classical equations for the Yang–Mills field, construct the Hamiltonian, having defined the conjugate momenta, and then quantise the theory by using the canonical commutation relations. The classical equation of motion is (2.28)

$$\partial_\mu E_{\mu\nu} = -J_\nu^{\text{matter}} - A_\mu \times E_{\mu\nu} ,$$

where

$$E_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu \times A_\nu , \tag{8.4}$$

The $E_{\mu\nu}$ tensor has six (Lorentz) components. Since we have singled out the time axis by our gauge choice, consider the ti component of $E_{\mu\nu}$. ($i = x, y, z$). From (8.4)

$$E_{ti} = \frac{\partial A_i}{\partial t} \equiv \mathcal{E}_i , \tag{8.5}$$

where by definition $\mathcal{E}_i$ is the colour electric field (named using the analogy with QED). Similarly the ij components are (from (8.4))

$$E_{ij} = \partial_i A_j - \partial_j A_i + A_i \times A_j \equiv \epsilon_{ijk} B_k , \tag{8.6}$$

where by definition B_k is the colour magnetic field. Using this definition, (8.6) gives

$$\underset{\sim}{B} = \operatorname{curl} \underset{\sim}{A} + \tfrac{1}{2} \underset{\sim}{A} \overset{\times}{\times} \underset{\sim}{A} . \tag{8.7}$$

Look at the i component of (8.4). This now reads

$$\frac{\partial^2 A_i}{\partial t^2} + (\operatorname{curl} \underset{\sim}{B})_i = -J_i - (\underset{\sim}{A} \overset{\times}{\times} \underset{\sim}{B})_i , \tag{8.8}$$

where we have introduced the notation that the upper vector symbols refer to

isospace and the lower symbols to real space.

These three equations can be written generally as

$$\frac{\partial^2 \underset{\sim}{A}}{\partial t^2} + \operatorname{curl} \underset{\sim}{B} = -\underset{\sim}{J} - \underset{\sim}{A} \overset{\times}{\times} \underset{\sim}{B} . \tag{8.9}$$

The fourth equation comes from taking the time component of (8.3); this is

$$\frac{\partial}{\partial t}(\underset{\sim}{\nabla} \cdot \underset{\sim}{A}) = -\rho - \underset{\sim}{A} \overset{\times}{\cdot} \underset{\sim}{\mathcal{E}} , \tag{8.10}$$

where we have put

$$\rho \equiv J_0 . \tag{8.11}$$

If we solve (8.9) we do not have to solve (8.10), because (8.10) is practically a consequence of (8.9); we can see this by taking the divergence of (8.9) and re-arranging the terms:

$$\frac{\partial}{\partial t}\left[\frac{\partial}{\partial t}\underset{\sim}{\nabla} \cdot \underset{\sim}{A}\right] = \frac{\partial}{\partial t}\left[-\underset{\sim}{A} \overset{\times}{\cdot} \underset{\sim}{\mathcal{E}}\right] \; -\underset{\sim}{\nabla} \cdot \underset{\sim}{j} - \underset{\sim}{A} \overset{\times}{\cdot} \underset{\sim}{j} . \tag{8.12}$$

This is the same as the time derivative of 8.10 provided

$$\frac{\partial \rho}{\partial t} = \underset{\sim}{\nabla} \cdot \underset{\sim}{j} + \underset{\sim}{A} \overset{\times}{\cdot} \underset{\sim}{j} . \tag{8.13}$$

The matter will be such that we have current conservation (in the covariant sense) and hence this equation will be satisfied. In principle, integrating (8.12) with respect to time to get (8.10) (using (8.13)), we could generate a function of x, y, and z. We assume that the initial conditions are such that this function is zero; it will then stay zero for all time. We can therefore forget about (8.10) in solving the theory.

We can solve (8.9) for $A_i(\underset{\sim}{x}, t)$: (8.9) is just a complicated non-linear differential equation for A_i. The dynamical features of the theory are simple since the only derivative of A_i which appears is the second time one, and this appears linearly. Can we find a Lagrangian from which (8.9) comes? It's easy, because we know all about Newton's laws and to get a term like $\partial^2 A_i/\partial t^2$ in the equation of motion, we know that we must have a term like $(\partial A_i/\partial t)^2$ in the Lagrangian. This is just like having the square of the velocity in the Lagrangian; the equation of motion then has a term linear in the acceleration. In order to generate the rest of the terms in the equation of motion we proceed as follows: the term J_i is generated by a Lagrangian term of the form $J_i \cdot A_i$, and it turns out that the remaining field pieces i.e. $\underset{\sim}{\nabla} \times \underset{\sim}{B} + \underset{\sim}{A} \overset{\times}{\times} \underset{\sim}{B}$ can be derived from a term of the form $\underset{\sim}{B} \cdot \underset{\sim}{B}$. So the Lagrangian has the form

$$-\frac{1}{2}\frac{\partial A_i}{\partial t}\cdot\frac{\partial A_i}{\partial t}-\tfrac{1}{2}B_i\cdot B_i+J_i\cdot A_i\,. \tag{8.14}$$

This is clearly the same as the Lagrangian

$$-\tfrac{1}{4}E_{\mu\nu}\cdot E_{\mu\nu}+\text{Matter}\,, \tag{8.15}$$

when we put $A_0 = 0$ and use the definition of B_i given earlier. What about the Hamiltonian? Nothing could be simpler; define the momentum conjugate to A_i by

$$\Pi_i=\frac{\delta\mathcal{L}}{\delta\dot{A}_i}=-\frac{\partial A_i}{\partial t}. \tag{8.16}$$

To get the Hamiltonian take the square of the momentum and add the potential energy, viz.

$$H=\tfrac{1}{2}\Pi_i\cdot\Pi_i+\tfrac{1}{2}B_i\cdot B_i-J_i\cdot A_i\,. \tag{8.17}$$

We now quantise the theory by imposing the canonical quantisation conditions, and treating Π and A as operators

$$[\Pi_i^a(\underset{\sim}{x},t),A_j^b(\underset{\sim}{y},t)]=\mathrm{i}\delta^{ab}\delta^{ij}\delta^3(\underset{\sim}{x}-\underset{\sim}{y})\,. \tag{8.18}$$

What would the diagrams look like for such a theory? We can get the propagator from (8.9) by looking at the linear part. This is substituting for $\underset{\sim}{B}$ using (8.7)

$$\frac{\partial^2\underset{\sim}{A}}{\partial t^2}+\operatorname{curl}(\operatorname{curl}\underset{\sim}{A})=-\underset{\sim}{J}-\underset{\sim}{A}\overset{\times}{\times}\underset{\sim}{B}-\tfrac{1}{2}\operatorname{curl}(\underset{\sim}{A}\overset{\times}{\times}\underset{\sim}{A})\,. \tag{8.19}$$

All the terms on the left side are linear in $\underset{\sim}{A}$ and all the terms on the right are higher order. We can solve this to obtain the propagator which is

$$\frac{\delta_{ij}-(k_ik_j/\omega^2)}{\omega^2-k_ik_i}\,. \tag{8.20}$$

This proof is left as an exercise; it is simple to complete using the method we described in sect. 6.

What about the coupling terms? They are the same as before except that there are no time components. This set of rules then gives the correct answers for Yang–Mills problems.

No fooling around! No ghosts! No nothing! Isn't that wonderful? Yes it's wonderful, but it's also mysterious because all the rules depend on choosing a time axis and it's not clear that the answer to a physical problem will be relatavistically invariant. So what? We know that it must be relativistically invariant be-

cause the theory we started with was. True, but the theory diverges and we encounter the delicate problem of imposing renormalisation without destroying Lorentz invariance – a problem which took 20 years to solve in the case of QED. It shouldn't take us this long to solve the problem in Yang–Mills as we're much smarter now.

Despite these problems these are perfectly good rules and we can go away and calculate with them.

Suppose we try to compare closed loops with this propagator and with the propagator we had earlier viz.

$$\delta_{\mu\nu}/k^2 \,. \tag{8.21}$$

In using (8.21) we are counting the contribution from A_0; but there are superfluous degrees of freedom; therefore we must subtract something to get rid of them – this something is the ghost. In some sense therefore, the ghost is the difference between the result one gets by using (8.20) and (8.21). What we subtract doesn't look relativistically invariant. The subtraction will be different depending on the orientation of the axes; in this case we have singled out the time axis and therefore if we rotate the axes things will change. The problem of showing that the results are independent of the method of subtraction was solved by Fadde'ev and Popov [26].

9. The equivalence of different gauges

The purpose of this chapter is to make clear the ideas involved in Fadde'ev and Popov's demonstration of the equivalence of different gauges. I shall not prove anything in great detail as this would take too long. We will start with the formalism of sect. 8, the $A_0 = 0$ gauge, and attempt to transform it so as to obtain the formalism in different gauges; in particular we will try to obtain the system in which the propagator is $\delta_{\mu\nu}/k^2$ ($\partial_\mu A_\mu = 0$ gauge). The simplest thing to do would be to write down the rules in the $A_0 = 0$ gauge and then transform them to the rules for the $\partial_\mu A_\mu = 0$ gauge. In order to get more generality and check that the formalism works in all orders we will use a more central approach via path integrals. What we shall do is to construct a path integral for the $A_0 = 0$ gauge, transform the path integral to the other gauge, and then get the rules for the other gauge from the path integral. We shall see that these rules are the same as those obtained in sects. 6 and 7.

As everybody knows, if we have a Lagrangian of the form

$$L = \frac{m\dot{q}^2}{2} - V(q) \tag{9.1}$$

Then the amplitude for the particle to go along a particular trajectory $q(t)$ is [18]

$$\exp(\mathrm{i}S/\hbar)\,, \tag{9.2}$$

where S is the classical action obtained from (9.1). The total amplitude for the particle to go from a to b is obtained by summing over all trajectories connecting a and b, viz.

$$\text{Amplitude} \propto \int \exp\left(\frac{\mathrm{i}}{\hbar}\int\left[\frac{m\dot{q}(t)^2}{2} - V(q(t))\right]\mathrm{d}t\right)\mathcal{D}q(t)\,. \tag{9.3}$$

If we have a system of several variables $q_i(t)$, (9.3) becomes

$$\text{Amplitude} \propto \int \exp\left(\frac{\mathrm{i}}{\hbar}\int \sum_i\left[\frac{m_i\dot{q}_i(t)^2}{2} - V(q_i(t))\right]\mathrm{d}t\right)\mathcal{D}q_i(t)\,. \tag{9.4}$$

In a field theory the variable i becomes continuous and is denoted by $\underset{\sim}{x}$. Eq. (9.4) then becomes

$$\text{Amplitude} \propto \int \exp\left(\frac{\mathrm{i}}{\hbar}\int \mathcal{L}(A)\mathrm{d}^3\underset{\sim}{x}\mathrm{d}t\right)\mathcal{D}A(\underset{\sim}{x}, t)\,, \tag{9.5}$$

where $\mathcal{L}(A)$ is the Lagrangian density for the theory and the integral over $\underset{\sim}{x}$ replaces the sum over i in the limit i continuous

$$\mathcal{L}(A) = \left(\frac{\partial A(\underset{\sim}{x}, t)}{\partial t}\right)^2 - V(A)\,. \tag{9.6}$$

Eq. (9.5) means this: if A is distributed in some way in spacetime, there is a certain action for this distribution which we can calculate. The classical theory corresponds to the distribution which makes the action a minimum. Quantum mechanics corresponds to adding the amplitude corresponding to the actions for all possible distributions. We can deduce the classical limit by noticing that in the integration over field distributions, the amplitudes corresponding to near stationary values of the action add coherently whereas those far from the minimum are changing so rapidly that the amplitudes from neighbouring distributions tend to cancel. In the limit of action $S/\hbar$ being very large only paths near the minimum contribute to the integral over distributions and so we obtain the classical theory.

We now return to the Lagrangian density which we derived for the $A_0 = 0$ gauge in the previous section (8.14) and substitute it in (9.5). This gives

$$\int \exp\left(\mathrm{i}\int\left[-\frac{1}{2}\frac{\partial \underset{\sim}{A}(\underset{\sim}{x}, t)}{\partial t} : \frac{\partial \underset{\sim}{A}(\underset{\sim}{x}, t)}{\partial t} - \tfrac{1}{2}\underset{\sim}{B} : \underset{\sim}{B} + \underset{\sim}{J} : \underset{\sim}{A}\right]\mathrm{d}^4x\mathcal{D}\underset{\sim}{A}(\underset{\sim}{x}, t)\mathcal{D}_{\text{Matter}}\,. \tag{9.7}$$

$\mathcal{D}\underset{\sim}{A}(\underset{\sim}{x}, t)$ is defined to mean $\Pi_{i,a,\underset{\sim}{x},t} \mathrm{d}A_i^a(\underset{\sim}{x}, t)$ and in $\mathcal{D}_{\text{Matter}}$ we have to integrate over the spin $\frac{1}{2}$ ψ field. This means that we must use the path integral using Grassman algebra, which we mentioned earlier.

Unfortunately from the path integral viewpoint, the expression is not manifestly Lorentz invariant. However we can write (9.7) as follows

$$\int \exp\left\{ i \int [-\tfrac{1}{4} E_{\mu\nu} \cdot E_{\mu\nu} + J_\mu \cdot A_\mu] \mathrm{d}^4 x \right\} \delta(A_0(\underset{\sim}{x}, t)) \mathcal{D}A_\mu(x, t) \mathcal{D}_{\text{Matter}} , \tag{9.8}$$

where $\delta(A_0(\underset{\sim}{x}, t))$ is defined to mean $\Pi_{a,\underset{\sim}{x},t} \delta(A_0^a(\underset{\sim}{x}, t))$ and in this expression the functional integral runs over all four space-time components of $A_\mu(\underset{\sim}{x}, t)$. Although (9.8) and (9.7) are equivalent (by construction) we note that in (9.8):

(i) $E_{\mu\nu} \cdot E_{\mu\nu}$ is invariant under all gauge transformations;

(ii) $J_\mu \cdot A_\mu$ is also invariant under gauge transformations provided that the matter fields transform in the correct way;

(iii) $\mathcal{D}A_\mu$ is also invariant under an infinitesimal gauge transformation

$$A_\mu \to A_\mu + \partial_\mu \boldsymbol{\alpha} + \boldsymbol{\alpha} \times A_\mu . \tag{9.9}$$

We can see that (iii) is true by the following argument. Since we are considering $\mathcal{D}A_\mu$, i.e. an infinitesimal, we can drop the $\partial_\mu \boldsymbol{\alpha}$ term since it is higher order in infinitesimal. $\mathcal{D}A_\mu$ is a volume element in isospace

$$\mathcal{D}A_\mu = \prod_{\mu,\underset{\sim}{x},t} \mathrm{d}A_\mu^1(\underset{\sim}{x}, t) \mathrm{d}A_\mu^2(\underset{\sim}{x}, t) \mathrm{d}A_\mu^3(\underset{\sim}{x}, t) . \tag{9.10}$$

The remaining gauge transformation in (9.9) ($\boldsymbol{\alpha} \times A_\mu$) is a rotation in isospace and volume elements are invariant under such transformations (cf. $\mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z$ which is invariant under a rotation in three dimensional Euclidean space). Hence $\mathcal{D}A_\mu$ is invariant.

We have therefore shown that (9.8) is gauge invariant apart from the δ-function. What happens to it under a gauge transformation? Under a gauge transformation (not necessarily infinitesimal),

$$A_0 \to G A_0 , \tag{9.11}$$

which in general contains the space as well as time components of A_μ. Since any A_μ distribution can be gauge transformed so that $A_0 = 0$, the δ-function is really more universal than it looks. Consequently we can get to any other gauge by suitably choosing G.

As a specific example, let us choose to transform to the gauge

$$\partial_\mu A_\mu = 0 \,. \tag{9.12}$$

We will find that when we transform to this gauge,

$$\delta(A_0) \to \delta(\partial_\mu A_\mu)\mathcal{F}(A) \tag{9.13}$$

and we will also find that the factor $\mathcal{F}(A)$ can be written in such a way that it exactly reproduces the ghost term (Lagrangian (7.30)) which we required in the $\partial_\mu A_\mu = 0$ gauge.

There are several mathematical tricks which we will need in what follows. Because of the vast number of indices flying around, we will often use a symbolic notation.

Mathematical trick 1. Since almost all the questions about path integrals which are important are concerned with changes in the fields, the overall normalisation of the path integrals is irrelevant. We can, therefore, ignore any constants (including infinite ones) which appear as multiplicative factors outside the path integral provided that they do not depend on the fields.

Mathematical trick 2. Suppose that we have a number of variables x_i and a number of functions y_i, which depend on the x_i. If we wish to transform from the x_i to the y_i then the δ-function transform as follows:

$$\delta(y_i - y_i(0)) = \left[\mathrm{Det}\,\frac{\partial y_i}{\mathrm{d}x_j}\right]^{-1}_{x_j=0} \delta(x_j) \,, \tag{9.14}$$

or

$$\delta(x_j) = \left[\mathrm{Det}\,\frac{\partial y_i}{\mathrm{d}x_j}\right]_{x_j=0} \delta(y_i - y_i(0)) \,,$$

where as before the δ-functions are the generalised functions, i.e. $\delta(y_i)$ means $\Pi_i \delta(y_i)$. This can be seen by extending the result for one variable $y = f(x)$ viz.

$$\delta(y) = \frac{1}{|f'(x)|}\delta(x) \,, \tag{9.15}$$

to several variables.

Mathematical trick 3. Consider the integral

$$\int \exp(-\tfrac{1}{2} cp^2)\,\mathrm{d}p = \sqrt{2\pi/c} \,. \tag{9.16}$$

We drop the $\sqrt{2\pi}$ factors in what follows since they will be irrelevant because of mathematical trick 1. Let us now have several variables and look at the integral

$$\int \exp\left(-\tfrac{1}{2}\sum_i p_i c_i' p_i\right) \prod_i \mathrm{d}p_i \,. \tag{9.17}$$

This clearly has the value $1/\sqrt{\Pi_i c_i'}$.

If we do a linear transformation on these p_i's:

$$p_i \to \sum_j a_{ij} p_j \,, \tag{9.18}$$

then (9.17) becomes

$$\int \exp\left(-\tfrac{1}{2}\sum_{ij} p_i c_{ij} p_j\right) \prod_i \mathrm{d}p_i = \frac{1}{\sqrt{\mathrm{Det}\, c_{ij}}} \tag{9.19}$$

where the c's and the c''s are related by (9.18). Eq. (9.19) is like a path integral in quantum mechanics for a Bose particle interacting in a way described by the matrix c_{ij} (provided that we make a Wick rotation of the time axis $\mathrm{d}t \to \mathrm{i}\mathrm{d}t$). If we suppose that p_i is complex then we are really integrating over twice as many variables and the square root disappears:

$$\int \exp\left(-\tfrac{1}{2}\sum_{ij} p_i^* c_{ij} p_j\right) \prod_i \mathrm{d}p_i = \frac{1}{\mathrm{Det}\, c_{ij}} \,. \tag{9.20}$$

If we try to do the same integral for particles Q_i with Fermi statistics (i.e. we introduce a Grassman algebra), we find that, with a suitable definition of what we mean by an integral in this case,

$$\int \exp\left\{-\tfrac{1}{2}\sum_{ij} Q_i^* c_{ij} Q_j\right\} \prod_i \mathrm{d}Q_i = \mathrm{Det}\, c_{ij} \,. \tag{9.21}$$

It will turn out that (9.21) is just the reason that the ghost has Fermi statistics.

When we transform from the $\delta(A_0)$ to the $\delta(\partial_\mu A_\mu)$ in the functional integral, we will pick up a Jacobian similar to that in (9.14). Symbolically we can think of the variable x_i as being $A_0(\underset{\sim}{x}, t)$ and y_i as being $\partial_\mu A_\mu(\underset{\sim}{x}, t)$ (in this case $y_i(0)$ is zero). We will then use (9.21) to turn the determinant into a path integral. Another equation is useful when it comes to finding the coefficients c_{ij}; consider

$$y_i(x_j + p_j) = y_i(x_j) + c_{ij} p_j \,, \tag{9.22}$$

where we have defined $c_{ij} = (\mathrm{d}y_i/\mathrm{d}x_j)_{x_j=0}$. On multiplying by p_i^* we get

$$p_i^* [y_i(x_j + p_j) - y_i(x_j)] = p_i^* c_{ij} p_j \,. \tag{9.23}$$

The term on the right has the same form as the expression in (9.20). We are trying to transform from the gauge $A_0 = 0$ to the gauge $\partial_\mu A_\mu = 0$ and there-

fore we need the matrix

$$\Delta(\partial_\mu A_\mu)/\Delta(A_0)\,, \tag{9.24}$$

which is analogous to the matrix $\mathrm{d}y_i/\mathrm{d}x_j$. This matrix $\Delta(\partial_\mu A_\mu(y))/\Delta(A_0(x))$ has infinite dimension corresponding to the continuous variables x and y, and in addition has isospin labels. If we try to evaluate (9.24) directly, we will run into difficulties since the two gauges are linked by a finite gauge transformation which depends on the original fields. (We know that finite gauge transformations are difficult to handle.) The derivative $\mathrm{d}y_i/\mathrm{d}x_j$ can fortunately be obtained in an indirect way. Suppose y_i and x_j are changed by changing a variable t. We may be able to evaluate $\mathrm{d}y_i/\mathrm{d}x_j$ by first evaluating $\mathrm{d}y_i/\mathrm{d}t$ and $\mathrm{d}x_j/\mathrm{d}t$ and then using

$$\frac{\mathrm{d}y_i}{\mathrm{d}x_j} = \frac{\mathrm{d}y_i}{\mathrm{d}t} \cdot \frac{\mathrm{d}t}{\mathrm{d}x_j}. \tag{9.25}$$

It was the very clever idea of Fadde'ev to extend this idea to evaluate (9.24): in this case he took t to be an infinitesimal gauge transformation. If we do an infinitesimal gauge transformation on A_0 (gauge parameter $\boldsymbol{\alpha}$) we get:

$$A_0' = A_0 + \boldsymbol{\alpha} \times A_0 + \partial_0 \boldsymbol{\alpha}\,. \tag{9.26}$$

Transform $\partial_\mu A_\mu$ by an infinitesimal gauge transformation with parameter $\boldsymbol{\beta}$:

$$\partial_\mu A_\mu' = \partial_\mu A_\mu + \boldsymbol{\beta} \times \partial_\mu A_\mu + \partial_\mu \beta \times A_\mu + \partial_\mu \partial_\mu \boldsymbol{\beta}\,. \tag{9.27}$$

As we are trying to use an equation of the form (9.25) to evaluate the matrix (9.24), we would naively think that we should take $\boldsymbol{\alpha} = \boldsymbol{\beta}$. But wait a minute. What we really need is for the gauge parameter $\boldsymbol{\beta}$ to be $\boldsymbol{\alpha}$ gauge transformed by the finite gauge transformation G, defined by (9.11). (All the fields are rotated by the gauge transformation G and so what we really want is to know how the fields A_0 and $\partial_\mu A_\mu$ change in terms of the variables $\boldsymbol{\alpha}$ and $G\boldsymbol{\alpha}$ respectively.) However we will use $\boldsymbol{\alpha}$ for $\boldsymbol{\alpha}$ gauge transformed for ease of notation in what follows. (This sublety is not vital to the argument and must not be allowed to confuse us.) Since $A_0 = 0$ in the original gauge, the change in the original A_0 produced by $\boldsymbol{\alpha}$ (from (9.26)) is

$$\partial_0 \boldsymbol{\alpha}. \tag{9.28}$$

The matrix (9.23) is to be considered as the product of the two factors

$$\frac{\Delta(\partial_\mu A_\mu)}{\Delta(\boldsymbol{\alpha})} \quad \text{and} \quad \frac{\Delta(\boldsymbol{\alpha})}{\Delta(A_0)}. \tag{9.29}$$

So we have really factored the matrix which described how one gauge varies

with respect to the other gauge into two parts, each of which tells us how one of the gauges varies with respect to the infinitesimal gauge transformation $\boldsymbol{\alpha}$. The second of these two factors is independent of the fields since $\Delta(A_0)$ is independent of the fields (9.28). Using maths trick 1 we can drop this factor. The change in $\partial_\mu A_\mu$, is from (9.27):

$$\partial_\mu \boldsymbol{\alpha} \times A_\mu + \partial_\mu \partial_\mu \boldsymbol{\alpha} , \tag{9.30}$$

where we have used the fact that $\partial_\mu A_\mu = 0$ and we have put $\boldsymbol{\beta} = \boldsymbol{\alpha}$. This depends on the field A_μ and so we cannot get rid of the $\Delta(\partial_\mu A_\mu)/\Delta(\boldsymbol{\alpha})$ term in the same way as we got rid of $\Delta(A_0)/\Delta(\boldsymbol{\alpha})$.

We can now write (9.8) as

$$\int \exp\left\{ i \int [-\tfrac{1}{4} E_{\mu\nu} \cdot E_{\mu\nu} + J_\mu \cdot A_\mu] \, d^4x \right\} \delta(\partial_\mu A_\mu \mathcal{F}(A_\mu) \mathcal{D} A_\mu \mathcal{D}_{\text{Matter}} , \tag{9.31}$$

where $\mathcal{F}(A_\mu)$ is the Jacobian for the transformation from $\delta(A_0)$ to $\delta(\partial_\mu A_\mu)$ which we have identified as

$$\text{Det} \frac{\Delta(\partial_\mu A_\mu)}{\Delta(\boldsymbol{\alpha})} . \tag{9.32}$$

This can be written in the form

$$\int \exp\left\{ \frac{i}{2} \int P^+ (\partial_\mu \partial_\mu - A_\mu \times \partial_\mu) P \, d^4x \right\} \mathcal{D} P , \tag{9.33}$$

where $\boldsymbol{P}$ is a Fermi particle. We can show this as follows. Let (9.32) be symbolically Det c_{ij}; this can be written, using maths trick 3 (9.21) as

$$\int \exp(\tfrac{1}{2} P_i^+ c_{ij} P_j) \prod_i dP_i . \tag{9.34}$$

Use (9.23) to re-write the exponent of this, giving

$$\int \exp\{-\tfrac{1}{2} P_i^* [y_i(x_j + P_j) - y_i(x_j)]\} \prod_i dP_i . \tag{9.35}$$

We now convert back from the symbolism to the A's, putting as before $y_i \to \partial_\mu A_\mu$ and $x_i \mapsto A_0$:

$$\int \exp\{-\tfrac{1}{2} P^* \Delta(\partial_\mu A_\mu)\} \mathcal{D} P . \tag{9.36}$$

Substituting for $\Delta(\partial_\mu A_\mu)$ from (9.30) and carrying out a Wick rotation, we get eq. (9.33). Integrating by parts in (9.33) and noting that $\partial_\mu A_\mu = 0$ by the choice of gauge, we get

$$\int \exp\left\{\frac{\mathrm{i}}{2}\int [\partial_\mu P^{+\cdot} \partial_\mu P + \partial_\mu P^{+\cdot}(A_\mu \times P)]\,\mathrm{d}^4x\right\}\mathcal{D}P\,, \tag{9.37}$$

and thus we have obtained the same propagator and coupling as we did in sect. 7.

In general, if we wish to transform to the gauge $f(A_\mu) = 0$, we first calculate

$$f(A_\mu + D_\mu \boldsymbol{\alpha}) = f(A_\mu) + f'(A_\mu) D_\mu \boldsymbol{\alpha}\,, \tag{9.38}$$

cf. (9.27) for the case $f(A_\mu) = \partial_\mu A_\mu$. Going through the above procedure

$$\delta(A_0) \to \delta[f(A_\mu)] \exp\left\{\frac{\mathrm{i}}{\xi}\int P^+ [f'(A_\mu) D_\mu] P\,\mathrm{d}^4x\right\}\mathcal{D}P\,. \tag{9.39}$$

To summarise, we have transformed the path integral (9.7) to

$$\int \exp\left\{\mathrm{i}\int [-\tfrac{1}{4}E_{\mu\nu}{}^{\cdot} E_{\mu\nu} + J_\mu{}^{\cdot} A_\mu + \tfrac{1}{2}\partial_\mu P^{+\cdot} \partial_\mu P + \tfrac{1}{2}\partial_\mu P^{+\cdot}(A_\mu \times P)]\,\mathrm{d}^4x\right\}$$
$$\times \delta(\partial_\mu A_\mu)\mathcal{D}A_\mu \mathcal{D}P\, \mathcal{D}_{\text{Matter}}\,. \tag{9.40}$$

People have tried to make rules from this, but there is an easier way. From (9.14), we note that the determinant is independent of the value of $y_i(0)$ in the δ-function. Hence if instead of choosing the gauge $\partial_\mu A_\mu = 0$ (i.e. $y_i(0) = 0$) we chose the gauge

$$\partial_\mu A_\mu(\underset{\sim}{x}, t) = \boldsymbol{\beta}(\underset{\sim}{x}, t) \quad (\text{i.e., } y_i(0) = \beta)\,, \tag{9.41}$$

the only changes this will produce will be to replace $\delta(\partial_\mu A_\mu)$ in (9.40) by $\delta(\partial_\mu A_\mu - \boldsymbol{\beta})$, and add an additional piece to 9.30. The last piece produces an additional term

$$\tfrac{1}{2}P^{+\cdot}(P \times \boldsymbol{\beta}) \tag{9.42}$$

which will cancel the extra term appearing in 9.37 when we integrate by parts since now $\partial_\mu A_\mu = \boldsymbol{\beta}$. We now that we can multiply 9.48 by the constant

$$\exp\left(-\frac{\mathrm{i}}{2}\int \boldsymbol{\beta}^2\right)\mathcal{D}\boldsymbol{\beta}\,, \tag{9.43}$$

as this does not produce any difference in the theory.

Integrating over $\boldsymbol{\beta}$ using $\delta(\partial_\mu A_\mu - \boldsymbol{\beta})$ we obtain the following path integral

$$\int \exp\left\{\mathrm{i}\int [-\tfrac{1}{4}E_{\mu\nu}{}^{\cdot} E_{\mu\nu} - \tfrac{1}{2}(\partial_\mu A_\mu)^{\cdot}(\partial_\nu A_\nu) + \tfrac{1}{2}\partial_\mu P^{+\cdot} \partial_\mu P\right.$$
$$\left. + \tfrac{1}{2}\partial_\mu P^{+\cdot}(A_\mu \times P)]\,\mathrm{d}^4x\right\}\mathcal{D}A_\mu \mathcal{D}P \mathcal{D}_{\text{Matter}}\,. \tag{9.44}$$

We have recovered the Lagrangian (7.33) which led to the rules in the $\partial_\mu A_\mu = 0$ gauge, and we have therefore shown that the gauges $A_0 = 0$ and $\partial_\mu A_\mu = 0$ are equivalent.

It is possible to choose other gauges by replacing the term $\frac{1}{2}(\partial_\mu A_\mu)\cdot(\partial_\nu A_\nu)$ by $(\lambda/2)(\partial_\mu A_\mu)\cdot(\partial_\nu A_\nu)$, where λ takes any value between 0 and 1. This of course will produce a different effective Lagrangian. (What people really do when they choose a gauge is to construct an effective Lagrangian and then calculate with it.) This produces a propagator for the A_μ of the form

$$\frac{\delta_{\mu\nu} - (1-\lambda)(k_\mu k_\nu / k^2)}{k^2} \tag{9.45}$$

which of course must produce the same physics. The modified propagator is useful when it comes to proving the renormalisability of gauge theories (about which I would have said more if I had more time) since if we take the $\lambda = 1$ gauge we can then use naive power counting (i.e. the smart thing to do). Naive power counting tells us whether something has a danger of diverging – it doesn't prove anything: theories which look divergent by naive power counting may not be, due to cancelations among the divergences. In some gauges naive power counting makes the theory look horribly divergent but we know that this cannot be the case. In the $\lambda = 1$ gauge, there are no powers of momenta in the numerator of the propagator and so naively the power counting is better.

Fadde'ev and Popov, in fact, proceeded in the following way, without choosing a specific gauge. Blindly if we were to write the simple simplest possible path integral for the theory, it would be

$$\int \exp\left\{ i \int [-\tfrac{1}{4} E_{\mu\nu}\cdot E_{\mu\nu} + J_\mu \cdot A_\mu] \mathrm{d}^4 x \right\} \mathcal{D}A_\mu \mathcal{D}_{\text{Matter}} \,. \tag{9.46}$$

However, if we make a gauge transformation on the exponent, it does not change; so when we integrate over the A_μ's which are connected by this gauge transformation, we will get infinity. An analogy is when we evaluate the integral

$$\int \exp\{-\alpha(x^2 + 2xy + y^2)\} \mathrm{d}x \, \mathrm{d}y \tag{9.47}$$

Changing variables to $x - y$ and $x + y$, this becomes, up to a constant,

$$\int \exp\{-\alpha(x+y)^2\} \mathrm{d}(x+y) \mathrm{d}(x-y) \,. \tag{9.48}$$

Since the exponent is independent of $(x - y)$, the integral over $(x - y)$ diverges. A similar thing happens in the path integral except that the situation is more complicated: the Jacobian in the change of variables depends on A_μ and this is what introduces the ghosts we have been discussing.

References

[1] C.N. Yang and R. Mills, Phys. Rev. 96 (1954) 191.
[2] J. Goldstone, Nuovo Cimento 19 (1961) 15.
[3] P.W. Higgs, Phys. Rev. Letters 12 (1964) 132.
[4] P.W. Higgs, Phys. Rev. Letters 13 (1964) 508.
[5] T.W. Kibble, Phys. Rev. 155 (1967) 1554.
[6] C.H. Llewellyn Smith, Phys. Letters B46 (1973) 233.
[7] S. Weinberg, Phys. Rev. Letters 19 (1967) 1264.
[8] A. Salam, in: Elementary Particle Theory, ed. N. Svartholm (Almquist and Forlag, Stockholm, 1968).
[9] J.D. Bjorken and S.D. Drell, Relativistic Quantum Mechanics (McGraw-Hill, 1968).
[10] E.S. Abers and B.W. Lee, Phys. Reports C9 (1973) 1.
[11] R.P. Feynman and A.R. Hibbs, Quantum Mechanics and Path Integrals (McGraw-Hill, 1965).
[12] L. Susskind, this volume, p. 207.
[13] R.P. Feynman, M. Kislinger and F. Ravndal, Phys. Rev. D3 (1971) 2706.
[14] P.W. Higgs, Phys. Rev. 155 (1966) 1554.
[15] B. Renner, Current Algebras and Their Applications (Pergamon, 1968).
[16] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman Lectures in Physics, Vol. 3 (1965).
[17] D. Rogovin and M. Scully, Phys. Reports 25 (1976) 175.
[18] R.P. Feynman, Rev. Mod. Phys. 20 (1948) 367.
[19] C. DeWitt-Morette, private communication (1951); General Relativity and Gravitation Journal, to be published.
[20] C. Garrod, Rev. Mod. Phys. 38 (1966) 483.
[21] P.A.M. Dirac, Principles of Quantum Mechanics (Oxford University Press, 1957).
[22] J. Schwinger, in: Theoretical Physics, Trieste Lectures, 1962 (IAEA, Vienna, 1963).
[23] R.P. Feynman, Acta Phys. Polonica 26 (1963) 697.
[24] B.S. DeWitt, Phys. Rev. 162 (1967) 1195, 1239.
[25] G. 't Hooft, Nucl. Phys. B35 (1971) 167.
[26] L.D. Faddeev and V.N. Popov, Phys. Letters B25 (1967) 29.
[27] V.N. Popov and L.D. Faddeev, Perturbation Theory for Gauge Invariant Fields, Kiev I.T.P. Report (unpublished).

PHYSICAL REVIEW D VOLUME 18, NUMBER 9 1 NOVEMBER 1978

Quantum-chromodynamic approach for the large-transverse-momentum production of particles and jets

R. P. Feynman, R. D. Field, and G. C. Fox
California Institute of Technology, Pasadena, California 91125
(Received 15 May 1978)

It is shown that if, in a calculation of high-transverse-momentum ($p_\perp$) meson production in hadron-hadron collisions, one includes not only the scale-breaking effects that might be expected from asymptotically free theories but also the effects due to the transverse momentum of quarks in hadrons *and* further adds contributions from quark-gluon and gluon-gluon scattering to those of quark-quark scattering *then* the results are not inconsistent with the data. The approach yields the correct magnitude and an apparent approximate $1/p_\perp^8$ behavior in accord with single-particle data for the energy range currently observed. Two-particle correlations are examined. Because of scale-breaking effects and the presence of gluons, the theory does not have the problem of predicting too many away-side hadrons at large $p_\perp$ as did an earlier quark-quark scattering "black-box" approach. We conclude that the quantum-chromodynamics approach is in reasonable accord with the data although theoretical uncertainties (especially at low $p_\perp$) make incontrovertible conclusions impossible at present. Crucial tests of the theory require higher $p_\perp$ than are now available; estimates for this region are made.

I. INTRODUCTION

We investigate whether the present experimental behavior of mesons with large transverse momentum in hadron-hadron collisions is consistent with the theory of quantum-chromodynamics (QCD) with asymptotic freedom, at least as the theory is now partially understood. It is shown that if things behave more or less according to current theoretical ideas, the experimental data at high $p_\perp$ would be explicable with reasonable choices for currently unknown quantities (such as the distribution of gluons in the proton and the fragmentation functions describing gluon jets). The theory of QCD might provide an adequate explanation of all the experimental results that we have discussed in previous papers (hereafter referred to as FF1[1] and FFF[2]).

We and others[3-9] investigated this large-$p_\perp$ experimental behavior phenomenologically. In particular, the view that large-$p_\perp$ mesons are generated by hard large-angle collisions between quarks present in the initial hadrons has been found to be very fruitful. The outgoing quarks are assumed not to come out freely, but to distribute their momentum over a series of hadrons that form a "jet" in the general direction of the "original quark."

This view when compared in detail with all available experiments is found to be successful in many regards. In particular, the distributions of particles in single-arm experiments and the large ratio of jet to particle cross sections are successfully interpreted or predicted. To do this, the differential cross section for quark-quark collisions, $d\hat{\sigma}/d\hat{t}$ ("black-box" cross sections), was taken to vary as $f(\hat{t}/\hat{s})\hat{s}^{-4}$ in disagreement with field theory which expects $\hat{s}^{-2}$ behavior. The size and angular dependence, 2300 mb/$(-\hat{s}\hat{t}^3)$, was chosen empirically. In addition, in FF1 and FFF, the effects of collisions of gluons (as constituents within hadrons) were omitted. At that time, there was no experimental evidence to require their existence.

The $f(\hat{t}/\hat{s})\hat{s}^{-4}$ behavior was chosen as a direct result of assumptions that parton distributions scaled with energy and the observation that experiments done at two or more different center-of-mass energies W but at a fixed $x_\perp = 2p_\perp/W$ and fixed angle showed the invariant cross section varying as $p_\perp^{-8}$.

It is necessary to include a transverse momentum, $(k_\perp)_{h\to q}$, of the quarks within the initial hadrons to account for much of the data[2,10-14]; for example, observations in two-arm experiments of P_{out} (Ref. 15) (i.e., lack of coplanarity of the beam, target, and two outgoing hadrons).[16] Some of the apparent large $p_\perp$ of a hadron can be due to the initial transverse momentum of the incoming quarks. This vitiates the direct scaling connection between a $d\hat{\sigma}/d\hat{t} = f(\hat{t}/\hat{s})\hat{s}^{-n}$ and the invariant cross section behaving like $p_\perp^{-2n}$ at fixed $x_\perp$ and $\theta_{\text{c.m.}}$. However, as long as $\langle k_\perp\rangle_{h\to q}$ is 500 MeV or less, the effects on the single-particle invariant cross section are not great and in FFF we limited ourselves to values not higher than this.

There are, however, two serious discrepancies with experiment which indicate that, in spite of the successes, something is wrong with the black-box model. First, recently measured values of P_{out} seem to be higher than expected so that $\langle k_\perp\rangle_{h\to q}$ must exceed 500 MeV. That would mean

that the scaling argument leading to $\hat{s}^{-4}$ for the quark-quark cross section was wrong in the range of data used. That is, the $p_\perp^{-8}$ behavior of experiment would be *accidental* and not fundamental. For example, a quark-quark cross section varying only as $\hat{s}^{-3}$ could (with a $\langle k_\perp \rangle_{h\to q}$ big enough to explain the large P_{out}'s observed) produce an apparent $p_\perp^{-8}$ hadron scaling.

Second, in two-arm experiments with events triggered on one side by a high-$p_\perp$ hadron (the "toward" side), the number of particles with large $p_\perp$ on the opposite side (the "away" side) was experimentally only about $\frac{1}{3}$ of the number predicted. Some of this is accounted for by increasing $\langle k_\perp \rangle_{h\to q}$, but not all, by far. The only explanation available in the approach is that the outgoing momentum on the away side is distributed more softly (distributed among more hadrons of lower momentum) than is typical of a quark; that is, there is more than one component contributing to the high-transverse-momentum jets. This is evidence for the need to include gluons as well as quarks in the description of high-$p_\perp$ phenomena. It requires both gluons and the assumption that gluons fragment into a distribution of hadrons of lower average momentum than does a quark.

These discrepancies lead us to include gluons in an analysis of high-$p_\perp$ hadron-hadron collisions, and to the further suggestion that QCD field theory might not be inconsistent with what is observed. Although with large values of $\langle k_\perp \rangle_{h\to q}$, the scattering cross section behaving like $f(\hat{t}/\hat{s})\hat{s}^{-3}$ will yield a $p_\perp^{-8}$ behavior over the range of data, this still differs from the naive field-theory expectation of $\hat{s}^{-2}$. (Including gluons does not help produce a $p_\perp^{-8}$ behavior.) But in the theory of QCD, there are a number of small scale-breaking effects to notice. The effective coupling constant falls logarithmically with energy. The incoming parton distributions should not scale, but at high x should fall and at small x rise as Q^2 increases. Effects in this direction are already seen in ep and μp scattering and have been analyzed in Refs. 17–20. An analogous modification of the fragmentation functions $D_q^h(z)$ is also expected theoretically. None of these effects alone change the effective apparent $p_\perp$ power index N_{eff} in $p_\perp^{-N_{\text{eff}}}$ very much and yet they all work in the same direction and together, as we show, they can change N_{eff} from the naive 4 to about 6 in the energy range of present experiments. (The scale breaking due to the large value of $\langle k_\perp \rangle_{h\to q}$ then brings N_{eff} to about 8 over this range.)

Thus the possibility exists that QCD can provide the full explanation of all the high-energy high-$p_\perp$ experimental results. We analyze this possibility in this paper. Some of our findings have been presented in Refs. 21 and 22. We wish to explain our approach in detail here. The net result is to demonstrate that this possibility is very real.[23]

The main problem in such an analysis is that no complete calculation of a prediction for QCD for any phenomenon—even qualitative ones such as the confinement of quarks—has yet been made. At present, the mathematical complexities are still too great. However, at very high energy or high momentum transfer Q, the theory is asymptotically free; the effective coupling constant falls with increasing Q^2. As emphasized by Politzer,[26] this permits calculation of those parts of a collision involving high Q^2. Yet every real process involves high and low Q^2 together and the precise separation of these parts and hence exact definition of the theory for hadron-hadron collisions is a problem for the (we hope near) future. We shall proceed here in a preliminary way.

What we do is include all the ingredients thought to be present from existing ideas on QCD. We assume that the effective coupling behaves as $\alpha_s(Q^2) = 12\pi/(25 \ln Q^2/\Lambda^2)$ with Λ determined from the scale breaking observed in ep and μp collisions. The distribution of constituents i (quarks and gluons) in the proton $G_{p\to i}(x, Q^2)$ and their fragmentation functions $D_i^h(z, Q^2)$ are given a Q^2 dependence in accord with QCD analyses of ep and μp collisions. The theory gives formulas for scaling violations (Q^2 dependences), but the functions must be known at some nominal reference momenta, say Q_0^2. The distributions $G_{p\to i}(x, Q_0^2)$ are determined from fits to the ep and μp data. For the quark fragmentation functions $D(z, Q_0^2)$ we use the distributions of our previous quark-jet paper[27] and take $Q_0^2 = 4$ GeV^2. We hope that data on the quark fragmentation functions from e^+e^-, ep, μp, or νp experiments will soon be available to test the Q^2 dependence expected from QCD and to allow for a more precise determination of them.

For the fundamental constituent cross sections for quark-quark, quark-gluon, and gluon-gluon scattering, we have taken the first-order perturbation scattering expected from QCD[28,29] (see Table I) and normalized absolutely by the effective coupling $\alpha_s(Q^2)$. This replaces the arbitrary size, energy dependence, and angular dependence of the black box in our previous papers.

Thus it would seem that we have little freedom of arbitrary choice. This would be true were it not for the distributions of gluons in the proton $G_{h\to g}(x, Q_0^2)$, and the distribution of hadrons in a gluon jet $D_g^h(z, Q_0^2)$, at the reference momenta. These two functions are not constrained much by other experiments and are thus essentially arbitrary. We have chosen these with an eye to experiment. In particular, we have chosen $D_g^h(z, Q_0^2)$

TABLE I. Cross sections for the various constituent quark-quark, quark-gluon, and gluon-gluon subprocesses.[a] The differential cross section is given by $d\hat{\sigma}/d\hat{t} = \pi\alpha_s^2(Q^2)|A|^2/\hat{s}^2$, where $\alpha_s(Q^2)$ is the effective coupling given by Eq. (3.1).

	Subprocess	$\lvert A\rvert^2$
1.	$q_iq_j \to q_iq_j$ $q_i\bar{q}_j \to q_i\bar{q}_j$ $(i \neq j)$	$\frac{4}{9}\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}$
2.	$q_iq_i \to q_iq_i$	$\frac{4}{9}\left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}+\frac{\hat{s}^2+\hat{t}^2}{\hat{u}^2}\right)-\frac{8}{27}\frac{\hat{s}^2}{\hat{u}\hat{t}}$
3.	$q_i\bar{q}_i \to q_i\bar{q}_i$	$\frac{4}{9}\left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}+\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}\right)-\frac{8}{27}\frac{\hat{u}^2}{\hat{s}\hat{t}}$
4.	$q_i\bar{q}_i \to gg$	$\frac{32}{27}\left(\frac{\hat{u}^2+\hat{t}^2}{\hat{u}\hat{t}}\right)-\frac{8}{3}\left(\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}\right)$
5.	$gg \to q_i\bar{q}_i$	$\frac{1}{6}\left(\frac{\hat{u}^2+\hat{t}^2}{\hat{u}\hat{t}}\right)-\frac{3}{8}\left(\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}\right)$
6.	$q_ig \to q_ig$	$-\frac{4}{9}\left(\frac{\hat{u}^2+\hat{s}^2}{\hat{u}\hat{s}}\right)+\left(\frac{\hat{u}^2+\hat{s}^2}{\hat{t}^2}\right)$
7.	$gg \to gg$	$\frac{9}{2}\left(3-\frac{\hat{u}\hat{t}}{\hat{s}^2}-\frac{\hat{u}\hat{s}}{\hat{t}^2}-\frac{\hat{s}\hat{t}}{\hat{u}^2}\right)$

[a] This table is identical to that in Ref. 29.

"softer" than $D_q^h(z, Q_0^2)$ because of experimental features of high-$p_\perp$ processes, and our success depends on this choice in several of the comparisons.

The theory of QCD also may explain how P_{out} can be so large.[26,30-33] For example, sometimes two quarks hit and scatter to two quarks plus a relatively hard radiated gluon, so that the two outgoing quark jets are out of momentum coplanarity by the momentum of the gluon. These effects can be calculated, and we are engaged in such calculations. What we have done here is a temporary expedient. We have simply taken the $k_\perp$ distribution measured for $\mu^+\mu^-$ pairs in pp collisions (where similar gluon emissions are possible) as a measure of an "effective" $k_\perp$ of quarks in the initial hadrons to mimic the effect of such $2-3$ constituent processes. This is not precisely correct and the Q^2 and x dependence of the high-$k_\perp$ tail to the effective transverse momentum of quarks in the hadrons is not handled properly in this manner. We hope to improve on this at a later date.

We begin in Sec. II by reviewing the successes of the quark-quark black-box approach and to examine closely its failures. The ingredients used in our QCD approach to high-$p_\perp$ processes are explained in Sec. III and the results presented in Sec. IV. We reserve Sec. V for summary and conclusions. The agreement with experiment is very satisfactory. Quantum chromodynamics might well be the correct theory behind these phenomena.

II. THE QUARK-QUARK SCATTERING BLACK-BOX APPROACH[1-10]

A. Successes

In spite of the rather *ad hoc* way in which we adjusted the quark-quark scattering cross section $d\hat{\sigma}/d\hat{t}$, many predictions of the black-box approach did not depend sensitively on it and were in agreement with experiment. As discussed in the summary of FFF, the conclusions that did not depend strongly on the precise value of the internal transverse momenta of the quarks with hadrons were the most successful. They include:

(1) Predictions for the large-$p_\perp$ single-particle ratios were quite successful. The π^+/π^- (and K^+/K^-) ratios were predicted to become larger at high $x_\perp$ since in this region the constituents that collide are predominantly u quarks that fragment more often into π^+ (K^+) than π^- (K^-). Similarly, the mixing of the η meson implied that η/π^0 should be about $\frac{1}{2}$ at high $x_\perp$ also in agreement with data. The high-$x_\perp$ K^+/π^+ ratio was used to deduce that $D_u^{K^+}(z)/D_u^{\pi^+}(z)$ must be about $\frac{1}{2}$ at large z. Recent lepton data are consistent with this deduction.[34]

(2) The number of π^0's[35] and jets[36-38] produced with a pion or proton beam is consistent with the expectation that the quarks in a pion carry, on the average, more momentum than they do in a proton.

(3) The quark-scattering approach predicted that the cross section for producing a jet of hadrons is considerably larger than that for producing a single meson at the same $p_\perp$.[39] For example, the model predicted a jet (quark) to single particle (π^0) ratio of about 370 at $x_\perp = 0.4$ and $\theta_{c.m.} = 90°$. The expectations were in good agreement with subsequent jet trigger experiments.[37,40]

(4) A distinctive feature of the model was that high-$p_\perp$ particles are not isolated but members of a cluster (jet) of particles representing the fragmentation of the quark. In single-particle triggers, one expected to see the remainder of the jet as an enhancement of associated particles in roughly the same direction as the trigger. With one additional assumption[2,41] about the character of fragmentation, the number of particles accompanying a large-$p_\perp$ trigger could be successfully understood.

(5) There is now considerable experimental support for the overall four jet structure shown in Fig. 1 for a large-$p_\perp$ event. These agreements between the quark scattering approach and experi-

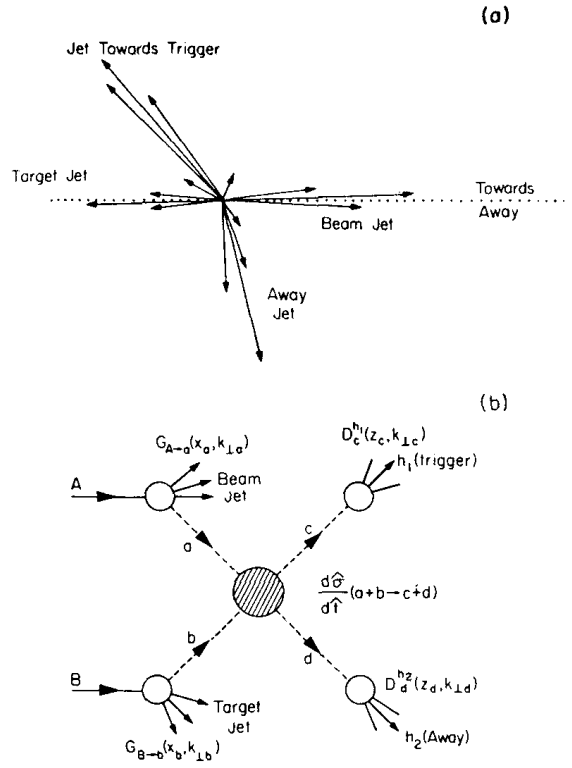


FIG. 1. (a) Illustration of the four-jet structure resulting from a beam hadron (entering at left along dotted line) colliding with a target hadron (entering at right along dotted line) in the c.m. frame: two jets with large $p_\perp$ (collection of particles moving roughly in the same direction), one called the "toward" (trigger) side and one on the "away" side; and two jets with small $p_\perp$ that result from the breakup of the beam and target hadrons (usually referred to as the "soft hadronic" background). (b) Illustration of the underlying structure of the large-$p_\perp$ process $A+B\to h_1+h_2+X$. The large-$p_\perp$ trigger hadron h_1 occurs as the result of a large-angle scattering of constituents $(q_a+q_b\to q_c+q_d)$, followed by the decay or fragmentation of constituent c into a towards-side jet of hadrons (one being the trigger h_1) and constituent d into an away-side jet of hadrons (one being h_2). The quantities x_a, x_b, $k_{\perp a}$, $k_{\perp b}$ are the longitudinal fraction of the incoming hadrons A, B momentum and perpendicular momentum of constituents a, b and z_c, z_d $k_{\perp c}$, $k_{\perp d}$ are the fraction of the outgoing constituents longitudinal momentum and perpendicular momentum carried by the detected hadrons h_1 and h_2.

ment indicate strongly that quarks play an important role in the production of high-$p_\perp$ mesons. In particular, they show that the mesons responsible for high-$p_\perp$ triggers probably arise from quarks that have fragmented in a manner similar to that observed in lepton-initiated processes.

B. Failures

As noted in the Introduction, the quark scattering black-box model does not agree in detail with all the results of high-$p_\perp$ experiments. It disagrees with data in the following ways.[38]

(1) Large-$p_\perp$ events are far less coplanar than first expected from a two-to-two scattering subprocess as shown in Fig. 1. Our first guess[2,41] that $\langle k_\perp\rangle_{h\to q}\approx\langle k_\perp\rangle_{q\to h}=330$ MeV resulted in a too narrow away-side $P_{\rm out}$ distribution. Even our final choice in FFF of $\langle k_\perp\rangle_{h\to q}=500$ MeV yields more coplanarity than seen in recent experiments.

(2) The quark-quark scattering model predicts too many high-$p_\perp$ particles on the away-side of a large-$p_\perp$ trigger. For example, the number of away hadrons with $z_p\geq 0.5$ (Ref. 42) for a $p_\perp=4.5$ GeV/c trigger at $\theta=45°$ and $W=53$ GeV is predicted in FFF to be 3–4 times larger than seen experimentally by the CCHK group.[15]

(3) Not only does the model predict too many away-side particles, it predicts many more positives than negatives on the away side. For a trigger $p_\perp$ of 3.0 GeV/c, $\theta_{\rm c.m.}=90°$ and $W=53$ GeV/c, the model predicts about 50% more positives than negatives on the away side with $p_\perp$ (away) >1.5 GeV/c. Recent data from the CERN ISR (Ref. 43) show about equal away-side positives and negatives under these ciruumstances. These are serious problems for the model. The last two imply that (at the small-$x_\perp$ values probed by ISR experiments) the recoiling away-side parton does not fragment in a manner similar to that observed by lepton experiments. Furthermore, as discussed in FFF Table 4, we cannot simply increase $\langle k_\perp\rangle_{h\to q}$ to improve our agreement with (1) and (2). The value of 500 MeV was as large as we could take in FFF without spoiling our agreement with the energy dependence of the single-particle cross section. We feel that it is not useful at present to try to "fiddle" the quark scattering model to make it agree with recent experiments, particularly since there is an emerging candidate theory of strong interactions that apparently has the features necessary to repair the failures of the black-box model.

III. THE INGREDIENTS TO THE QCD APPROACH

A. Effective coupling $\alpha_s(Q^2)$

The effective strong-interaction coupling constant falls logarithmically with increasing Q^2, where Q is some characteristic momentum in a collision. In general, the effective quark-quark-gluon coupling is expected to have the form

$$\alpha_s(Q^2)\equiv\frac{g^2}{4\pi}=\frac{12\pi}{(33-2n_f)[\ln(Q^2/\Lambda^2)+C]}\,, \qquad (3.1)$$

where n_f is the number of quark flavors (we use $n_f=4$). The constant C represents corrections that, in general, differ from process to process

but can, in principle, be calculated (although this might be quite difficult in practice).[44] The quantity Λ is an unknown scale factor that can be determined from the amount of "scale breaking" observed in a given experiment. Analysis of the scale breaking in ep and μp collisions indicates that Λ is in the range 0.3 to 0.7 GeV/c (with $C=0$).[17-20]

For ep collisions, Q is the four-momentum transfer from the electron to the quark. On the other hand, the correct kinematic quantity to use for Q^2 in the constituent subprocess shown in Fig. 1 is not known. This problem is, of course, related to the unknown $\pm C$ in (3.1). For definiteness, we will take $C=0$ and choose

$$Q^2 = 2\hat{s}\hat{t}\hat{u}/(\hat{s}^2+\hat{t}^2+\hat{u}^2), \tag{3.2}$$

where $\hat{s}$, $\hat{t}$, and $\hat{u}$ are the usual Mandelstam s, t, and u invariants but for the constituent subprocess. This form for Q^2 is symmetric in $\hat{s}$, $\hat{t}$, and $\hat{u}$ and approaches $-\hat{t}$ in the case $\hat{t} \ll \hat{s}$. This uncertainty in the form for Q^2 and, correspondingly, the lack of knowledge of $\pm C$ makes predictions at low Q^2 (i.e., low $p_\perp$) in hadron-hadron collisions a bit uncertain.

If the distribution of quarks within the proton, $G_{p\to q}(x)$, and the fragmentation of quarks in hadrons, $D_q^h(z)$, both scale, then the invariant cross section $Ed\sigma/d^3p$ for producing a large-$p_\perp$ meson reflects directly the energy dependence of the quark-quark cross section $d\hat{\sigma}/d\hat{t}$. Thus if the latter behaves as $h(\hat{t}/\hat{s})/\hat{s}^n$, then the former behaves as $f(x_\perp, \theta_{\text{c.m.}})/p_\perp{}^{2n}$. The cross section for the scattering of partons in field theory (see Table I) with α_s = constant yield $2n = N_{\text{eff}} = 4$ where we define

$$N_{\text{eff}} = -\ln(\sigma_1/\sigma_2)/\ln(p_{\perp 1}/p_{\perp 2}), \tag{3.3}$$

where $\sigma_{1,2}$ is the invariant cross section (at fixed $x_\perp$) at $p_{\perp 1,2}$. Including an α_s that depends on Q^2 according to (3.1) and (3.2) produces an $Ed\sigma/d^3p$ that decreases faster than $1/p_\perp{}^4$ at small $p_\perp$ ($N_{\text{eff}} \approx 4.8$ for $2 \leq p_\perp \leq 10$ GeV/c at $x_\perp = 0.2$) but approaches the $1/p_\perp{}^4$ behavior at large $p_\perp$. This can be seen by the dot-dash curve in Fig. 2 where we plot $p_\perp{}^8$ times the predicted $Ed\sigma/d^3p$ arising from quark-quark scattering at $x_\perp = 0.2$ versus $p_\perp$ using $\Lambda = 0.4$ GeV/c. One sees that including the $\alpha_s(Q^2)$ dependence brings one a small way toward the flat $(1/p_\perp{}^8)$ dependence seen experimentally at $p_\perp \lesssim 6$ GeV/c.

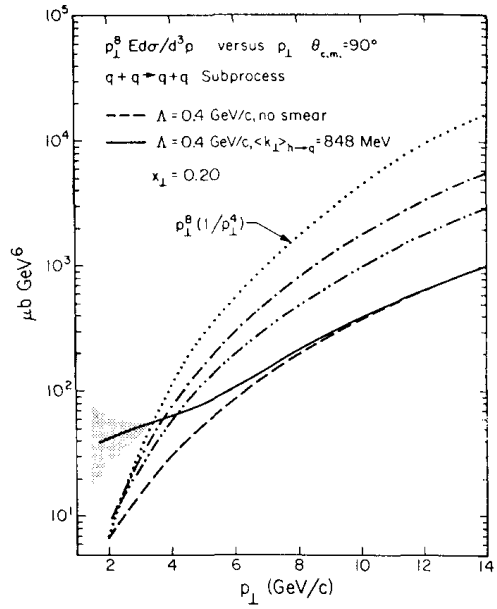


FIG. 2. The behavior of $p_\perp{}^8$ times the invariant cross section $Ed\sigma/d^3p$ for $pp \to \pi^0 + X$ at $\theta_{\text{c.m.}} = 90°$ and $x_\perp = 0.2$ arising from the QCD subprocess $q+q \to q+q$ calculated with $\Lambda = 0.4$ Gev/c. If the quark distributions within protons, $G_{p\to q}(x)$, and the quark fragmentation functions, $D_q^{\pi^0}(z)$, scale and if the strong interaction coupling α_s is constant, then $p_\perp{}^8 Ed\sigma/d^3p$ behaves like $p_\perp{}^4$ at fixed $x_\perp$ and $\theta_{\text{c.m.}}$ (dotted line). Allowing α_s to depend on Q^2 according to (3.1) yields the dot-dashed curves. Including the expected Q^2 dependence $\alpha_s(Q^2)$ and $G_{p\to q}(x, Q^2)$ results in the dot-dot-dashed curves. Finally, allowing $\alpha_s(Q^2)$, $G_{p\to q}(x, Q^2)$ and $D_q^{\pi^0}(z, Q^2)$ all to vary with Q^2 in a manner expected from QCD results in dashed curve and the solid curve is the result after smearing with $\langle k_\perp \rangle_{h\to q} = 848$ MeV and $\langle k_\perp \rangle_{q\to h} = 439$ MeV. The shaded area represents uncertainties due to the way in which one cuts off the low $\hat{s}$ and $\hat{t}$ singularities in $d\hat{\sigma}/d\hat{t}$. Above $p_\perp \gtrsim 3.5$ GeV/c at $\theta_{\text{c.m.}} = 90°$, the results are insensitive to the details of this cutoff procedure.

B. The quark and gluon distributions $G(x, Q^2)$

In the QCD approach, the "effective" quark distributions in a proton, $G_{p\to q}(x, Q^2)$, do not scale. The influence of this on $\nu W_2(x, Q^2)$ for ep and μp scattering has been studied[17-20] and may account for the lack of scaling seen in these experiments over the range $4.0 \lesssim Q^2 \lesssim 20.0$ GeV^2. We use the formulation of Ref. 18 to extrapolate these functions to the higher-Q^2 region needed in analyzing high-$p_\perp$ hadron data (see Table II for the range of Q^2 sampled).

In an asymptotically free field theory, scale violations are generated by gluon corrections typified by gluon bremstrahlung from a quark and by quark-anciquark pair creation by a gluon. One can predict the behavior of the constituent distributions $G_i(x, Q^2)$ given that they are known at some reference momenta Q_0 [large enough so that $\alpha_s(Q_0{}^2)$ is small enough to make perturbation theory applicable]. Following Ref. 18, the moments of the distributions

TABLE II. The mean values of z_c, Q_x, and Q^2 resulting for $pp \to \pi^0 + X$ at 90° in the QCD approach with $\Lambda = 0.4$ GeV/c, and where z_c is the fraction of the constituent momenta carried by the trigger hadron (see Fig. 1); Q_x is the component of momentum of the constituent scattering toward the trigger ($Q_x = k_{xa} + k_{xb}$, see Table I and Fig. 2 of FFF); Q^2 is defined by Eq. (3.2).

W (GeV)	$P_\perp$ (GeV/c)	$\langle z_c \rangle$	$\langle Q_x \rangle$ (GeV)	$\langle Q^2 \rangle$ (GeV^2)
53	2.0	0.68	1.25	7.9
53	4.0	0.75	0.86	33.7
53	7.0	0.81	0.61	96.5
53	9.0	0.84	0.54	149.0
19.4	1.94	0.74	1.53	4.8
19.4	3.0	0.86	1.77	7.6
19.4	4.0	0.89	1.76	15.5
19.4	6.0	0.92	1.46	42.9
19.4	7.0	0.94	1.46	58.5
500	10.0	0.58	0.16	525
500	30.0	0.69	0.10	3003
1000	10.0	0.53	0.21	751
1000	30.0	0.61	0.10	4184

$$M_i(n, Q^2) = \int_0^1 x^n G_i(x, Q^2) dx \qquad (3.4)$$

are given in terms of the moments at Q_0 by

$$M_j(n, Q^2) = \sum_{i=1}^{9} M_i(n, Q_0^2) R_{ij}(n, Q^2, Q_0^2), \qquad (3.5)$$

where $R_{ij}(n, Q^2, Q_0^2)$ is a known matrix (depending on Λ) and i corresponds to the constituent types (u, d, s, c, $\bar{u}$, $\bar{d}$, $\bar{s}$, $\bar{c}$, gluon). The final resulting distributions at Q^2 are calculated by inverting (3.4) by an inverse Mellin transform [Eq. (13) of Ref. 18].

Figure 3 shows the predicted Q^2 behavior of $\nu W_2(x, Q^2)$ resulting from an analysis of the ep and μp data. The x dependence of the parton distributions at the reference momentum, $G_i(x, Q_0^2 = 4\ \text{GeV}^2)$, was chosen to agree with experiment. Unfortunately, the analysis of ep and μp is relatively insensitive to the input gluon distribution with the proton. We take $xG_1(x, Q_0^2) = xG_{p\to g}(x, Q_0^2) = (1+9x)(1-x)^4$ at the reference momentum. It integrates to a total momentum for gluons within the proton of 50%. The resulting Q^2 dependence of $G_{p\to g}(x, Q^2)$ is also shown in Fig. 3. Both $\nu W_2(x, Q^2)$ and $xG_{p\to g}(x, Q^2)$ exhibit a rise at small x and a decrease at large x as Q^2 increases. The effect is particularly dramatic for the latter.

The QCD interpretation of the ep and μp inelastic-scattering data has some ambiguities because one expects, not only the logarithmic scale breaking shown in Fig. 3, but also other corrections falling more rapidly with Q^2. The latter would be unimportant at very large $Q^2 \gtrsim 50$ (GeV/c)^2 but are important in the Q^2 range probed by the current data. One example of such a correction is the $O(m^2/Q^2)$ correction (m is proton mass) generated by using x' and not x as the argument of the structure functions. Here

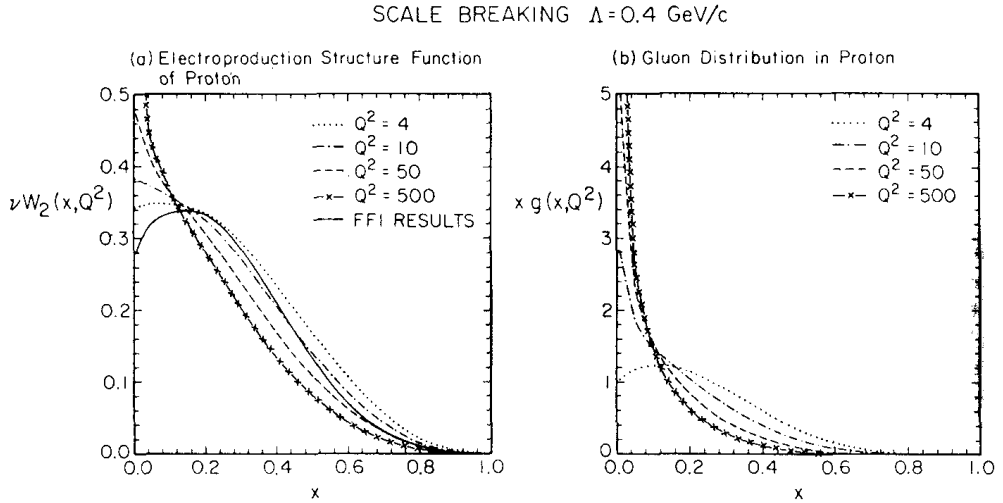


FIG. 3. (a) The predicted Q^2 dependence (scale breaking) of the electroproduction structure function for the proton $\nu W_2(x, Q^2)$ arising from the constituent (quarks, antiquarks and gluons) distributions $G_i(x, Q^2)$ used in this analysis. The distributions at high Q^2 are calculated from the distributions at the reference momentum $Q_0^2 = 4$ GeV^2 using a QCD moment analysis with $\Lambda = 0.4$ GeV/c. In asymptotically free theories, one expects a decrease in the number of high-x constituents and an increase in the number of low-x constituents as Q^2 increases. Also shown is the value of $\nu W_2(x)$ (independent of Q^2) used in the quark-quark black box model of FF1. (b) The predicted Q^2 dependence of the distribution of gluons within the proton $xG_{p\to g}(x, Q^2) = xg(x, Q^2)$ used in this analysis. The distribution at high Q^2 is calculated in terms of a distribution at the reference momentum $Q_0^2 = 4$ GeV^2 given by $xg(x, Q_0^2) = (1+9x)(1-x)^4$.

$$x' = \frac{x}{1+m^2x/Q^2} \sim x - m^2x^2/Q^2 . \qquad (3.6)$$

This leads to scale breaking that is about half the observed amount in the large $x \gtrsim 0.3$ region. However, one can construct variables of a similar type (for instance x_{super} introduced by Atwood[45]) that can describe all the observed large-x scale breaking in terms of $O(m^2/Q^2)$ corrections. Such models give essentially no scale breaking at low x and the important feature of the QCD approach is that the same value of Λ describes the scale breaking at both low and high x. This is shown in Fig. 4. It should be noted that the fit to the low-x data is poorer than that at high x although the trend with Q^2 is given well in both cases. This is because the low-x data is new data that was not available when the parameters of the QCD solution were determined.[46] In Ref. 18, we not only considered structure functions that were a function of x but also a more complicated and probably more realistic formalism developed by Georgi and Politzer.[17] This includes $O(m^2/Q^2)$ terms not only in the argument of the structure functions [similar to (3.6) above] but also as overall multiplicative factors. It turns out that the $O(m^2/Q^2)$ terms tend to cancel among themselves and this formalism is phenomenologically equivalent to the simple formulation we use here. In particular, essentially the same value of $\Lambda \approx 0.5$ GeV/c is found in the best fit of both formalisms.

If one includes the scale-breaking effects of $G_i(x, Q^2)$ in addition to the running coupling constant $\alpha_s(Q^2)$, the resulting $pp \to \pi^0 + X$ cross section arising from the quark-quark subprocess has an N_{eff} equal to about 5.0 and 5.5 for the range $2.0 \leq p_\perp \leq 10.0$ at $x_\perp = 0.2$ and 0.5, respectively. The scale breaking of $G_i(x, Q^2)$ has little effect at $x_\perp = 0.2$ (see Fig. 2) because at this $x_\perp$ one is sensitive to $G_i(x, Q^2)$ near the values of x that are stationary as Q^2 increases.

C. The fragmentation functions $D_i^h(z, Q^2)$

The experimentally measurable constituent fragmentation functions $D_i^h(z, Q^2)$ (here i refers to a gluon or a $u, d, s, c, \bar{u}, \bar{d}, \bar{s}, \bar{c}$ quark) are expected, in asymptotically free theories, to show scale breaking (Q^2 dependences) similar to that predicted for $\nu W_2(x, Q^2)$.[26,47] The moments of the fragmentation functions to a given hadron h given by

$$\bar{M}_i^h(n, Q^2) = \int_0^1 z^n D_i^h(z, Q^2) dz , \qquad (3.7)$$

are given in terms of the moments at some reference momentum Q_0 by an equation similar to

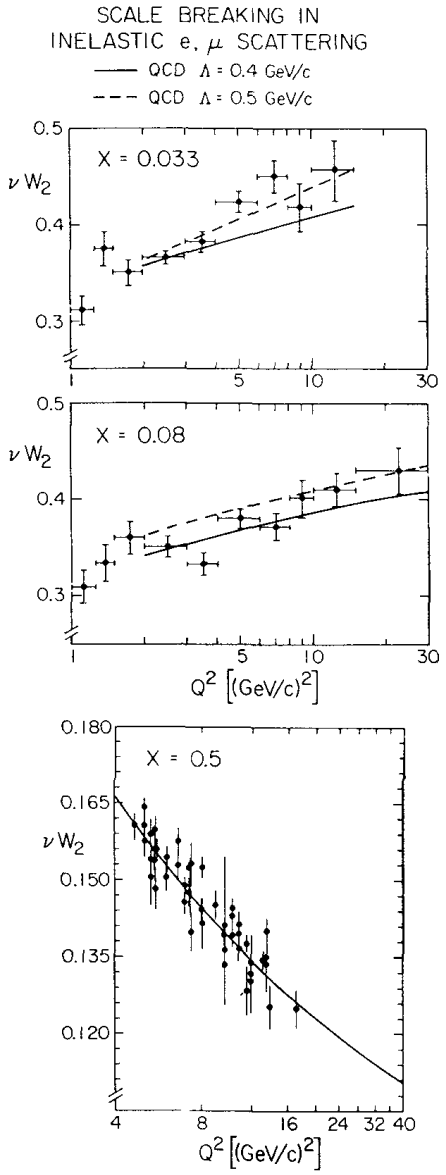


FIG. 4. Comparison of the scale-breaking effects (Q^2 dependence) expected from an asymptotically free theory with data on ep and μp inelastic scattering at $x = 0.033$ and 0.08 (Ref. 20) and at $x = 0.5$ (Ref. 72). The theory comes from the analysis of Ref. 18 using $\Lambda = 0.4$ GeV/c (solid curve) and $\Lambda = 0.5$ GeV/c (dashed curve).

(3.6). Namely,

$$\bar{M}_j^h(n, Q^2) = \sum_{i=1}^{9} \bar{M}_i^h(n, Q_0^2) \bar{R}_{ij}(n, Q^2, Q_0^2) , \qquad (3.8)$$

where the matrix $\bar{R}_{ij}$ is simply related (but not equal to) R_{ij}.[48] One then uses the Mellin-transform technique of Ref. 18 to invert (3.7) and obtains $D_i^h(z, Q^2)$ in terms of these functions at the reference momentum Q_0^2 (we take $Q_0^2 = 4$ GeV^2).

As explained by Gross,[49] the Q^2 dependence of a distribution function depends on its shape at the reference momenta. The faster the function $D(z, Q_0^2)$ falls off with increasing z at large z, the faster the large z points fall as Q^2 is increased above Q_0^2. It is important when considering the fragmentation functions to distinguish the distribution of primary (or direct) mesons from the final net distribution (which includes decay products). The above moment analysis should be applied to the former not the latter. This is, of course, a bit difficult since we do not know experimentally exactly how many resonances are in the quark jets at Q_0^2. What we shall do is to use the results of Ref. 27 (hereafter called FF2) at the reference momentum $Q_0^2 = 4$ GeV^2. The distribution of primary mesons at Q_0^2 is then given by

$$D_q^h(z, Q_0^2) = A_q^h f(1-z) + B^h \bar{F}(z), \tag{3.9}$$

with A_q^h and B^h given in Table I of FF2 and

$$\bar{F}(z) = F(z) - f(1-z) \tag{3.10}$$

with

$$f(1-z) = f(\eta) = 1 - a + 3a\eta^2 .$$

The parameter a is chosen to be 0.77 and $F(z)$ is given by Eq. (2.23) in FF2. The distribution of primary mesons in a gluon jet at Q_0^2 is assumed to have the form

$$D_g^h(z, Q_0^2) = B^h F_g(z), \tag{3.11}$$

where we arbitrarily take

$$F_g(z) = 3(1-z)^2/z, \tag{3.12}$$

where we have assumed, as discussed earlier, that the gluon fragmentation function falls off faster with increasing z than do the quark functions.

We take the primary mesons precisely as explained in FF2 as being either pseudoscalars (π, K, etc.) or vector mesons (ρ, K^*, etc.) with equal likelihood. We use the QCD moment method to calculate the primary meson (pseudoscalar plus vector meson) distributions at any desired Q^2. The resonances are then allowed to decay and we form the total net (direct + indirect) distributions at that Q^2. Typical results are shown in Fig. 5.

The effect on the predicted large-$p_\perp$ invariant cross section of including scale violations of the fragmentation functions is that the N_{eff} now becomes 5.8 and 6.4 between $p_\perp = 2$ and 10 GeV/c at $x_\perp = 0.2$ and 0.5, respectively. Large-$p_\perp$ results are particularly sensitive to scale violations of the $D(z, Q^2)$ function since these violations are largest at high z (see Fig. 5) and this is precisely the region sampled by the calculations.

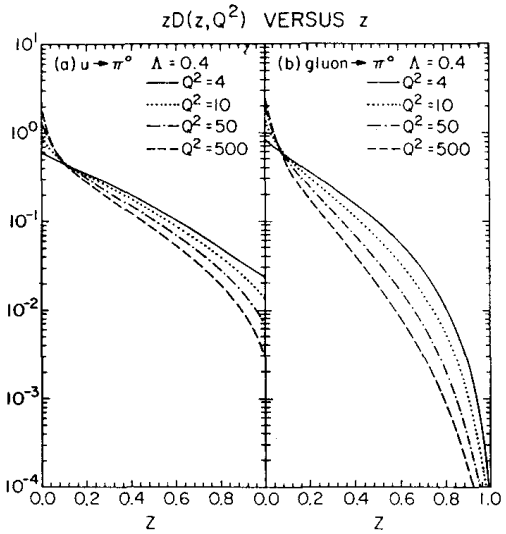


FIG. 5. The Q^2 dependence of the fragmentation function for a u quark to a π^0, $D_u^{\pi^0}(z, Q^2)$, expected from an asymptotically free theory. The distributions at high Q^2 are claculated from the distribution at the reference momentum $Q_0^2 = 4$ GeV^2 using $\Lambda = 0.4$ GeV/c, where $D_q^h(z, Q_0^2)$ is taken from the analysis in FF2. (b) Same as (a) but for the gluon fragmentation function $D_g^{\pi^0}(z, Q^2)$.

D. Transverse momentum

As we learned in FFF, constituents have a large internal transverse momentum inside the proton. Such effects (called smearing) are particularly important for large-$p_\perp$ calculations, due to the "trigger bias" which selects the configuration in which the initial quarks (or gluons) are already moving toward the trigger[11-15,41] (see Fig. 3 of FFF). In QCD, this transverse momentum of the partons can arise from two sources illustrated in Fig. 6.

Firstly, in a proton beam, quarks are confined in the transverse direction to within the proton radius. Therefore, from the uncertainty principle, they must have some transverse momentum. This momentum is intrinsic to the basic parton "wave function" inside the proton. As illustrated in Fig. 6(a), one might expect the wave function to have a term where the trigger parton $k_\perp$ is balanced by another constituent (or constituents) which has the opposite $k_\perp$ and most of the remaining longitudinal momentum. Consider now the plane formed by the beam, target, and a 90° trigger hadron (called the x-z plane in Fig. 6). Typically, the trigger arises from the fragmentation of a constituent with $k_{\perp x} > 0$ which is balanced by the remaining constituents having $k_{\perp x} < 0$. One expects to see this negative $k_{\perp x}$ as a shift in the beam and target jets at large $|x_\parallel|$. This shift (i.e., nonzero $\langle k_{\perp x} \rangle$) of the beam jet as one in-

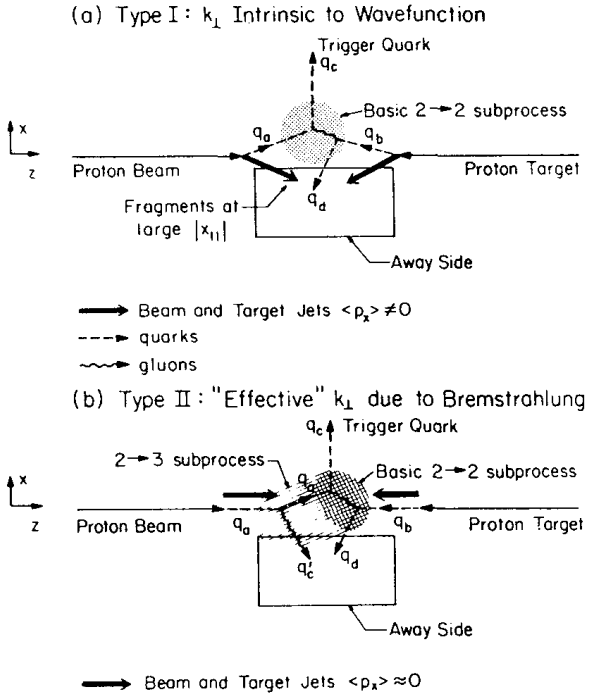


FIG. 6. (a) Illustration of the nonperturbative component of the transverse momentum of quarks within proton that is intrinsic to the wave function of the proton. One expects this transverse momentum to be balanced by the remaining constituents in the proton which can, in turn, fragment into particles at high x_{11}. The away-side consists of the recoiling quark q_d and two slightly shifted jets, one from the beam and one from the target. (b) Illustration of the perturbative component to the transverse momentum of a quark with a hadron which is due to the bremstrahlung of a gluon before the basic $2 \to 2$ scattering occurs. In this case, the trigger quark is balanced by two away-side jets, one from the quark q_d and from the radiated gluon q_c'.

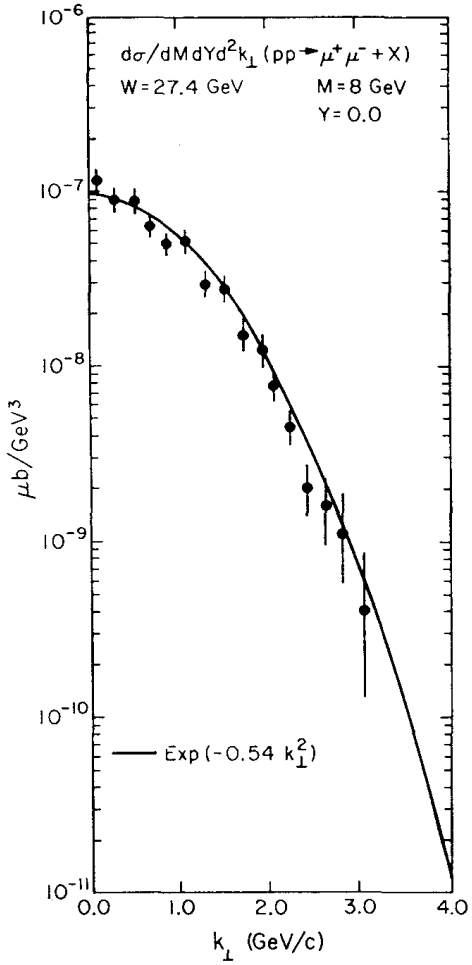


FIG. 7. The transverse-momentum spectrum, $d\sigma/dMdYd^2k_\perp$, of muon pairs in pp collisions at $W = 27.4$ GeV, $M_{\mu\mu} = 8$ GeV, and rapidity $Y = 0$ from Ref. 51. Also shown is a Gaussian fit of the form $\exp(-0.54\ k_\perp^2)$ which yields $\langle k_\perp \rangle_{\mu\mu} \simeq 1.2$ GeV and is interpreted as implying $\langle k_\perp \rangle_{h \to q} = 848$ MeV.

creases the $p_\perp$ of a 90° trigger has recently been observed by the British-French-Scandinavian (BFS) group at ISR[50] (see Fig. 14 or Ref. 22).

Secondly, in QCD, one expects to receive an "effective" $k_\perp$ of quarks in protons due to the bremstrahlung of gluons. This perturbation term is illustrated in Fig. 6(b). It corresponds to including two particle to three or more particle processes $(2-3)$ rather than just the two particle to two particle $2-2$ scatterings. For such subprocesses, the $k_\perp$ of the quark q_a is balanced by a gluon jet on the away side which subsequently fragments into many low-momentum hadrons. In addition, the mean value of the effective $k_\perp$ is expected to depend on the x value of quark q_a and on the Q^2 for the processes. Separating the origin of the transverse momenta into Types I and II as seen in Fig. 6 is a bit artificial since both mechanisms occur simultaneously.

The effective constituent transverse momentum is directly observed in the Drell-Yan process $pp \to \mu^+\mu^- + X$. Current data[51] indicate that $\langle k_\perp \rangle_{\mu^+\mu^-}$ is about 1.2 GeV (see Fig. 7). There has been much speculation about how much of the dimuon $k_\perp$ spectra shown in Fig. 7 is due to the wave function (Type I) and how much is explained by QCD perturbation calculations (Type II).[26,30-33] The latter predicts a high-$k_\perp$ tail to the distribution that falls roughly like a power and a mean that depends both on x and Q^2 of the muon pair. For the present analysis, we shall parameterize the transverse momentum of the constituents in protons by a Gaussian with $\langle k_\perp \rangle_{h \to q} = 848$ MeV which produces for the Drell-Yan subprocess the curve shown in Fig. 7. We shall take this distribution to be independent of x and Q^2 and to be the

same for quarks, antiquarks, and gluons in the proton. In so doing, we are not handling properly the x and Q^2 dependence of the high-$k_\perp$ tail expected from QCD bremstrahlung. At a later date, we hope to calculate and include the $2 \to 3$ processes expected by QCD. For the present, we merely use the data in Fig. 7 to give us an "effective" $k_\perp$ distribution and include only $2 \to 2$ subprocesses.

In a manner similar to that illustrated in Fig. 6, the emission of gluons after the hard-scattering ($2 \to 2$) subprocesses induce an "effective" $k_\perp$ of the hadrons that fragment from the outgoing quarks because one is really seeing two jets rather than one. As for the quark distributions in the proton, we do not include these effects (we also neglect the interferences that arise between the amplitude for emitting gluon before and after the hard $2 \to 2$ process) and for the present take the transverse momentum distribution of hadrons from outgoing quarks (and gluons) to be a Gaussian with $\langle k_\perp \rangle_{q \to h} = 439$ MeV independent of z or Q^2 as in FF2. Again, this is not precisely correct and we hope to improve this in later work.

E. The cross sections $d\hat{\sigma}/d\hat{t}$

In the QCD approach, one includes not only the contributions from quark-quark scattering but also the contributions from quark-gluon and gluon-gluon scattering. We include all seven processes: $qq \to qq$, $\bar{q}q \to \bar{q}q$, $\bar{q}\bar{q} \to \bar{q}\bar{q}$, $gq \to gq$, $g\bar{q} \to g\bar{q}$, $gg \to \bar{q}q$, $\bar{q}q \to gg$, and $gg \to gg$, where g is a gluon. Each $2 \to 2$ differential cross section, $d\hat{\sigma}/d\hat{t}$, is calculated to first order in perturbation theory with an effective coupling constant $\alpha_s(Q^2)$ as in (3.1). These cross sections have been calculated previously by Cutler and Sivers[28] and by Combridge, Kripganz, and Ranft[29] and for completeness are given in Table I.[52] All these cross sections behave as $\hat{s}^{-2}$ at fixed $\hat{t}/\hat{s}$ (and for constant α_s) so that including gluons does not help in changing N_{eff} from 4 to 8 but gluons are important in increasing the magnitude of the low-$x_\perp$ cross section to agree with data.[13,28,29] In addition, we will see that gluons play an important role in understanding the high-$p_\perp$ correlation data.

Including gluons, unfortunately, introduces an uncertainty (at low $x_\perp$) in the high-$p_\perp$ predictions. As explained in Secs. III B and III C, the gluon distribution in a proton and the gluon fragmentation functions are essentially unknown. We only know that the total momentum carried by quarks and gluons in a proton is that of the proton and similarly the total energy carried by all the hadron fragments from a gluon is the gluon energy. Many of our high-$p_\perp$ predictions depend on these unknown distributions, for the QCD quark-gluon and gluon-gluon scattering cross sections are large. If one accepts QCD as the correct description for high-$p_\perp$ processes, then one could eventually hope to use the hadron-hadron data to help determine these functions. For the present, we simply calculate with our initial guesses for $G_{p \to g}(x, Q_0^2)$ and $D_g^h(z, Q_0^2)$ and do *not* attempt to find the optimal forms for these functions by fitting high-$p_\perp$ data.

As discussed in great detail in FFF, cross sections of the type given in Table I are not adequate once one allows for a nonzero transverse momentum of constituents in the hadrons or of hadrons from in the outgoing jets.[53] This is because they diverge at $\hat{s}$, $\hat{u}$, or $\hat{t}$ equal zero which can occur once $\langle k_\perp \rangle_{h \to q}$ or $\langle k_\perp \rangle_{q \to h}$ is nonzero. To remove this unwanted singularity in the integral, we simply replace $\hat{s}$, $\hat{t}$, and $\hat{u}$ in Table I by $\hat{s} + M_0^2$, $M_0^2 - \hat{t}$, and $M_0^2 - \hat{u}$, respectively, with $M_0^2 = 1.0$ GeV^2.[54] Because we are generating the transverse momentum of the constituents by a Gaussian, the results for hadron production at large $p_\perp$ are not sensitive to this ad hoc cutoff procedure. With the large-$k_\perp$ damping, characteristic of a Gaussian, once one has removed the infinity at say $\hat{t} = 0$ (by whatever means), one never samples this low-$\hat{t}$ region when calculating high-$p_\perp$ meson production. This would not be true for a transverse-momentum distribution falling off less rapidly with $k_\perp$ (for example, like a power) so that the large-$k_\perp$ tails expected by the QCD processes in Fig. 6 will have to be handled differently in the future.

Because $\langle k_\perp \rangle_{h \to q}$ is large, one does begin to become sensitive to the low-$\hat{s}$ and -$\hat{t}$ cutoff at low $p_\perp$. This is illustrated in Fig. 2 where we show the results for $pp \to \pi^0 + X$ at $x_\perp = 0.2$ and $\theta_{\text{c.m.}} = 90°$ arising from quark-quark scattering before and after smearing. The shaded area represents uncertainties due to varying the low-$\hat{s}$ and -$\hat{t}$ cutoffs. The region of $p_\perp \lesssim 3.5$ GeV/c cannot, at present, be used to quantitatively test the QCD ideas since this region is sensitive to the cutoff procedure.[55] For $p_\perp \gtrsim 3.5$ GeV/c, on the other hand, the manner of cutoff is not important and the result depends only on the amount of smearing (i.e., on $\langle k_\perp \rangle_{h \to q}$).

IV. RESULTS

A. The single-particle cross section

1. $p_\perp^{-8}$ behavior

Figure 8 shows a comparison of the predicted and experimental behavior of $p_\perp^8$ times $Ed\sigma/d^3p$ for $pp \to \pi + X$ at $\theta_{\text{c.m.}} = 90°$ and $x_\perp = 0.2$, 0.35, and 0.5 versus $p_\perp$. The dot-dashed and solid curves are the final results before and after smearing, respectively, with $\Lambda = 0.4$ GeV/c. The dashed

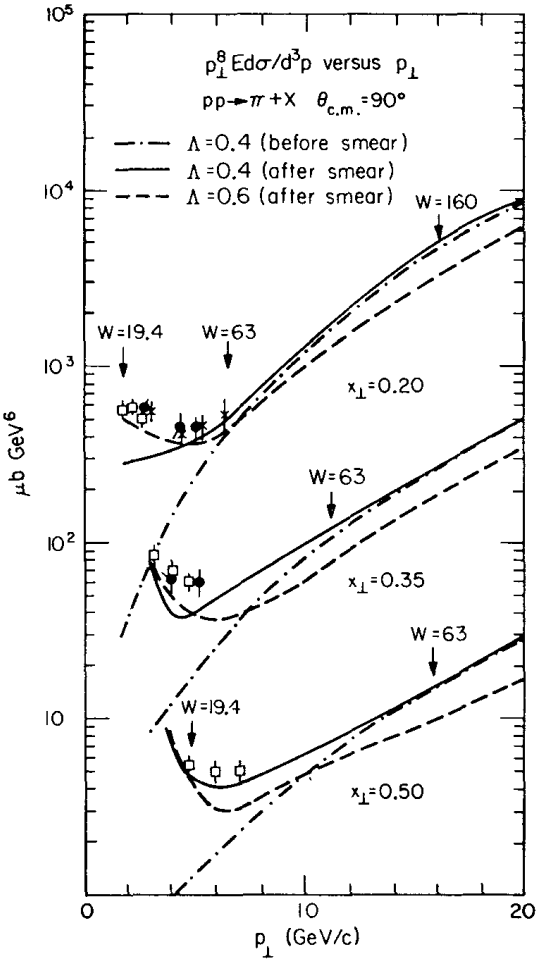


FIG. 8. The data on $p_\perp^8$ times $Ed\sigma/d^3p$ for large $p_\perp$ pion production at $\theta_{c.m.} = 90°$ and *fixed* $x_\perp = 0.2$, 0.35, and 0.5 versus $p_\perp$ (open squares: Ref. 73, solid dots: Ref. 74, crosses: Ref. 75) compared with the predictions (with absolute normalization) of a model that incorporates all the features expected from QCD. The dot-dashed and solid curves are the results before and after smearing, respectively, using $\Lambda = 0.4$ GeV/c and the dashed curves are the results using $\Lambda = 0.6$ GeV/c (after smearing).

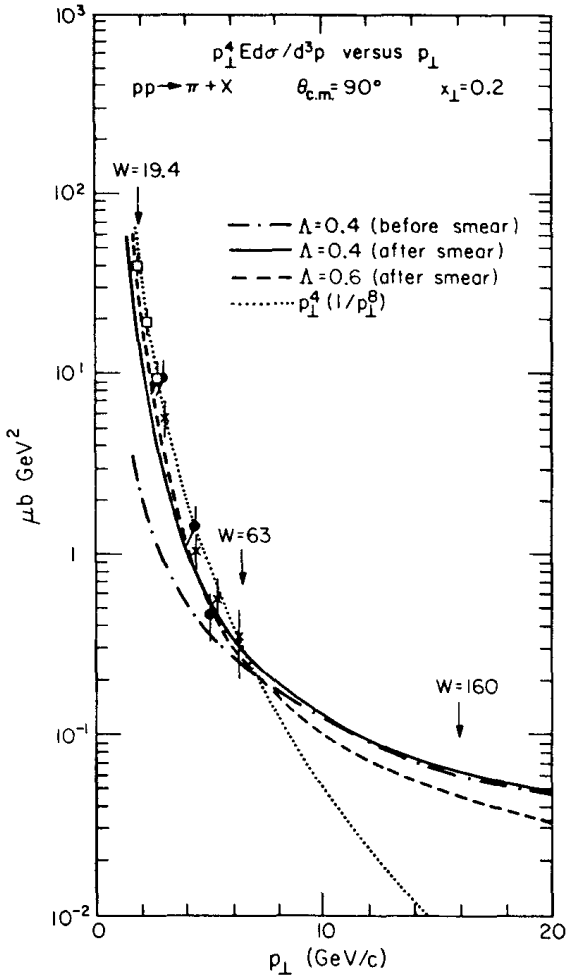


FIG. 9. The same as Fig. 8 except we now plot $p_\perp^4$ times $Ed\sigma/d^3p$ versus $p_\perp$ at $x_\perp = 0.2$ and $\theta_{c.m.} = 90°$. One clearly sees the asymptotic approach to an $Ed\sigma/d^3p$ α $p_\perp^{-4}$ behavior at fixed $x_\perp$.

curves are the results (after smearing) using $\Lambda = 0.6$ GeV/c. The effect of smearing (at fixed $x_\perp$) is to increase the low-$p_\perp$ predictions (by about a factor of 10 at $p_\perp = 2$ GeV/c and $x_\perp = 0.2$) while not affecting much the high-$p_\perp$ region. For the range $2.0 \leq p_\perp \leq 6.0$ GeV/c at $x_\perp = 0.2$, and $4.0 \leq p_\perp \leq 10.0$ GeV/c at $x_\perp = 0.5$, the results are roughly independent of $p_\perp$ (when multipled by $p_\perp^8$). However, this $p_\perp^{-8}$ behavior is only a "local" effect. It holds only over a small range of $p_\perp$ (at low $p_\perp$); the region depending somewhat on $x_\perp$. As $p_\perp$ increases, the predictions approach the expected $p_\perp^{-4}$ behavior. This can be seen more clearly in Fig. 9 where we plot the predictions and data times $p_\perp^4$ at $x_\perp = 0.2$ and $\theta_{c.m.} = 90°$. The behavior becomes $p_\perp^{-4}$ only asymptotically, but by $p_\perp = 10$ GeV/c at this $x_\perp$, it is fairly close (about $p_\perp^{-5}$).

As illustrated in Fig. 2, the low-$p_\perp$ region is sensitive to the small-$\hat{s}$ and -$\hat{t}$ cutoff employed.[55] However, because of the Gaussian falloff of the transverse momentum distributions, the results are completely insensitive to the form of the cutoff for $p_\perp \gtrsim 3.5$ GeV/c at $\theta_{c.m.} = 90°$. For example, Table II shows that the constituent subprocess has a mean momentum $\langle Q_x \rangle = 1.76$ GeV/c for a $p_\perp = 4$ GeV/c trigger at $W = 19.4$ GeV but even for this large "trigger bias", only 12% of the total $pp \to \pi^0 + X$ cross section arises from the region $|\hat{t}| < 10$ GeV^2 and none arises for $|\hat{t}| < 5$ GeV^2.

The data on $Ed\sigma/d^3p$ at *fixed* $W = 19.4$ and 53 GeV versus $p_\perp$ are compared with the theory in

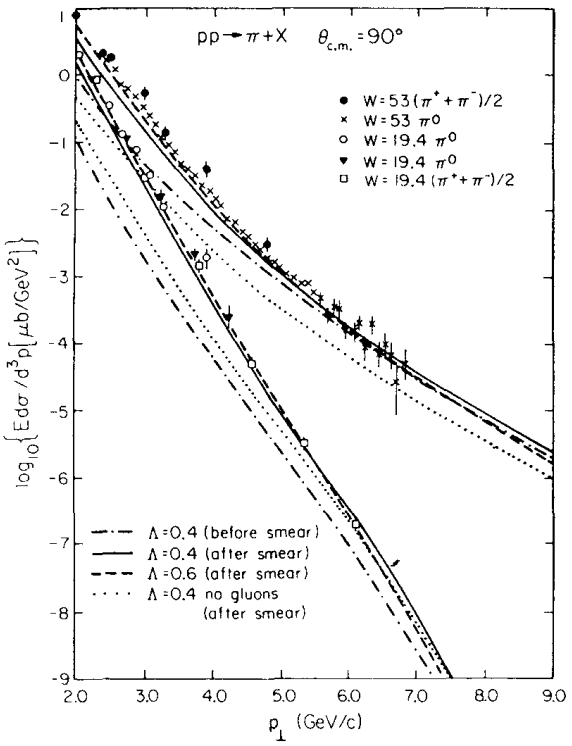


FIG. 10. Comparison of a QCD model (normalized absolutely) with data on large-$p_\perp$ pion production in proton-proton collisions as $W=\sqrt{s}$ =19.4 and 53 GeV/c with $\theta_{c.m.}=90°$ (open squares: Ref. 73, solid dots: Ref. 74, crosses: Ref. 75, solid triangles: Ref. 76, open circles: Ref. 77). The dot-dashed and solid curves are the results before and after smearing, respectively, using $\Lambda=0.4$ GeV/c and $\langle k_\perp\rangle_{h\to q}=848$ MeV and the dashed curves for $\Lambda=0.6$ GeV/c (after smearing). The contribution arising from quark-quark, quark-antiquark, and antiquark-antiquark scattering (i.e., no gluons) is shown by the dotted curves (after smearing).

Fig. 10. The agreement is quite remarkable. It is almost as good as the black-box model (Fig. 13 of FF1) where we chose the behavior of $d\hat{\sigma}/d\hat{t}$ and the normalization to fit the data. The results before smearing are also shown (dot-dashed curves). Smearing has little effect for $p_\perp \gtrsim 4.0$ GeV/c at $W=53$ GeV but has a sizable effect (even at $p_\perp=6.0$ GeV/c) at $W=19.4$ GeV due to the steepness of the cross section at this low energy. The contributions to the total invariant cross section from quark-quark elastic scattering (plus $q\bar{q}\to q\bar{q}$ and $\bar{q}\bar{q}\to\bar{q}\bar{q}$) are shown in Fig. 10 (dotted curve). Gluons make important contributions to the cross section at small $x_\perp$ ($x_\perp \lesssim 0.4$).

We cannot at this time say whether the slight disagreement in the normalization of the theory seen in Figs. 8, 9, and 10 at low $x_\perp$ (about a factor of 2 at $p_\perp=2$ GeV/c and $W=53$ GeV) is significant. At these low values of $x_\perp$ and $p_\perp$, the theory cannot at present be calculated precisely as the results depend sensitively on the unknown gluon distributions, the shape of the transverse-momentum distributions of the quark within the hadrons, the nature of the low-$\hat{s}$ and -$\hat{t}$ cutoff, our choice for Q^2, and possibly higher-order corrections [such as the $\pm C$ in Eq. (3.1)]. It may be that all of the invariant cross section down to $p_\perp$'s as low as 1.5 or 2.0 GeV/c is due to the scattering of quarks and gluons as described by QCD. On the other hand, it may be that other nonleading constituent subprocesses such as the ones estimated by Blankenbecler, Brodsky, and Gunion[56] make some contributions in the range of $1.5 \lesssim p_\perp \lesssim 4.0$ GeV/c with QCD dominating at the higher $p_\perp$'s.

2. *Angular dependence*

In FF1 we used the data shown in Fig. 11 to deduce that $d\hat{\sigma}/d\hat{t} \propto 1/\hat{s}\hat{t}^3$ was the preferred form for the angular dependence of quark-quark scattering. We must now see if the QCD predictions are

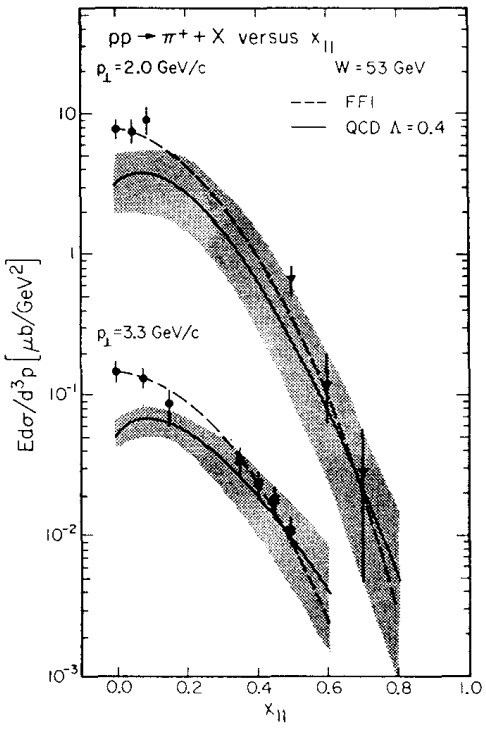


FIG. 11. Data on the $x_\parallel$ dependence of the invariant cross section for $pp\to\pi^+ + X$ at $W=53$ GeV and $p_\perp=2.0$ and 3.3 GeV/c (solid dots: Ref. 74, solid triangles: Ref. 78, solid Squares: Ref. 79) compared with the results of a QCD model with $\Lambda=0.4$ GeV/c (solid curves). At these low-$p_\perp$ values, the predictions are sensitive to the low $\hat{s}$ and $\hat{t}$ cutoff of $d\hat{\sigma}/d\hat{t}$. At any fixed $p_\perp$, the uncertainty due to the cutoff procedure (illustrated by the shaded areas) is greater at large $x_\parallel$. Also shown (dashed curves) are the results from the quark-quark black-box model of FF1 which was adjusted to fit these data.

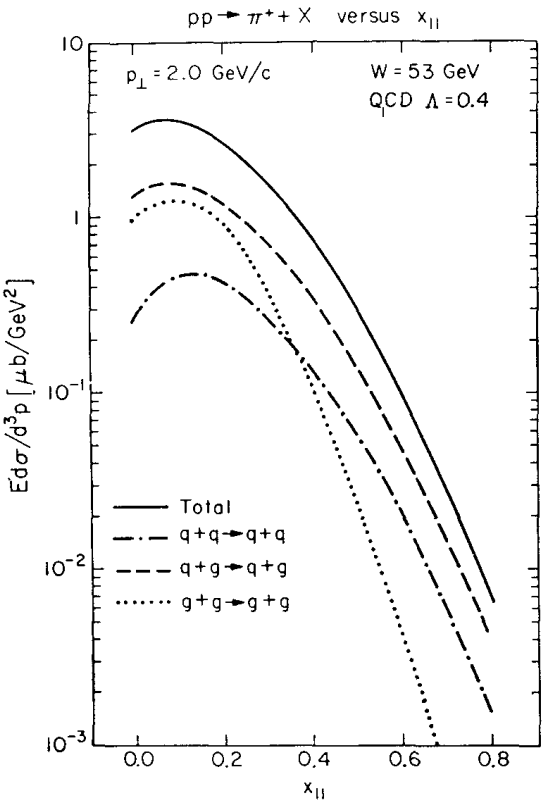


FIG. 12. Contributions to the total $Ed\sigma/d^3p$ for $pp \to \pi^+ + X$ arising from the various QCD subprocesses, $q+q \to q+q$, $q+g \to q+g$, and $g+g \to g+g$ at $W=53$ GeV, $p_\perp=2.0$ GeV/c versus $x_\parallel$.

consistent with the same data. Figure 11 shows the QCD results with $\Lambda = 0.4$ GeV/c (after smearing) for the $x_\parallel$ dependence of $Ed\sigma/d^3p$ for $pp \to \pi^+ + X$ at $p_\perp = 2$ and 3.3 GeV/c and $W = 53$ GeV. As noted earlier, the predictions are a bit low at this low $p_\perp$; however, the QCD $x_\parallel$ dependence is not in gross disagreement with the data. Unfortunately, comparison with these low-$p_\perp$ data is not very significant. First, with the large value of $\langle k_\perp \rangle_{k \to q}$ we are now using, the results for these $p_\perp$ values are sensitive to the manner in which we remove the singularity in $d\hat{\sigma}/d\hat{t}$ at $\hat{t}$ (or $\hat{u}$) equal zero. This is particularly true at large $x_\parallel$ where the constituent scattering occurs at small $\hat{t}$ (or $\hat{u}$) values. We have tried to indicate this uncertainty by the shaded region in Fig. 10. Secondly, since this region is so sensitive to smearing, the results depend on our assumption that $\langle k_\perp \rangle_{k \to q}$ is independent of the x_q of the quark q. Modifying this assumption could change the resulting $x_\parallel$ dependence.

Furthermore, in the QCD approach, the $x_\parallel$ dependence of the invariant cross section does not directly reflect the angular dependence of $d\hat{\sigma}/d\hat{t}$. Figure 12 shows that the various QCD subprocesses have differing $x_\parallel$ dependences so that the result depends not only on the various $d\hat{\sigma}/d\hat{t}$ but also on the amount of each term. For example, the gluon-gluon scattering contributions fall off fast with increasing $x_\parallel$. We could cause the predicted $Ed\sigma/d^3p$ to fall off more rapidly with $x_\parallel$ by increasing the amount of gluon-gluon scattering [by changing the relatively unknown functions $G_{p \to g}(x, Q^2)$ and $D_g^{\pi^+}(z, Q^2)$].

3. Particle ratios

As mentioned earlier, and as shown in Fig. 10, gluons make an important contribution to the single-particle invariant cross section at low $x_\perp$. However, since the gluon fragmentation function has been chosen to be considerably smaller at large z than the quark fragmentation function, an experiment demanding a large-$p_\perp$ meson trigger is "biased" in favor of the toward-side constituent being a quark rather than a gluon. Table III gives the fraction of the single-particle cross section arising from the various combinations of toward and away constituents. At $W = 53$ and $p_\perp \geq 4.0$ GeV/c, 45% of the π^0 cross section arises from a toward-side quark having scattered off a recoiling gluon while 27% arises from the quark recoiling off another quark (a total of 72%). Tables III and IV show that for $x_\perp \gtrsim 0.3$, the toward constituent for single-particle triggers is almost always a quark. This quark, however, scatters off both quarks and gluons in the other proton so that the recoiling constituent is quite often a gluon. At $W = 53$ GeV, a π^0 trigger with $p_\perp = 4.0$ GeV/c, the toward constituent is a quark (or antiquark) 72% of the time while the away constituent is a gluon 62% of the time.

This "bias" for quarks rather than gluons in single-particle triggers means that the predictions for ratios of different kinds of particles will not be very different from those of the quark-quark scattering black-box approach. As shown in Fig. 13, $pp \to (\pi^+/\pi^-) + X$ ratio predictions of QCD are smaller due to the contamination from the gluon decays (gluons fragment into equal numbers of positives and negatives). They are in equally good agreement with data. The QCD particle ratios do not "scale" as did the FF1 results (i.e., they are a function of $x_\perp$ and W at fixed $\theta_{\text{c.m.}}$). The curve displayed in Fig. 13(a) is calculated at $W = 19.4$ GeV and increases slightly as W increases (by about 20% in going from $W = 19.4$ to 53 GeV/c).

B. The jet cross section

The "bias" in favor of toward-side quarks does not occur when one triggers on jets rather than

TABLE III. Fraction of the total $pp \to \pi^0 + X$ 90° invariant cross section arising from various combinations of toward-side (trigger) and away-side partons. Results are shown for the cases where the π^0 trigger came from a quark or gluon and the recoiling constituent (away-side) is a quark or gluon or either (sum). The notation is "toward ⟷ away" and "quark" refers to $q + \bar{q}$.

W (GeV)	$P_\perp$ (GeV/c)	$x_\perp$	quark ⟷ quark	Toward-side quark: quark ⟷ gluon	quark ⟷ anything	gluon ⟷ quark	Toward-side gluon: gluon ⟷ gluon	gluon ⟷ anything
53	2.0	0.08	0.16	0.41	0.57	0.15	0.28	0.43
53	4.0	0.15	0.27	0.45	0.72	0.11	0.17	0.28
53	7.0	0.26	0.39	0.48	0.87	0.07	0.06	0.13
53	9.0	0.34	0.45	0.48	0.93	0.04	0.03	0.07
19.4	1.94	0.20	0.18	0.36	0.54	0.15	0.31	0.46
19.4	3.0	0.31	0.27	0.54	0.81	0.08	0.11	0.19
19.4	4.0	0.41	0.37	0.52	0.89	0.06	0.05	0.11
19.4	6.0	0.62	0.55	0.42	0.97	0.02	0.01	0.03
500	10	0.04	0.24	0.32	0.56	0.15	0.29	0.44
500	30	0.12	0.34	0.42	0.76	0.13	0.11	0.24

on single particles and, as can be seen in Table IV, gluons make up a sizable fraction of the total jet cross section. With our guesses for the gluon distributions, gluons are responsible for 73% of the jet triggers at $p_\perp = 4$ GeV/c, $W = 53$ GeV, and $\theta_{\text{c.m.}} = 90°$. Even at higher-$x_\perp$ values such as $p_\perp = 6.0$ GeV/c, $W = 19.4$ GeV, $\theta_{\text{c.m.}} = 90°$, gluons still make up 45% of the jets. Because of the presence of gluon jets and because the quark fragmentation functions $D_q^h(z, Q^2)$ are smaller at high z due to scale-breaking effects, the jet to single-π^0 ratio is now predicted to be larger than it was for the quark-quark black-box approach. This is seen in Fig. 14 where we plot the invariant cross section for $pp \to \text{Jet} + X$ divided by $pp \to \pi^0 + X$ at $\theta_{\text{c.m.}} = 90°$ versus $x_\perp$. In the QCD approach, this ratio no longer "scales." It is a function not only of $x_\perp$ (at fixed $\theta_{\text{c.m.}}$), but also of W. We show results for $W = 19.4$, 53, and 500 GeV.

As noted in Ref. 57, the quark scattering model in FF1 and FFF agreed quite well with the jet cross section observed experimentally at $W = 19.4$ GeV and $3 \leq p_\perp \leq 6.0$ GeV/c. We might now be concerned that the QCD results for the jet cross section are larger than FF1 by a factor of about 5 in this region. We have, however, previously been somewhat naive when comparing theory with experiment. What we show in Fig. 14 is the cross section for producing a quark (or gluon) with a given momentum (divided by the π^0 cross section at the same momentum). However, as we noticed in Ref. 27, quarks of a given momentum (equal to their energy) cannot produce jets with the momentum of all particles equal to the energy of all particles. Our jet model in FF2 gives $E_{\text{tot}} - p_{z_{\text{tot}}} \approx 1.2$ GeV for quark jets. Since the cross section for producing jets falls so steeply, the cross section for producing a jet with a given $p_{z_{\text{tot}}}$ is considerably smaller than that for producing one with a given E_{tot}. As explained in Ref. 41, it is the former that is more closely connected to what is measured experimentally. At $W = 19.4$ GeV/c, we estimate that the cross section to produce a jet where $p_{z_{\text{tot}}} = 5$ GeV/c at 90° is about 10 times smaller than the cross section to produce a jet whose $E_{\text{tot}} = 5$ GeV/c. If we correct for this effect, the new QCD prediction at $W = 19.4$ GeV is within a factor of 2 of the old (incorrectly interpreted) FF1 results that appeared to agree so well with the data.

The difference between E_{tot} and $p_{z_{\text{tot}}}$ of a jet arises, of course, from low-momentum particles that have energy due to their mass (or $k_\perp$) but have little momentum p_z. This is tangled with the experimental uncertainty in all hadron-jet experiments concerning low-$p_\perp$ particles. One cannot be sure that one is not losing the low-$p_\perp$

TABLE IV. Fraction of the $pp \to \pi^0 + X$ and $pp \to \text{jet} + X$ cross section at $\theta_{\text{c.m.}} = 90°$ arising from the case where the toward-side (or trigger) constituent is a u, d, or antiquarks $\bar{Q} = \bar{u} + \bar{d} + \bar{s}$ or a gluon. Also shown is the fraction arising from the case where the recoiling or away side constituent is a u, d, or antiquarks $\bar{Q}$ or a gluon.

			Toward side				Away side			
W (GeV)	$P_\perp$ (GeV/c)	Trigger	u	d	$\bar{Q}$	Gluon	u	d	$\bar{Q}$	Gluon
53	2	π^0	0.20	0.20	0.13	0.43	0.14	0.09	0.06	0.70
53	2	jet	0.12	0.10	0.14	0.64	0.18	0.09	0.08	0.63
53	4	π^0	0.36	0.25	0.09	0.27	0.23	0.09	0.05	0.62
53	4	jet	0.12	0.09	0.05	0.73	0.20	0.08	0.02	0.69
53	7	π^0	0.55	0.26	0.05	0.13	0.28	0.14	0.05	0.53
19.4	3.0	π^0	0.46	0.29	0.06	0.19	0.20	0.11	0.03	0.66
19.4	3.0	jet	0.13	0.08	0.04	0.74	0.21	0.08	0.02	0.68
19.4	6.0	π^0	0.73	0.24	0	0.03	0.42	0.15	0	0.43
19.4	6.0	jet	0.37	0.17	0	0.45	0.36	0.14	0	0.50
500	10	π^0	0.18	0.17	0.18	0.44	0.21	0.08	0.09	0.60
500	10	jet	0.05	0.05	0.06	0.81	0.13	0.11	0.07	0.66
500	30	π^0	0.37	0.23	0.14	0.24	0.29	0.12	0.05	0.53
500	30	jet	0.11	0.08	0.10	0.69	0.21	0.11	0.06	0.60

jet particles that are not well collimated or gaining low-$p_\perp$ background from the beam and target jets in Fig. 1. Only by doing a very careful analysis, including the precise acceptances of a given experiment, can one distinguish between the results of FF1 and the new QCD approach in spite of their rather large differences. One might hope someday to distinguish experimentally between gluon and quark jets. The gluon jets are assumed to have a higher multiplicity of particles each with lower momentum on the average. In addition, unlike the quark jets discussed in Ref. 27, gluon jets will carry on the average no net charge (or strangeness, etc.).

C. The toward-side correlations

The jet physics, discussed in the previous subsection, directly tests that particles at high $p_\perp$ are not produced singly but, rather, are members of a cluster. This aspect of constituent models can also be tested in single-particle triggers by observing the accompanying particles produced near the trigger (in phase space). Experimentally, one observes an enhancement of particles with high $p_\perp$ accompanying the high-$p_\perp$ trigger which was predicted correctly from the quark scattering model[2,41] together with an assumption about the double fragmentation function, $D_q^{h_1,h_2}(z_1, z_2)$.[27] In the QCD case, the trigger hadrons usually come from quarks rather than gluons and, furthermore, the mean value of z_c shown in Table II are similar to those in FFF.[58] Thus the new predictions, as shown in Fig. 15, are, in fact, very similar to the quark scattering results. However, the QCD results have an additional uncertainty due to our lack of knowledge of the gluon double fragmentation functions, $D_g^{h_1,h_2}(z_1, z_2, Q^2)$. In Fig. 15 where 18% of the triggers come from gluons, we have assumed that the accompanying particles were just given by the fragmentation of a gluon with momentum $(1 - z_c)$ times that of trigger gluon. This has the feature of giving the correct multiplicity for the produced hadrons; however, it is surely not exact. For example, for a π^+ trigger, one would expect to find more high-$p_\perp$ π^-'s than π^+'s in the accompanying particles. In Ref. 22, an upper estimate for the effect of the gluon decays was deduced by assuming that the accompanying hadrons came from a $\bar{u}$ (or d) jet carrying all of the remaining momentum. Comparing with the CCHK data, this upper estimate and the lower estimate gotten by dropping gluon term completely roughly bracket the data.

In summary, we find that the QCD calculations have somewhat greater uncertainty than those in FFF coming from the gluon contribution. However, both models are in good agreement with the same-side correlation data.

D. Away-side correlations

1. *Away-side multiplicity $n(z_p)$*

An important consequence of the QCD approach is that the number of away-side hadrons with large-$p_\perp$ $(-p_x)$ is predicted to be considerably

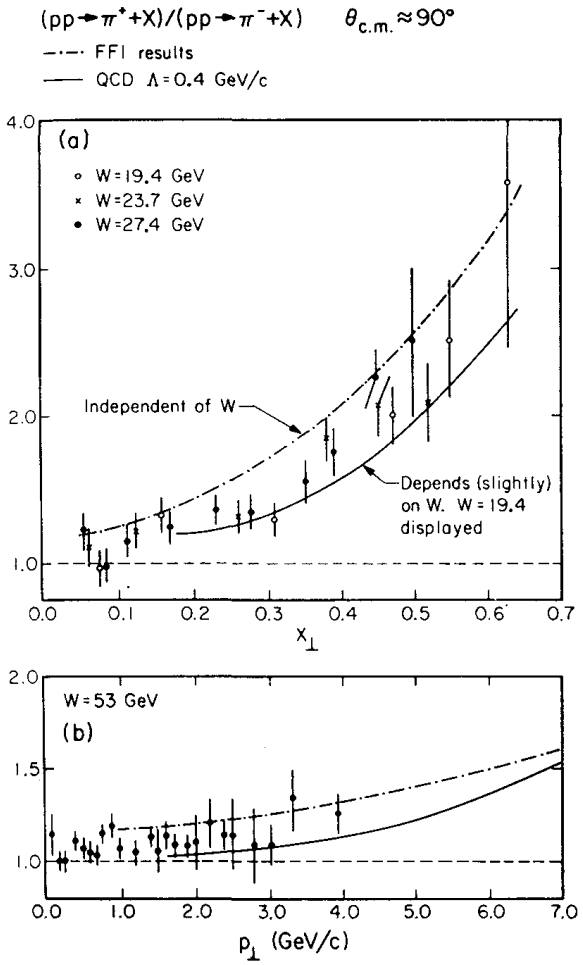


FIG. 13. (a) Comparison of the data (Refs. 73 and 80), of the π^+/π^- ratio in pp collisions at $\theta_{c.m.}=90°$ versus $x_\perp$ with the quark-quark scattering model of FF1 (which is independent of W at fixed $x_\perp$) and the QCD results using $\Lambda=0.4$ GeV/c. The QCD results are plotted for $W=19.4$ GeV and are not precisely independent of W. The π^+/π^- ratio increases at fixed $x_\perp$ and $\theta_{c.m.}$ as W increases (by about 20% in going from 19.4 to 53 GeV). (b) Comparison of the data (Ref. 74) at $W=53$ GeV on $pp\to(\pi^+/\pi^-)+X$ at $\theta_{c.m.}=90°$ versus $p_\perp$ with the QCD results.

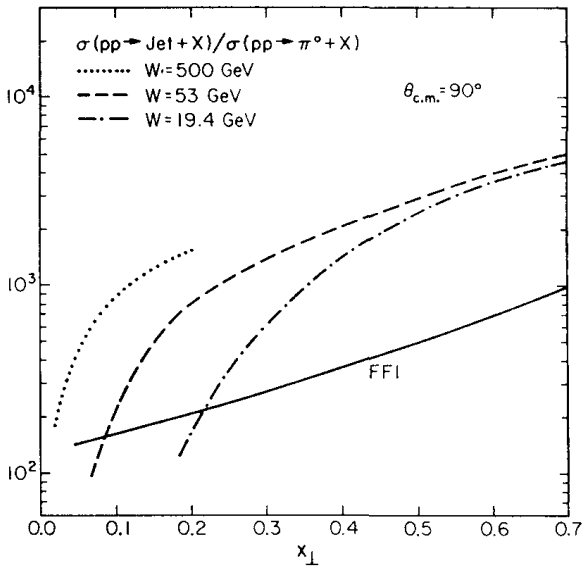


FIG. 14. Prediction of the jet-to-single-π^0 ratio $\theta_{c.m.}=90°$ versus $x_\perp$ for $W=500$, 53, and 19.4 GeV from the QCD approach using $\Lambda=0.4$ GeV/c. The jet cross section is defined as the cross section for producing a parton (quark+antiquark+gluon) with the given $x_\perp$. Also shown is the prediction from the quark scattering model of FF1 which is independent of W at fixed $x_\perp$ and $\theta_{c.m.}$.

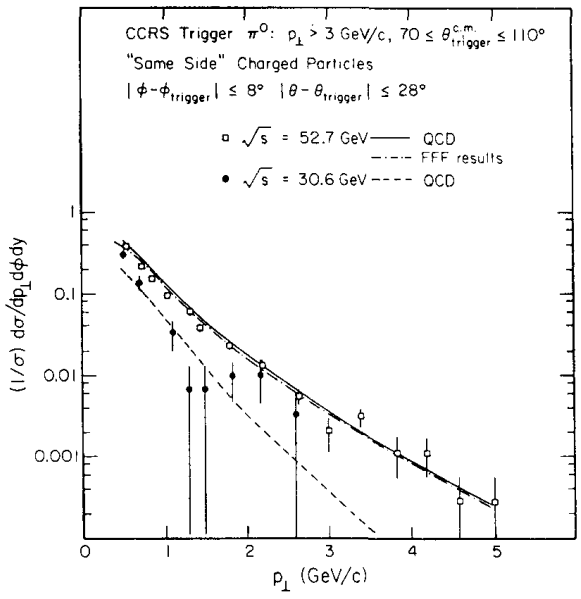


FIG. 15. Toward-side correlation measurements from CCRS collaboration (Ref. 75) together with the predictions of the QCD approach with $\Lambda=0.4$ GeV/c and the results of the quark-quark black-box model of FFF. Possible background contributions from the fragmentation of the beam and target are *not* included.

smaller than in the quark-quark scattering approach. Figures 16, 17, and 18 show that the number of away hadrons carrying a certain fraction z_p of the trigger momentum is predicted to be 3 to 4 times less than the FFF results, and now agrees quite well with experiment. This reduction in the away-side multiplicity function, $n(z_p)$ is due to three factors. First, we have increased $k_{h\to q}$ from 500 to 848 MeV. This results in the large Q_x values shown in Table I (compare with Table V in FFF) and thus to a reduction of $n(z_p)$. Second, the fragmentation functions $D_q^h(z, Q^2)$ de-

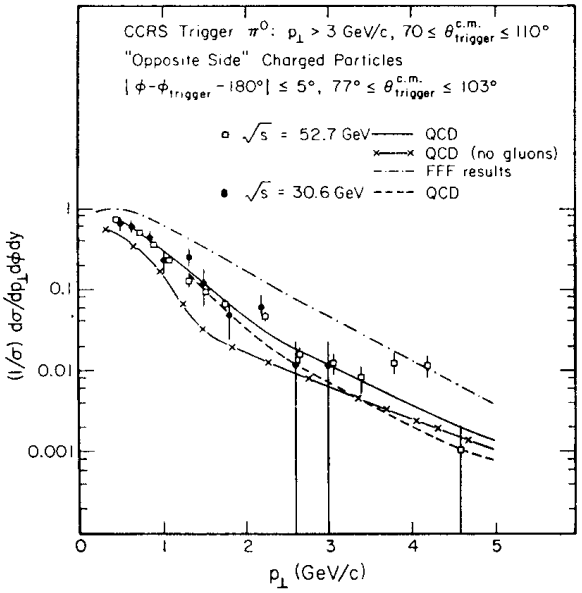


FIG. 16. Opposite or away-side correlation measurements from the CCRS collaboration (Ref. 75) together with the prediction of the QCD approach with $\Lambda = 0.4$ GeV/c and the results of the quark-quark black-box model of FFF. Possible background contribution from the beam and target jets have *not* been included.

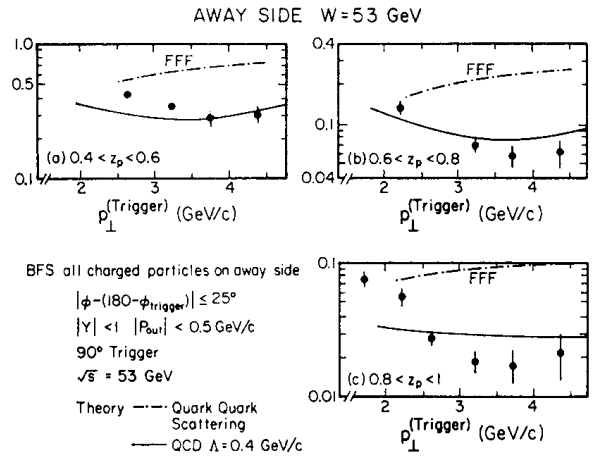


FIG. 17. The dependence on the trigger $p_\perp$ of the away-side hadron multiplicity $n(z_p) = (1/\sigma) d\sigma/dz_p$, where $z_p = -p_x$ (away)$/p_\perp$ (trig) from the British-French-Scandinavian collaboration (Ref. 43) on $pp \to h_1^\pm + h_2^\pm + X$ at $W = 53$ GeV, $\theta_1 = 90°$ and with an away-side acceptance of 25° in ϕ and $|Y_2| < 1$, $|P_{out}| < 0.5$ GeV/c. The predictions from the QCD approach with $\Lambda = 0.4$ GeV/c (solid curves) and the results of the quark-quark black-box model of FFF (dash-dot curves) are shown. Background contributions from the fragmentation of the beam and target (see Fig. 1 and Fig. 6) which might be important for low-$p_\perp$ triggers have *not* been included in either the QCD or FFF predictions.

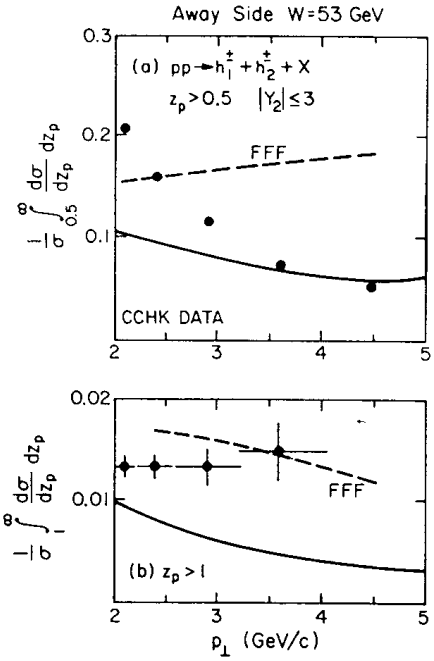


FIG. 18. The dependence on the trigger $p_\perp$ of the number of away-side hadrons per trigger with $z_p \ge 0.5$ (a) and $z_p \ge 1.0$ (b) from the CCHK collaboration (Ref. 15) on $pp \to h_1^\pm + h_2^\pm + X$ at $W = 53$ GeV, and θ_1 averaged over 45° and 20° with an away-side acceptance of 40° in ϕ and $|Y_2| \le 3$. The predictions from the QCD approach with $\Lambda = 0.4$ GeV/c (solid curves) and the results from the quark-quark black-box model of FFF (dashed curves) are shown. Background contributions from the beam and target jets (see Fig. 1 and Fig. 6) which might be important for low-$p_\perp$ triggers have *not* been included in either the QCD or FFF predictions.

crease as Q^2 increases (see Fig. 5) and are smaller at high z than the FFF values (which now correspond to $Q^2 = Q_0^2 = 4$ GeV^2). Finally, in the QCD approach, the away-side constituent is quite often a gluon (see Tables III and IV) which produces on the average fewer hadrons at large z_p than do quarks (see Fig. 19). However, as Table V and Fig. 16 show, the number of away hadrons with $z_p \ge 0.5$ arising from gluon jets is still about half the total. (The fraction decreases as $x_\perp$ increases.) This means that the away-side multiplicity $n(z_p)$ is sensitive to the essentially unknown gluon distributions $G_{p\to g}(x, Q^2)$ and $D_g^h(z, Q^2)$.

For both the QCD approach and the quark scattering model, the away-side multiplicity function, $n(z_p)$, is roughly independent of the trigger momentum over the range $2.0 \le p_\perp(\text{trig}) \le 6.0$ GeV/c at $W = 53$ GeV. This means that the rise in the data (Figs. 17 and 18) at small $p_\perp$(trig) must be ascribed to "background" from the beam and target jets (see Fig. 1). We do not know how to cal-

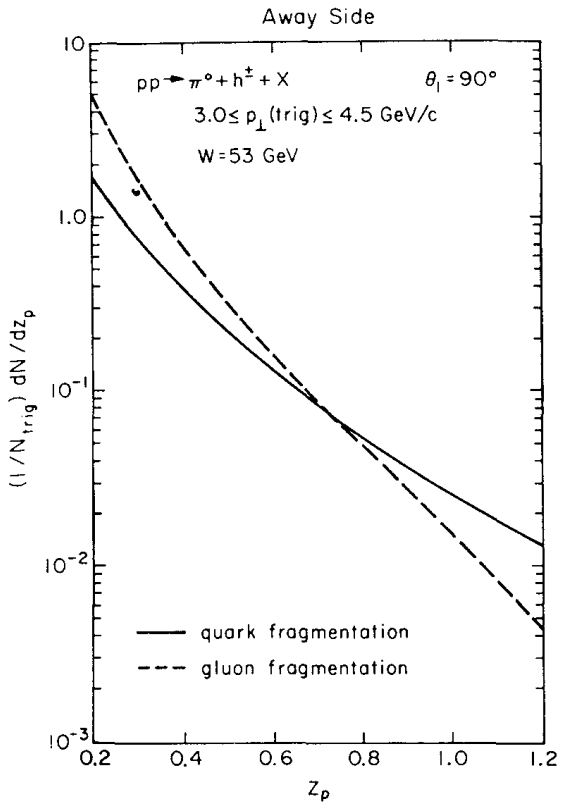


FIG. 19. The number of away-side charged hadrons per trigger, $n(z_p)$, arising from the case that the away-side constituent is a quark or antiquark (solid curve) or a gluon (dashed curve) from the QCD approach with $\Lambda = 0.4$ GeV/c. The results are calculated for $pp \to \pi^0 + h^\pm - X$ at $W = 53$ GeV, $\theta_1 = 90°$ with $3.0 \le p_\perp$ (trig) ≤ 4.5 GeV/c.

culate this properly at present, but estimates we have made indicate that this is indeed possible.[60] The rise at small $p_\perp$(trig) in the CCHK experiment (Fig. 18) is larger than that seen in experiment R-413 (Fig. 17) because the former has an away-side rapidity cut of $|Y| \le 3$ while the latter has $|Y| \le 1$. The CCHK experiment thus receives a larger background contamination at low $p_\perp$(trig) particularly from the Type I background shown in Fig. 6.

2. *Away-side particle ratios*

Another effect of the presence of gluons is that the away-side positive to negative particle ratios at the ISR (low $x_\perp$) are predicted to be considerably different than in FFF. Figure 20 shows that the QCD approach yields almost equal numbers of positives and negatives for $p_\perp$(away) > 1.5 GeV/c at $W = 53$ GeV and $3.0 \le p_\perp$(trig) ≤ 4.0 GeV/c in agreement with the recent ISR data.[43] If the away-side constituent is always a quark or antiquark as in FFF, then this ratio is predicted to be about 1.5 in gross disagreement with the experiment.[38] However, both the FFF model and the QCD approach predict little dependence of the away-side particle ratios on the type of trigger species (i.e., π^+, π^-, K^+, K^-). This is because the scattering forces do not involve flavor exchange (they are due to gluon exchange). Neither the QCD approach nor the FFF model can explain the apparently large increase in the away-side positive to negative ratio when triggering on K^- observed by R-413 (Fig. 20). The discrepancy can be seen more clearly in Fig. 21 where we compare the predictions for the away-side rapidity spectrum of positives and negatives for a π^- and K^- trigger with the preliminary R-413 data.[43]

This question of the flavor dependence of the constituent subprocesses is an important one. In models such as the constituent-interchange model (CIM),[56,61] the scattering forces arise from the exchange of quarks which carry flavor.[62] In these models, drastic changes can occur in the away-side particle ratios as one changes trigger species.[63] Figure 22 shows data from the Fermilab experiment E-494 (Ref. 64) on the away-side multiplicity of π^+, π^-, K^+, and K^- with $z_p \ge 0.5$ for a trigger meson of type π^+, π^-, K^+, K^- at W

TABLE V. Total away-side multiplicity (Ref. 59), $N(z_p \ge 0.5)$, per trigger for charged hadrons in the processes $pp \to \pi^0 + h^\pm + X$ at $\theta_1 = 90°$ predicted from the QCD approach with $\Lambda = 0.4$ GeV/c. Also shown are the individual contributions to the multiplicity for $z_p \ge 0.5$ from gluon and quark fragmentation. [The function $N(z_p^0, P_{\perp 1}, \theta_1)$ is defined by Eqs. (6.1) and (6.2) in FFF.]

W (GeV)	$P_\perp$ (GeV/c)	$x_\perp$	Quark fragmentation	Gluon fragmentation	Total $z_p \ge 0.5$
53	2.0	0.08	0.047	0.083	0.130
53	3.4	0.13	0.052	0.050	0.102
53	4.5	0.17	0.058	0.047	0.105
53	5.3	0.20	0.063	0.038	0.101
53	9.3	0.35	0.060	0.022	0.082
53	13.2	0.50	0.058	0.014	0.072

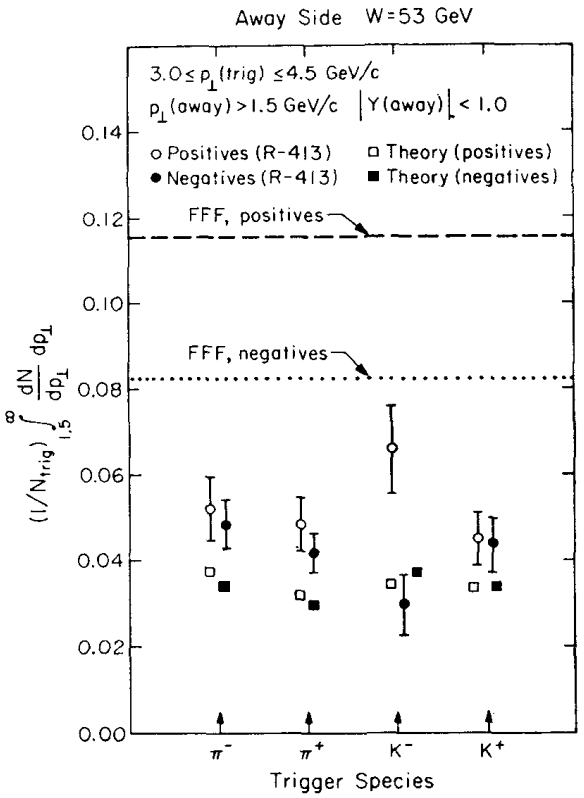


FIG. 20. The number of away-side positive and negative hadrons with $p_\perp$ (away) >1.5 GeV/c per trigger with $3.0 \leq p_\perp$ (trig) ≤ 4.5 GeV/c from the BFS collaboration (R-413) (Ref. 43) on $pp \to h_1 + h^\pm + X$ where $W = 53$ GeV and $\theta_1 = 90°$ and $|Y_2| < 1.0$. The results for π^-, π^+, K^-, and K^+ triggers are shown and compared to the predictions of the quark-quark black-box model of FFF and the QCD approach with $\Lambda = 0.4$ GeV/c (open and solid squares). Background contributions from the beam and target jets (see Figs. 1 and 6) have *not* been included in either the QCD or FFF predictions.

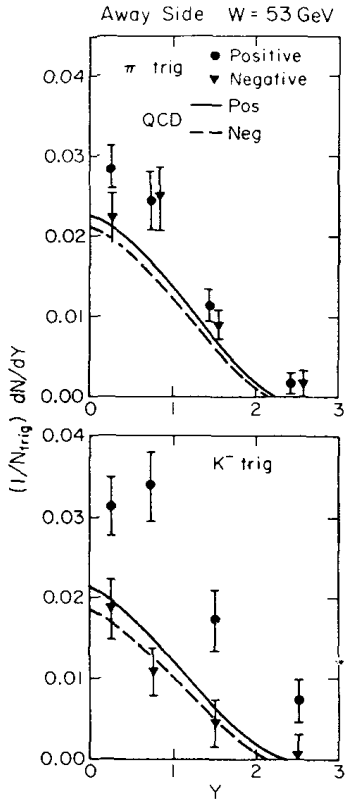


FIG. 21. The away-side rapidity distributions, $(1/N_{trig})\, dN/dY$, of positive and negative hadrons with $p_\perp$ (away) >1.5 GeV/c for 90° π^- and K^- triggers in the range $3.0 \leq p_\perp$(trig) ≤ 4.5 GeV/c from the BFS collaboration (R-413) (Ref. 43). Predictions of the QCD approach with $\Lambda = 0.4$ GeV/c are shown where background contributions from the beam and target jets (see Figs. 1 and 6) have not been included.

$= 27.4$ GeV and $3.0 \leq p_\perp(\text{trig}) \leq 5.0$ GeV/c compared to the QCD predictions. The agreement is quite good. The away-side ratios are roughly independent of the trigger species and given approximately by the single-particle ratios (shown by the wiggly arrows along the side) which is just as expected for a flavorless-exchange constituent subprocess. There is a slight disagreement for the K^- trigger but the data *do not* show the large positive plus negative sum seen by R-413. In fact, the away-side number of K^+ mesons with a K^- trigger is correctly predicted. The data in Fig. 22 from E-494 are taken off a beryllium target and there are A-dependence corrections[64] (we have made no A-dependence correction to our theoretical predictions) that make direct comparison a bit dangerous. Because of this and because of the apparent disagreement between R-413 and E-494 for K^- triggers the question as to whether or not there is any evidence for flavor exchange in the constituent subprocess is unsettled.

3. P_{out}

Due to our use of $\langle k_\perp \rangle_{h \to q} = 848$ MeV and $\langle k_\perp \rangle_{q \to h} = 439$ MeV, the mean values of P_{out} are predicted to be considerably larger than the results of FFF ($\langle k_\perp \rangle_{h \to q} = 500$ MeV, $\langle k_\perp \rangle_{q \to h} = 330$ MeV). In Figs. 23 and 24, we compare both the new QCD results and the FFF results with the mean values of P_{out} obtained in the CCHK experiment.[15] The value of $\langle k_\perp \rangle_{h \to q} = 848$ MeV, obtained from the fit to the data on $pp \to \mu^+ \mu^- + X$ shown in Fig. 7, results in $\langle P_{out} \rangle$ values that agree better with the hadron experiments, although they are still a bit small. (Some of the discrepancies may be due to contributions from the beam and target jets omitted in our analysis.)

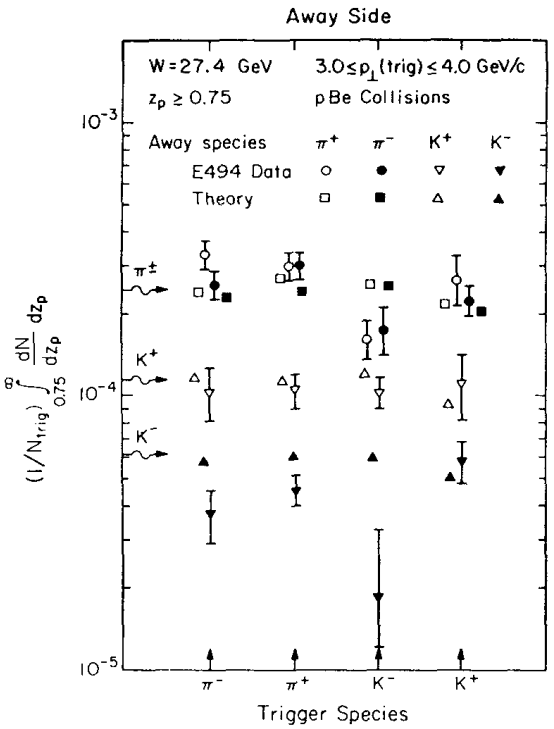


FIG. 22. The number of away-side mesons (of type π^+, π^-, K^+, K^-) with $z_p \geq 0.75$ per trigger (of type π^+, π^-, K^+, K^-) from the Fermilab experiment E494 (Ref. 64). The data are taken at $W = 27.4$ GeV with $3.0 \leq p_\perp(\text{trig}) \leq 4.0$ for proton-beryllium collisions and are compared with the prediction of the QCD approach (for proton-proton collisions) with $\Lambda = 0.4$ GeV/c. No A-dependence corrections have been made to the theory or the data.

4. Experimental tests for effects due to $\langle k_\perp \rangle_{h \to q}$

As seen in Figs. 2 and 8, the basic constituent subprocess of QCD (before smearing) behaves roughly like $1/p_\perp^6$ at fixed $x_\perp$ for $2 \leq p_\perp \leq 10$ GeV/c. The experimentally observed $1/p_\perp^8$ behavior is obtained by including the effects of smearing ($\langle k_\perp \rangle_{h \to q} \neq 0$) which raise the small-$p_\perp$ prediction while leaving the large-$p_\perp$ region essentially unchanged. This increase at small $p_\perp$, due to the "trigger bias" effect, can be partially removed by triggering on events with equally large $p_\perp$'s on the toward and away-side (i.e., $z_p \approx 1$).[65,66] Thus, in general, we expect the $p_\perp$ dependence of the two-particle back-to-back cross section to differ (in the region where smearing is an important effect) from that of the single-particle cross section. This is seen in Fig. 25 where we plot the two-particle back-to-back cross section dJ/dz_p at $z_p = 1$ (times $p_\perp^8$) versus $p_\perp$ at $x_\perp = 0.35$. It behaves roughly like $1/p_\perp^6$ over the range $4 \leq p_\perp \leq 6.0$ GeV/c whereas the single-particle cross-section results, when multiplied by $p_\perp^8$,

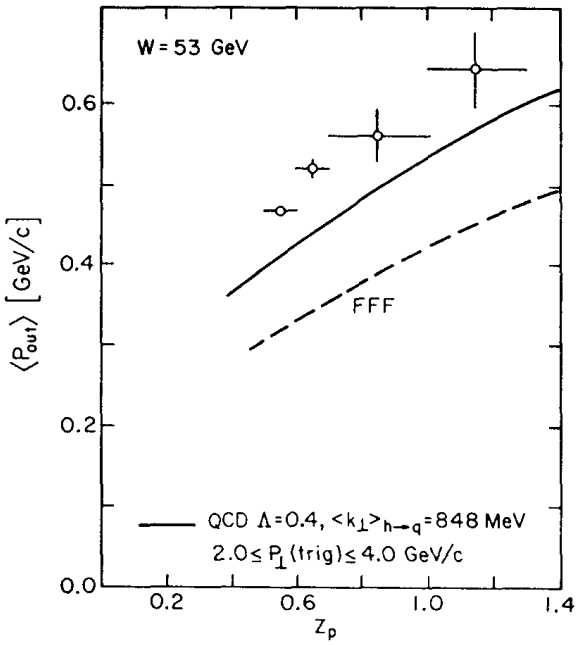


FIG. 23. The dependence on z_p of the mean value of the $|P_{out}|$ of away-side charged hadrons at $W = 53$ GeV and $2.0 \leq p_\perp(\text{trig}) \leq 4.0$ GeV/c with θ_1 averaged over 45° and 20° from the CCHK collaboration (Ref. 15) on $pp \to h_1^\pm + h_2^\pm + X$. The predictions from the QCD approach at $\theta_1 = 45°$ with $\Lambda = 0.4$ GeV/c and $\langle k_\perp \rangle_{h \to q} = 848$ MeV (solid curve) and the results of FFF (dashed curve) curve are shown.

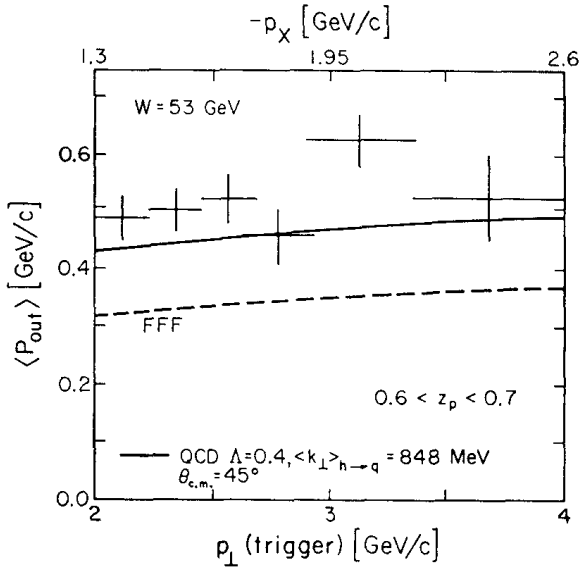


FIG. 24. The dependence on the trigger $p_\perp$ of the mean value of the $|P_{out}|$ of away-side charged hadrons with $0.6 \leq z_p \leq 0.7$ at $W = 53$ GeV and θ_1 averaged over 45° and 20° from the CCHK collaboration (Ref. 15) on $pp \to h_1^\pm + h_2^\pm = X$. The predictions for $\theta_1 = 45°$ from the QCD approach with $\Lambda = 0.4$ GeV/c and $\langle k_\perp \rangle_{h \to q} = 848$ MeV, $\langle k_\perp \rangle_{q \to h} = 439$ MeV (solid curve) and the FFF results (dashed curve) are shown.

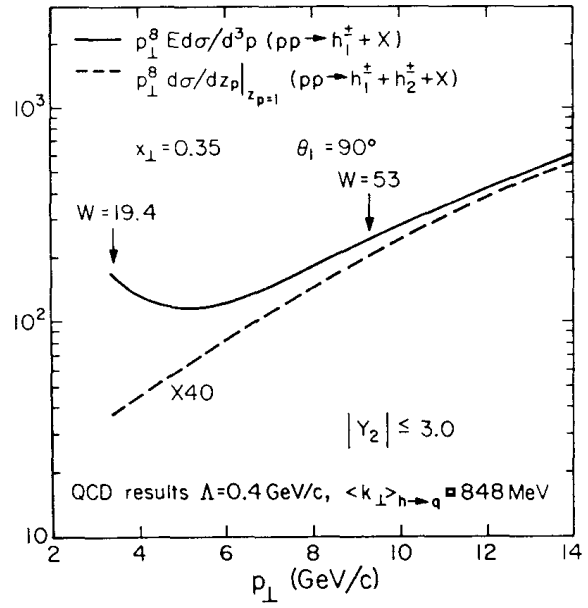


FIG. 25. Comparison of the behavior of $p_\perp^8$ times the single-charged-particle cross section $Ed\sigma/d^3p$ $(pp \to h_1^\pm + X)$, and $p_\perp^8$ times the two-particle back-to-back cross section $d\sigma/dz_p|_{z_p=1}$ $(pp \to h_1^\pm + h_2^\pm + X)$ at *fixed* $x_\perp = 0.35$ (times 40). The QCD predictions are calculated at $\theta_1 = 90°$ with $\Lambda = 0.4$ GeV/c and $\langle k_\perp \rangle_{h-q} = 848$ MeV.

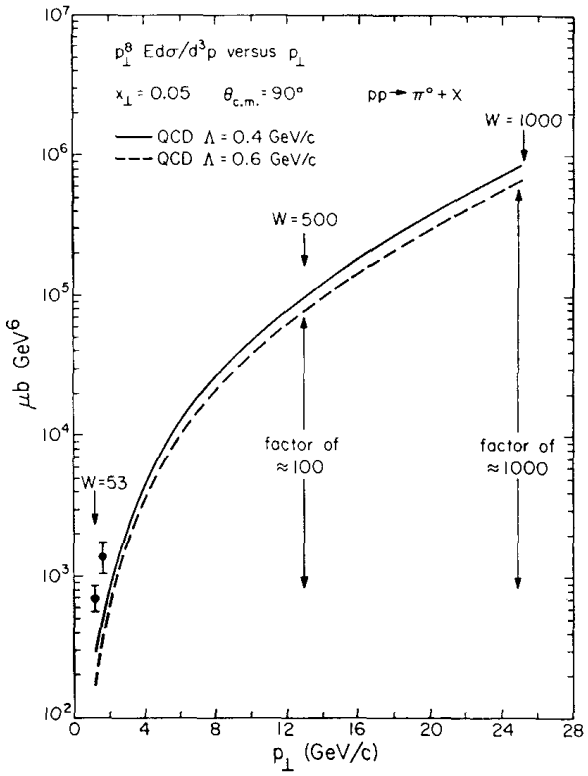


FIG. 26. The behavior of $p_\perp^8$ times the 90° single-π^0 cross section, $Ed\sigma/d^3p$, at $x_\perp = 0.05$ versus $p_\perp$ calculated from the QCD approach with $\Lambda = 0.4$ GeV/c (solid curve) and $\Lambda = 0.6$ GeV/c (dashed curve). The two low-$p_\perp$ data points are at $W = 53$ and 63 (Ref. 74). The predictions are a factor of 100 (1000) times larger than the flat $(p_\perp^{-8})$ extrapolation to $W = 500$ GeV (1000 GeV).

are roughly independent of $p_\perp$ over the range. The two-particle back-to-back cross section $d\sigma/dz_p(z_p \approx 1)$ reflects more closely the dependence on $p_\perp$ of the basic subprocess without the additional scale breaking due to smearing.

The predictions in Fig. 25 (and in all figures in this paper) are free from any beam and target jet background of the type discussed in Sec. IV D 1 above. As seen in Fig. 18, below $p_\perp = 3.5$ GeV/c this background is important and is presumably the cause of the rise of dN/dz_p at low $p_\perp$(trig). Any such increase of the expected dN/dz_p at low $p_\perp$ due to background would vitiate the comparison in Fig. 25 by making it behave similarly to the single-particle cross section. The test must be performed at $p_\perp$'s large enough so that the background contamination is negligible. This is why we calculated the results in Fig. 25 at $x_\perp = 0.35$ so that $p_\perp \gtrsim 4$ GeV/c.

E. Very-high-energy expectations

Figure 8 shows that the QCD predictions quickly deviate from a $1/p_\perp^8$ behavior (at fixed $x_\perp$) as the $p_\perp$ increases yielding a much larger cross section than expected from the black-box model. This is also seen in Fig. 26 where we plot the QCD predictions for $p_\perp^8$ times $Ed\sigma/d^3p$ versus $p_\perp$ at $x_\perp = 0.05$ and $\theta_{c.m.} = 90°$. At $W = 500$ GeV, the QCD results are a factor of 100 larger than a straight $(1/p_\perp^8)$ extrapolation and show a factor of 1000 increase at $W = 1000$ GeV. In Fig. 27 we display the predictions for 90° π^0 and jet production at fixed $W = 53$, 500, and 1000 GeV versus $p_\perp$. The preliminary high-$p_\perp$ data from CCOR (Ref. 67) at $W = 53$ GeV are also shown. The black-box model and the QCD predictions agree with each other and both agree with the data. By going to higher energy, one can easily discriminate between the two approaches. For example, at $W = 500$ GeV and $p_\perp = 30$ GeV/c, the π^0 (jet) cross section from QCD is roughly a factor of 100 (500) times larger than the FF1 results. In fact, the $p_\perp = 30$ GeV/c 90° π^0 cross section at $W = 500$ GeV is predicted in the QCD approach to be about the same magnitude as that measured at $p_\perp = 6.0$ GeV/c at Fermilab ($W = 19.4$ GeV).

These large single-particle and jet cross sections (see also Fig. 14) predicted by QCD, if correct, will make it very difficult, if not impossible, to find the W boson (and other new parti-

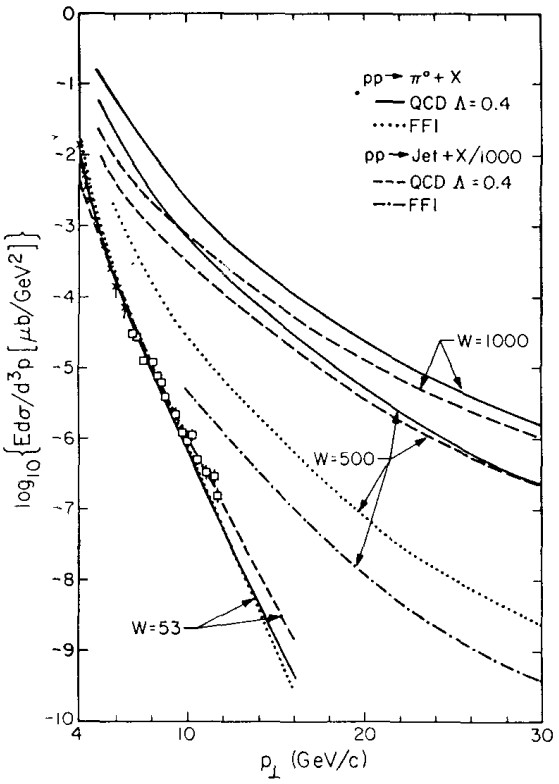


FIG. 27. Comparison of the results on the 90° π^0 cross section, $Ed\sigma/d^3p$, from the QCD approach with $\Lambda = 0.4$ GeV/c (solid curve) and the quark-quark black-box model of FF1 (dotted curves). Both models agree with the data at $W = 53$ GeV (crosses = Ref. 52) where the open squares are the "preliminary" data from the CCOR collaboration (Ref. 67) normalized to agree with the lower-$p_\perp$ experiments. The QCD approach results in much larger cross sections than the FF1 model at $W = 500$ and 1000 GeV. The FF1 results at 1000 GeV (not shown) are only slightly larger than the results at 500 GeV. Also shown are the cross sections for producing a jet at 90° (divided by 1000) as predicted by the QCD approach (dashed curves) and the FF1 model (dot-dashed curve).

cles) from its $q\bar{q}$ or jet-jet decay. Quigg[68] showed that even the black-box ($1/p_\perp^8$) model extrapolation led to a W signal that was, at best, 10 times the hadronic (jet-jet) background. The factor of 500 (1000) increase in this background at $W = 500$ (1000) GeV predicted is obviously fatal. However, if one indeed observes such large production rate for single particles and jets, then QCD will be verified and this may be as important as discovering the W boson.

It is not clear yet precisely what the quark and gluon jets will look like at very high $p_\perp$ (such as $p_\perp = 30$ GeV/c). If QCD is correct, they will certainly not look like the well collimated $\langle k_\perp \rangle_{q \to h} = 130$ MeV objects we use in this calculation and illustrated in Fig. 1. At $p_\perp = 30$ GeV/c, they should "appear" to be fatter. This is because as the $p_\perp$ of the outgoing quark is increased, it becomes increasingly more likely that it radiate a gluon and become two jets (one quark and one gluon). Then, this quark or gluon might radiate producing still more jets. The net result is that most of the time it will look as if there is one fat jet; however, occasionally when the radiation is hard enough, one will see the two or three distinct subjets.[69-71] Much theoretical effort is being focused on such questions and we should soon have a good idea of precisely what to expect at very high energies and $p_\perp$'s.

V. SUMMARY AND CONCLUSIONS

If this work is viewed as simply a comparison of one phenomenological model against another (e.g., the QCD approach versus the black-box model), not much can be said to favor one over another. It is true that the QCD approach has fewer free parameters in the parton cross sections $d\hat{\sigma}/d\hat{t}$, but there are more free choices in the gluon functions. More excuses are needed concerning background effects at low $p_\perp$, etc. It is true that the black-box pure quark scheme could not fit the away-side large-$p_\perp$ particle multiplicities and charge ratios, but it probably could be fixed up with the inclusion of gluons. It has become apparent that present high energies are not really high enough to isolate the manifold of effects (parton distributions, fragmentation functions, constituent cross sections, transverse momentum of partons, different kinds of constituents, etc.) that are mixed together so intimately in today's experiments. If the resolution of this would depend entirely on experiment, we shall have to end this long research with the tiresome and obvious call for still higher energies. At high $p_\perp$, predictions of the QCD approach are orders of magnitude greater than the black-box $p_\perp^{-8}$ extrapolations, so clear tests lie there.

But QCD is more than a phenomenological model. It is a precise and complete theory purporting to be an ultimate explanation of all hadronic experiments of all energies, high and low. There are many reasons to hope and expect it to be right. The question is, is it indeed right? Mathematical complexity has, so far, prevented us from quantitatively testing its correctness. What it predicts is not clearly known. Nevertheless, its property of asymptotic freedom leads us to expect that phenomena of high momentum transfer should be analyzable (by perturbation theory). Yet experiments at what was thought to be high enough $p_\perp$ seemed to show $p_\perp^{-8}$ behavior unlike the expected $p_\perp^{-4}$ (with possible logarithmetic modifications).

It was a mystery. Although many people said "perhaps the energy is not high enough," the remark was simply an article of faith; the mechanism leading to an apparent eighth power in the experimental region remained unknown.

We believe we have resolved this mystery, using the QCD theory itself to tell us what might happen in the range in question. There is, from the point of view of QCD, no mystery. The energy ($p_\perp$) is indeed too low and there are too many nonasymptotic effects acting. Results closer to a $p_\perp^{-4}$ falloff should appear only at much higher $p_\perp$ (see Fig. 9). Machines currently planned for these energies will resolve the question of models as soon as they are turned on.

On the other side, there is a great deal of data now available at energies and $p_\perp$ values in which asymptotic free field theory can make much more precise predictions than have yet been made. The QCD theory, unlike other phenomenological approaches, is complete mathematically so that a full discussion of theoretical predictions with limits of errors should be possible in the present range. These theoretical studies (perhaps starting at very high energy and working down) should be pursued vigorously. It is likely that among the present results of experiment, there are some that can contribute a more precise and definite test of QCD, if the theory could be developed a little further and made a bit more precise than we have done here. At the time of this writing, there is still no sharp quantitative test of QCD. An important test will come in connection with the phenomena of high $p_\perp$ discussed here.

ACKNOWLEDGMENTS

We acknowledge useful discussions with S. Brodsky, S. Ellis, H. D. Politzer, and D. Sivers. This work was supported in part by the U. S. Department of Energy under Contract No. EY76-C-03-0068.

[1]R. D. Field and R. P. Feynman, Phys. Rev. D 15, 2590 (1977).
[2]R. P. Feynman, R. D. Field, and G. C. Fox, Nucl. Phys. B128, 1 (1977).
[3]R. Baier, J. Cleymans, K. Kinoshita and B. Peterson, Nucl. Phys. B118, (1977).
[4]E. Fischbach and G. W. Look, Phys. Rev. D 15, 2576 (1977); G. W. Look and E. Fischbach, *ibid.* 16, 1369 (1977).
[5]W. Furmanski and J. Wosiek, Acta Phys. Pol. B8, 633 (1977); B8, 649 (1977).
[6]J. Kripfganz and J. Ranft, Nucl. Phys. B124, 351 (1977); E. M. Ilgenfritz, J. Kripfganz, H. J. Mohring, G. Ranft, J. Ranft, and A. Schiller, Acta. Phys. Pol. B9, 15 (1978); A. Schiller, E. M. Ilgenfritz, J. Kripfganz, H. J. Mohring, G. Ranft and J. Ranft, *ibid.* B9, 31 (1978).
[7]M. K. Chase and W. J. Stirling, Nucl. Phys. B133, 157 (1978).
[8]D. Duke, Phys. Rev. D 16, 1375 (1977).
[9]A. P. Contogouris and R. Gaskell, Nucl. Phys. B126, 157 (1977).
[10]P. V. Landshoff, talk presented at the Workshop on Future ISR Physics, 1977, edited by M. Jacob (unpublished).
[11]M. Fontannaz and D. Schiff, Nucl. Phys. B132, 457 (1978).
[12]M. K. Chase, University of Cambridge Report No. DAMTP 77/29, 1977 (unpublished).
[13]R. Raitio and R. Sosnowski, University of Helsinki Report No HU-TFT-77-22, invited talk given at the Workshop on Large $p_\perp$ Phenomena, University of Bielefeld, 1977 (unpublished).
[14]R. Baier and B. Petersson, University of Bielefeld Report No. BI-TP77/08 (unpublished).
[15]M. Della Negra *et al.* (CCHK Collaboration), Nucl. Phys. B127, 1 (1977).
[16]Effects due to the transverse momentum of the partons in the initial hadrons have been seen directly in the recent Fermilab jet-jet experiment of L. Cormell *et al.*, Fermilab Report No. 77-89, 1977 (unpublished).
[17]H. Georgi and H. D. Politzer, Phys. Rev. D 14, 1829 (1976).
[18]G. C. Fox, Nucl. Phys. B131, 107 (1977).
[19]A. J. Buras, E. G. Floratos, D. A. Ross, and G. T. Sachrajda, Nucl. Phys. B131, 308 (1977); A. J. Buras and K. J. F. Gaemers, *ibid.* 132, 249 (1978).
[20]H. L. Anderson, H. S. Matis, and L. C. Myrianthopoulos, Phys. Rev. Lett. 40, 1061 (1978); H. Anderson and T. Quirk, private communication.
[21]R. D. Field, Phys. Rev. Lett. 40, 997 (1978).
[22]G. C. Fox, invited talk presented at the Orbis Scientiae, Coral Gables, 1978 (unpublished).
[23]See also Refs. 13, 24, and 25.
[24]A. P. Contogouris, R. Gaskell, and A. Nicolaidis, McGill University report, 1977 (unpublished); A. P. Contogouris, R. Gaskell, and S. Papadopoulos, McGill University report 1978 (unpublished).
[25]S. D. Ellis, in *Particles and Fields—1977*, Proceedings of the Meeting of the APS Division of Particles and Fields, edited by P. A. Schreiner, G. H. Thomas, and A. B. Wickland (AIP, New York, 1978).
[26]H. D. Politzer, Phys. Rep. 14C, 129 (1974); invited talk at the Ben Lee Memorial Conference, 1977 [Report No. CALT-68-628 (unpublished)].
[27]R. D. Field and R. P. Feynman, Nucl. Phys. B136, 1 (1978).
[28]R. Cutler and D. Sivers, Phys. Rev. D 16, 679 (1977); 17, 196 (1978).
[29]B. L. Combridge, J. Kripfganz, and J. Ranft, Phys. Lett. 70B, 234 (1977).
[30]G. Altarelli, G. Parisi, and R. Petronzio, Phys. Lett. 76B, *ibid.* 351 (1978); G. Altarelli, G. Parisi, and R. Petronzio, 76B, 356 (1978).
[31]K. H, Craig and C. H. Llewellyn Smith, Phys. Lett. 72B, 349 (1978).
[32]H. Fritzsch and P. Minkowski, Phys. Lett. 73B, 180 (1978).
[33]K. Kajantie and R. Raitio, Nucl. Phys. B139, 72 (1978).
[34]J. F. Martin *et al.*, Phys. Rev. Lett. 40, 283 (1978).

[35]G. Donaldson *et al.*, Phys. Rev. Lett. 40, 917 (1978).
[36]C. Bromberg *et al.*, invited talk by E. Malamud at the VII International Symposium on Multiparticle Dynamics, Kayersberg, France, 1977 (unpublished).
[37]M. D. Corcoran *et al.*, Phys. Rev. Lett. 41, 9 (1978).
[38]See also G. C. Fox, in *Particles and Fields-1977*, proceedings of the Meeting of the Division of Particles and Fields of the APS, edited by P. A. Schreiner, G. H. Thomas, and A. B. Wicklund (AIP, New York, 1978).
[39]See also S. D. Ellis, M. Jacob, and P. V. Landshoff, Nucl. Phys. B108, 93 (1976); M. Jacob and P. Landshoff, *ibid.* B113, 395 (1976).
[40]C. Bromberg *et al.*, Nucl. Phys. B134, 189 (1978); C. Bromberg *et al.*, Phys. Rev. Lett. 38, 1447 (1977).
[41]G. C. Fox, in *Particles and Fields '76*, proceedings of the Annual Meeting of the Division of Particles and Fields of the APS, edited by H. Gordon and R. F. Peierls (BNL, Upton, New York, 1977), p. G1.
[42]We use the same notation as defined in Table I and Fig. 2 of FFF. The fractional momenta z_p is given by $z_p = -p_x(h_2)/p_\perp(h_1)$, where h_1 and h_2 are the trigger and away-side hadron, respectively, and p_x is the perpendicular component of momentum in the scattering plane formed by the trigger, beam, and target.
[43]H. Boggild (British-French-Scandinavian Collaboration), invited talk at the Eighth Symposium on Multiparticle Dynamics, Kaysersberg, France 1977 (unpublished); R. Moller, in *Leptons and Multileptons*, proceedings of the XII Recontre de Moriond, edited by J. Tran Thanh Van (Editions Frontieres, Paris, 1977).
[44]In addition to the $\ln(Q^2/\Lambda^2)$ variation in the denominator of Eq. (3.1), there is, in a more precise analysis, a term of the form $\ln[\ln(Q^2/\Lambda^2)]$ whose variation with Q^2 we have neglected. See W. E. Caswell, Phys. Rev. Lett. 33, 244 (1974).
[45]W. B. Atwood *et al.*, Phys. Lett. 64B, 479 (1976); W. B. Atwood, SLAC Report No. SLAC-185, thesis, 1975 (unpublished).
[46]Although in this paper we have calculated with the old solutions of Ref. 18, we have verified that the new data lead only to small changes in these solutions. In particular, the best value of Λ is unchanged.
[47]J. F. Owens, Phys. Lett. 76B, 85 (1978).
[48]The matrix $\bar{R}_{ij}$ is simply the transpose of R_{ij}.
[49]D. J. Gross, Phys. Rev. Lett. 32, 1071 (1974).
[50]M. G. Albrow *et al.*, Nucl. Phys. B135, 461 (1978).
[51]D. C. Hom *et al.*, Phys. Rev. Lett. 36, 1239 (1976), 37, 1374 (1976); S. W. Herb *et al.*, *ibid.* 39, 252 (1977); W. R. Innes *et al.*, *ibid.* 39, 1240 (1977).
[52]See also J. F. Owens, E. Reya, and M. Gluck, Phys. Rev. D 18, 1501 (1978).
[53]See also F. Halzen, G. A. Ringland, and R. G. Roberts, Phys. Rev. Lett. 40, 991 (1978).
[54]To avoid divergences due to increasing $\alpha_s(Q^2)$ at low Q^2 in (3.1), we never allow Q^2 to be smaller than our reference momentum $Q_0^2 = 4$ GeV^2 [i.e., Q^2 is given by (3.2) if it is greater than Q_0^2 and set equal to Q_0^2 elsewhere].
[55]At W = 19.4 GeV, $\theta = 90°$ and $p_\perp = 1.94$ GeV/c changing the cutoff M_0^2 from 1.0 GeV^2 to 0.3 GeV^2 causes a factor of 2 increase in the resulting invariant cross section for $pp \to \pi^0 + X$. Increasing M_0^2 to 3.0 GeV^2 decreases the cross section at this energy and $p_\perp$ by about a factor of 2.
[56]R. Blankenbecler, S. J. Brodsky, and J. F. Gunion, Phys. Rev. D 18, 900 (1978).
[57]R. D. Field, in *Leptons and Multileptons*, proceedings of the XII Recontre de Moriond, edited by J. Tran Thanh Van (Editions Frontieres, Paris, 1977).
[58]The quantity z_c is the fraction of the toward-side parton momentum carried by the trigger hadron (see Fig. 1).
[59]The away-side multiplicity distribution, $n(z_p) \equiv [1/N(\text{trig})]\, dN(\text{away})/dz_p$, is the number of away hadrons between z_p and $z_p + dz_p$ per trigger hadron, where z_p is defined in Ref. 42. See Eqs. (6.1) and (6.2) of FFF.
[60]See also the estimates made in Ref. 15.
[61]R. Blankenbecler, S. J. Brodsky, and J. F. Gunion, Phys. Rev. 106, 2652 (1972); Phys. Lett. 42B, 461 (1973); R. Blankenbecler and S. J. Brodsky, Phys. Rev. D 10, 2973 (1974); R. Blankenbecler, S. J. Brodsky, and J. Gunion, *ibid.* 12, 3469 (1975); R. Raitio and G. Ringland, *ibid.* 14, 2291 (1976).
[62]Some of the QCD terms in Table I such as $gg \to q\bar{q}$ do have a strong correlation between the toward and away quark flavor; however, these terms contribute little to the total single-particle rate and thus have only a very small effect.
[63]S. J. Brodsky, invited talk at the VIII International Symposium on Multiparticle Dynamics, Kaysersberg, France, 1977 [SLAC Report No. SLAC-PUB-2009 (unpublished)].
[64]R. J. Fisk *et al.*, Phys. Rev. Lett. 40, 984 (1978); R. J. Engelmann, private communication.
[65]B. Combridge, in ISR Discussion Meeting Between Theorists and Experimentalists Number 21, 1977, edited by M. Jacob (unpublished).
[66]Other ways of eliminating this trigger bias have been suggested by W. Ochs and L. Stodolsky, Phys. Lett. 69B, 225 (1977). Also see P. V. Landshoff and J. C. Polkinghorne, Phys. Rev. D 18, 3344 (1978).
[67]CERN-Columbia-Oxford-Rockefeller Experiment, reported by L. D. Lella in the Workshop on Future ISR Physics, 1977, edited by M. Jacob (unpublished).
[68]C. Quigg, Rev. Mod. Phys. 49, 297 (1977); also, R. F. Peierls, T. L. Trueman, and L. L. Wang, Phys. Rev. D 16, 1397 (1977).
[69]G. Sterman and S. Weinberg, Phys. Rev. Lett. 39, 1436 (1977).
[70]J. Ellis, M. K. Gaillard, and G. G. Ross, Nucl. Phys. B111, 253 (1976).
[71]A. De Rújula, J. Ellis, E. G. Floratos, and M. K. Gaillard, Nucl. Phys. B138, 387 (1978).
[72]The points at $x = 0.5$ correspond to all the SLAC data referred to in Table I of Ref. 18.
[73]J. W. Cronin *et al.* (CP Collaboration), Phys. Rev. D 11, 2105 (1975); D. Antreasyan *et al.*, Phys. Rev. Lett. 38, 112 (1977); 38, 115 (1977).
[74]B. Alper *et al.* (BS Collaboration), Nucl. Phys. B100, 237 (1975).
[75]F. W. Busser *et al.*, Nucl. Phys. B106, 1 (1976).
[76]G. Donaldson *et al.*, Phys. Rev. Lett. 36, 1110 (1976).
[77]D. C. Carey *et al.*, Fermilab Report No. FNAL-PUB-75120-EXP, 1975 (unpublished).
[78]M. B. Albrow *et al.* (CHLM Collaboration), Nucl. Phys. B73, 40 (1974).
[79]R. Cottrell *et al.*, Phys. Lett. 55B, 341 (1975).
[80]H. Frisch, in *Particles and Fields 76*, proceedings of the Annual Meeting of the Division of Particles and Fields of the APS, edited by H. Gordon and R. F. Peierls (BNL, Upton, New York, 1977), p. F59.

Nuclear Physics B136 (1978) 1–76

A PARAMETRIZATION OF THE PROPERTIES OF QUARK JETS *

R.D. FIELD and R.P. FEYNMAN
California Institute of Technology, Pasadena, California 91125, USA

Received 11 October 1977

A model is analyzed that provides a parametrization of the properties of the jet of mesons generated by a fast outgoing quark. It is assumed that the meson that contains the original quark leaves momentum and flavor to a remaining jet in which the particles are distributed (except for scaling of the energy and possible changes of flavor) like those of the original jet. One function, the probability $f(\eta)$ that the remaining jet has a fraction η of the momentum of the original jet, is chosen (as a parabola) so the final distribution of charged hadrons agrees with data from lepton experiments. All the properties of quark jets are determined from $f(\eta)$ and three parameters; the degree that SU(3) is broken in the formation of new quark-antiquark pairs ($\bar{s}s$ is taken as half as likely as $\bar{u}u$), the spin nature of the primary mesons (assumed to be vector and pseudoscalar with equal probability), and the mean transverse momentum given to these primary mesons. Monte Carlo methods are used to generate typical jets. Analytic approximations are also given. Many features of quark jets are examined. The distribution of momentum of various hadrons $D_q^h(z)$, the properties of the hadrons of largest momentum in the jet, correlations, rapidity-gap distributions, distribution of charge and of transverse momentum are some of the subjects discussed. The appearance of the jets to an instrument sensitive only to particles above some minimum momentum is also described. Although the model is probably not a true description of the physical mechanism responsible for quark jets, many predictions of the model seem quite reasonable, possibly much like real quark jets (except that the possibility of the emission of baryons is disregarded). The purpose of this work is to provide a model useful in the design of experiments in which quark jets may be observed, and further to provide a standard to facilitate the comparison of lepton-generated jets with the high-$p_\perp$ jets found in hadron collisions.

1. Introduction

Recent data from ISR [1,2] and Fermilab [3] indicate that the "jets" observed in large-$p_\perp$ hadron-hadron collisions are similar to those in processes initiated by leptons (i.e., e^+e^-, ep, and νp processes). The "jets" observed in both cases are thought to arise from quarks that fragment or cascade into a collection of hadrons moving in roughly the direction of the original quark.

* Work supported in part by the US Energy Research and Development Administration under Contract No. EY76-C-03-0068.

The experimental verification of this should proceed simultaneously in two directions. First, the detailed properties of the jets produced in lepton reactions must be examined. (Incidentally, we must confirm *via* charge properties, etc., that these jets could actually arise from quarks.) Secondly, the hadron-initiated large-$p_\perp$ jets must be compared in great detail to the lepton jets to determine if they are actually identical. (Some theorists believe that gluon jets will be produced at large $p_\perp$ in hadron-hadron collisions in addition to quark jets.) At the present time there is little data of either kind. However, hadron "jet trigger" experiments are proceeding or are planned. There is no comprehensive theory of the details of the jet structure that we should expect to observe. So, it will be difficult to know in, say, a hadron experiment what data to collect and how to summarize it so that it will be useful to compare to some future lepton experiment which will probably not measure precisely the same thing. We thought it might prove useful to have some easy-to-analyze "standard" jet structure to compare to. Thus, a hadron experiment could say "the real jets differ from the 'standard' in such and such a way", and the lepton experiment could then see whether they deviated from the same 'standard' in a similar way.

In a previous paper [4] (hereafter called FF1), we used limited experimental data aided by some theoretical ideas to suggest parametrizations for the functions $D_q^h(z)$, the mean number of hadrons of type h and momentum fraction z (per dz) in a jet initiated by a quark of flavor q with high momentum. The model presented here provides a new, much simpler, parametrization for these functions, making only negligible changes for those functions that were determined by experiment and otherwise in agreement with all the theoretical ideas in FF1.

A virtue of the new model is that it gives detailed answers to many other questions such as the number of correlated pairs of hadrons h_1, h_2 at z_1 and z_2, or the distributions of momentum gaps containing no hadrons, or the probability of observing various total jet charges, etc. In addition, one can ask for the probability that a quark of large momentum fragments so that the sum of the fractional momenta z_1, z_2, ... of all those hadrons with $z_i > z_{\min}$ is z. The latter is useful in analyzing "jet trigger" experiments. In this type of experiment, one triggers on a collection of particles that sum to give a large $p_\perp$. It is experimentally very difficult to define a "jet". One can never be sure that all the low-momentum particles from the quark are included *or* that one has not included some extra low-$p_\perp$ particles from the background of particles moving in the beam or target jets and not properly belonging to the transverse jet. These experiments would be much cleaner if one sets a threshold $p_{\perp_0}$, say, 500 MeV, for the transverse momenta of the particles whose total momentum makes up the trigger (so $z_{\min} = p_{\perp_0}/p_\perp(\text{quark})$).

We generate typical jets using Monte Carlo methods but provide an analytic approximation for the convenience of the reader. The predictions of the model are reasonable enough physically that we expect it may be close enough to reality to be useful in designing future experiments and to serve as a reasonable approximation to compare to data. We do not think of the model as a sound physical theory, and dis-

cuss why in a later section. Also, as worked out here, the model does not include baryons. We must imagine the jets to contain only mesons because we have so little knowledge at present of what baryons to expect and the character of the model does not make a clear suggestion. It also makes no clear suggestion of what correlations in transverse momentum of the hadrons to expect. We have added an additional assumption to determine this which seems reasonable to us but may be far from the physical situation. Some of the features of the transverse-momentum distributions are, however, strongly affected by the fact that the particles observed are often products of decays of higher resonances. These effects are not dependent on the details of our jet model and should be important also in analyzing the beam and target jets in ordinary inelastic hadron-hadron collisions.

Our quark-jet model involves one arbitrary function, the probability $f(z)$ that the hadron containing the original quark leaves the remaining jet a fraction z of its momentum. This ultimately determines the momentum distribution of hadrons. We have found that taking $f(z)$ to be a parabola with one adjustable parameter results in an adequate fit to the distribution of charged hadrons, $D_q^{h^+}(z) + D_q^{h^-}(z)$, observed in lepton experiments. All the properties of quark jets are then determined from $f(z)$ and three additional parameters; the degree that SU(3) is broken in the formation of new quark-antiquark pairs ($s\bar{s}$ is taken as half as likely as $\bar{u}u$), the spin of the primary mesons (assumed to be vector and pseudoscalar with equal probability), and the mean transverse momentum given to these primary mesons. This later parameter is determined by requiring that the final hadrons (after decay) have a mean transverse momentum of about 330 MeV.

2. The model

2.1. The ansatz

We assume that quark jets can be analyzed on the basis of a recursive principle. The ansatz is based on the idea that a quark of type "a" coming out at some momentum W_0 in the z direction creates a color field in which new quark-antiquark pairs are produced. Quark "a" then combines with an antiquark, say "$\bar{b}$", from the new pair $\bar{b}b$ to form a meson "$a\bar{b}$" leaving the remaining quark "b" to combine with further antiquarks. The "meson" $a\bar{b}$ may be directly observed as a pseudoscalar meson, or it may be a vector or higher-spin unstable resonance which subsequently decays into the observed mesons. To avoid complicating the ideas, we will call "$a\bar{b}$" the "primary" meson state and shall discuss secondary decay processes later. A "hierarchy" of primary mesons is formed of which $a\bar{b}$ is first in "rank", $b\bar{c}$ is second in rank, $c\bar{d}$ is third in rank, etc., as shown in fig. 1. (The "rank" in "hierarchy" should *not* be confused with order in momentum, but only order in the flavor relationships. The rank-2 primary meson may sometimes obtain a larger momentum than the rank-1 primary meson.)

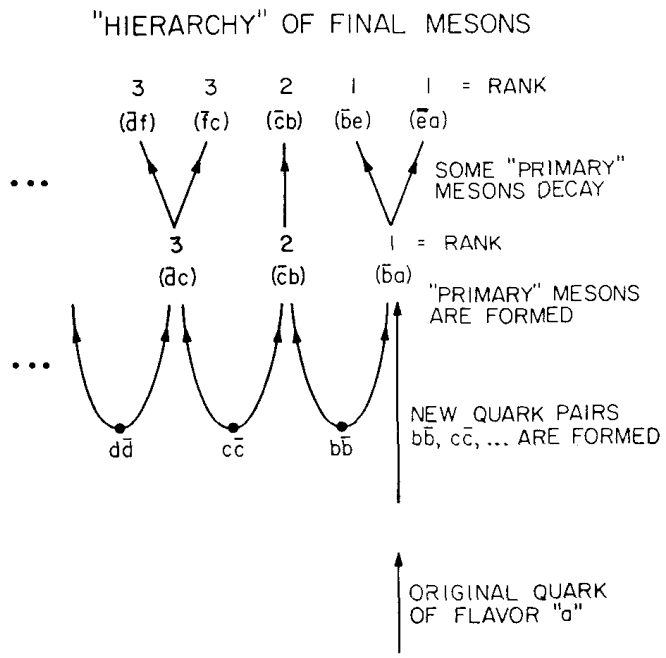


Fig. 1. Illustration of the "hierarchy" structure of the final mesons produced when a quark of type "a" fragments into hadrons. New quark pairs $b\bar{b}$, $c\bar{c}$, etc., are produced and "primary" mesons are formed. The "primary" meson $\bar{b}a$ that contains the original quark is said to have "rank" one and primary meson $\bar{c}b$ rank two, etc. Finally, some of the primary mesons decay and we assign all the decay products to have the rank of the parent. The order in "hierarchy" is *not* the same as order in momentum or rapidity.

The "chain decay" ansatz * assumes that, if the rank-1 primary meson carries away a momentum ξ_1 (from a quark jet of type "a" and momentum W_0) the remaining cascade starts with a quark of type "b" with momentum $W_1 = W_0 - \xi_1$ and the remaining hadrons are distributed in exactly the same way as the hadrons which come from a jet originated by a quark of type "b" with momentum W_1. It is further assumed that for very high momenta, all distributions scale so that they depend only on ratios of the hadron momenta to the quark momenta. Given these assumptions, complete knowledge of the structure of a quark jet is determined by one unknown function $f(\eta)$ and three parameters describing flavor, primary meson spin, and transverse momentum to be discussed later. The function $f(\eta)$ is defined by

$$f(\eta)\, d\eta = \text{the probability that the first hierarchy (rank-1) primary meson } \textit{leaves} \text{ the fraction of momentum } \eta \text{ to the remaining cascade,} \quad (2.1)$$

* We believe this recursive principle was first suggested by Krywicki and Petersson [6] and by Finkelstein and Peccei [7] in an analysis of proton-proton collisions.

and is normalized so that

$$\int_0^1 f(\eta)\, d\eta = 1 \,. \tag{2.2}$$

The rank-1 primary meson (with momentum fraction $z_1 = \xi_1/W_0$) contains the original quark "a" (see fig. 1) and $\eta = 1 - z_1$. Thus for an initial quark with momentum W_0, the probability that the first-hierarchy primary meson has momentum ξ_1 in $d\xi_1$ is $f(1 - \xi_1/W_0)\, d\xi_1/W_0$ and the probability that the rank-2 primary meson has momentum ξ_2 in $d\xi_2$ is $f(1 - \xi_2/W_1)\, d\xi_2/W_1$, where $W_1 = W_0 - \xi_1$, etc. The probability that we have a hierarchy sequence of primary mesons with the kth having momentum ξ_k in $d\xi_k$ is

$$\text{Prob}(\xi_1, \xi_2, ..., \xi_k)\, d\xi_1\, d\xi_2 \ldots d\xi_k \ldots = \prod_{i=1}^{\infty} f(\eta_i)\, d\eta_i \,, \tag{2.3}$$

where $\eta_i = W_i/W_{i-1}$ with $\xi_i = W_{i-1} - W_i$. That is, $\eta_i = 1 - \xi_i/W_{i-1}$ and $d\eta_i$ is to be replaced by $d\xi_i/W_i$ with $W_i = W_0 - \Sigma_{k=1}^{i} \xi_k$.

2.2. Single-particle decay distribution $F(z)$

The above ansatz leads to an obvious and simple Monte Carlo calculation of a jet as well as to a straightforward recursive integral equation. For example, if we define a single-particle distribution in the quark jet as

$F(z)\, dz$ = the probability of finding any primary meson (independent of hierarchy) with fractional momentum z within dz in a quark jet, (2.4)

then $F(z)$ must satisfy the following integral equation (take $W_0 = 1$)

$$F(z) = f(1 - z) + \int_z^1 f(\eta) F(z/\eta)\, d\eta/\eta \,, \tag{2.5}$$

where the limits are automatic since we define $f(1 - z) = 0$ and $F(z) = 0$ for $z > 1$ or $z < 0$. Eq. (2.5) arises because the primary meson might be the first in rank (with probability $f(1 - z)\, dz$) or if not, then the first-rank primary meson has left a momentum fraction η with probability $f(\eta)\, d\eta$, and in this remaining cascade the probability to find z in dz is $F(z/\eta)\, dz/\eta$ by the scaling principle. Dividing out the dz leaves eq. (2.5).

An integral equation for $F(z)$ given $f(\eta)$ as (2.5) involves only differences in rapidity ($Y_z = -\ln z$) • and hence can easily be analyzed by a Fourier transform

• In this paper, we will use the word rapidity to refer either to the "z rapidity" given by $Y_z = -\ln z$ in (3.4) *or* the true rapidity given by eq. (3.5). Usually the difference between the two will not be important; however, when it is, we label the former by Y_z and the latter by Y.

in rapidity, or equivalently by taking moments in z. If we define

$$M(r) = \int_0^1 z^r F(z)\, dz\,, \tag{2.6a}$$

$$C(r) = \int_0^1 \eta^r f(\eta)\, d\eta\,, \tag{2.6b}$$

then eq. (2.5) takes the form

$$M(r) = A(r) + C(r)\, M(r) \tag{2.7a}$$

or

$$M(r) = A(r)/(1 - C(r))\,, \tag{2.7b}$$

where

$$A(r) = \int_0^1 z^r f(1 - z)\, dz\,. \tag{2.7c}$$

The function $A(r)$ is, of course, related to $C(r)$ for they are determined by the same function, but the point we wish to make here is more general. The integral kernel of (2.5) can be inverted algebraically in moment space as

$$1/(1 - C(r)) = 1 + C(r)/(1 - C(r))\,. \tag{2.7d}$$

Thus an equation of the form

$$\phi(z) = a(z) + \int f(\eta)\, \phi(z/\eta)\, d\eta/\eta \tag{2.8a}$$

can be inverted to give

$$\phi(z) = a(z) + \int g(\eta)\, a(z/\eta)\, d\eta/\eta\,, \tag{2.8b}$$

where

$$\int \eta^r g(\eta)\, d\eta = C(r)/(1 - C(r))\,. \tag{2.8c}$$

For example, (2.5) is solved by

$$F(z) = f(1 - z) + \int_z^1 g(\eta)\, f(1 - z/\eta)\, d\eta/\eta\,, \tag{2.9}$$

which can be interpreted in the following way. If a particle is found at z, either it is a first-rank primary meson or else the particles of lower rank have left a momentum

η, and it is the first of those remaining (i.e., $f(1-z/\eta)\,dz/\eta$). That is, $g(\eta)\,d\eta$ has the significance of being the probability that all the primary mesons of lower rank than a given particle have left momentum fraction η of the original momentum of the jet.

The probability normalization of $f(\eta)$ given by (2.2) means that $C(0) = 1$. Then from eq. (2.5) (or (2.7b) since $A(1) = 1 - C(1)$), one finds that the total momentum of all the primary mesons in the quark jet, $M(1)$, is unity (i.e., equal to the original quark momentum). Namely,

$$\int_0^1 zF(z)\,dz = 1\ . \tag{2.10}$$

Another important point is that for small z, eq. (2.5) implies that $F(z)$ has the expected $R\,dz/z$ behavior (R = constant). This represents a uniform distribution in rapidity, $Y_z = -\ln z$. We see from (2.7b) that $M(r)$ diverges as $r \to 0$, since $C(0) = 1$, as R/r where

$$1/R = -dC(r)/dr|_{r=0}\ , \tag{2.11a}$$

or

$$1/R = -\int_0^1 \ln \eta f(\eta)\,d\eta\ . \tag{2.11b}$$

For small r, the integral (2.6a) is dominated by small z if $F(z)$ is R/z and is R/r. This may also be understood as follows. Since rapidity Y_z is $-\ln z$, $1/R$ from (2.11b) is the mean loss of rapidity, $\ln \eta$, per primary meson. Thus, deep in the cascade we must have R primary mesons per unit of rapidity $dY_z = dz/z$. For small η, in the plateau, $g(\eta)$ in eq. (2.8c) is given by $g(\eta) = R/\eta$.

2.3. Double-decay distribution $F_2(z_1, z_2)$

Suppose we want the probability of finding two primary mesons, one at z_1, the other at z_2, regardless of their rank in hierarchy. First, we define the function $F_2(z_1, z_2)$ by

$$F_2(z_1, z_2)\,dz_1\,dz_2 = \text{the probability of finding two primary mesons, one at } z_1 \text{ of any rank and one at } z_2, \text{ but of higher rank than the one at } z_1\ . \tag{2.12}$$

We can immediately write the integral equation

$$F_2(z_1, z_2) = f(1-z_1)\,F(z_2/(1-z_1))/(1-z_1) + \int f(\eta)\,F_2(z_1/\eta, z_2/\eta)\,d\eta/\eta^2\ . \tag{2.13}$$

The first term arises because z_1 might be the first-rank primary meson (probability

$f(1-z_1)\mathrm{d}z_1)$, and z_2 any primary meson in the following cascade (of total momentum $1-z_1$) and hence with probability $F(z_2/(1-z_1))\mathrm{d}z_2/(1-z_1)$. The second term arises if z_1 is not of rank one. The rank-one primary meson leaves the others with momentum η with probability $f(\eta)\mathrm{d}\eta$ and then we find z_1, z_2 with scaled probability $F_2(z_1/\eta, z_2/\eta)\,\mathrm{d}z_1/\eta\,\mathrm{d}z_2/\eta$. Eq. (2.13) can be solved by the techniques discussed in subsect. 2.2, whereupon one gets

$$F_2(z_1, z_2) = f(1-z_1)F(z_2/(1-z_1))/(1-z_1)$$

$$+\int_{z_1+z_2}^{1} g(\eta)\, f(1-z_1/\eta)\, F(z_2/(\eta-z_1))\,\mathrm{d}\eta/(\eta(\eta-z_1))\,. \tag{2.14}$$

The complete double-fragmentation function, where the meson at z_2 may be either of higher or lower rank than that at z_1, is then given by

$$\widetilde{F}_2(z_1, z_2) = F_2(z_1, z_2) + F_2(z_2, z_1)\,. \tag{2.15}$$

2.4. Choosing the form of $f(\eta)$

A form which makes the solution of the integral equation (2.5) the simplest is a power of η,

$$f(\eta) = (d+1)\eta^d\,. \tag{2.16}$$

In this case

$$C(r) = (d+1)/(r+d+1) \tag{2.17a}$$

or

$$C(r)/(1-C(r)) = (d+1)/r\,, \tag{2.17b}$$

so that

$$g(\eta) = (d+1)/\eta\,, \tag{2.18a}$$

and hence

$$zF(z) = (d+1)(1-z)^d\,. \tag{2.18b}$$

This means that

$$zF(z) = f(1-z)\,, \tag{2.18c}$$

which serves as a rather rough approximation for other forms of $f(\eta)$. This power form for $f(\eta)$ leads to a double-decay function of the form

$$F_2(z_1, z_2)\,\mathrm{d}z_1\,\mathrm{d}z_2 = f(1-z_1-z_2)\,\frac{\mathrm{d}z_1}{z_1}\,\frac{\mathrm{d}z_2}{z_1+z_2}\,, \tag{2.19a}$$

or disregarding order of rank,

$$\tilde{F}_2(z_1, z_2) = f(1 - z_1 - z_2)/z_1 z_2 \,, \tag{2.19b}$$

which can also serve as a rough approximation for other choices of $f(\eta)$ •. The inclusive probability to find N mesons with momenta $z_1, z_2, \ldots z_N$ is

$$\tilde{F}_N(z_1, z_2. \ldots z_N)\, dz_1 \ldots dz_N = (d+1)^{N+1} \left(1 - \sum_{i=1}^{N} z_i\right)^d \prod_{i=1}^{N} \frac{dz_i}{z_i} \,. \tag{2.20}$$

This is the manner in which mesons are distributed in the one-dimensional field-theory model of Casher, Kogut and Susskind [10]. In fact, it results simply from the assumption of particles produced independently at random having an *a priori* probability β to exist and then to be distributed uniformly in two-dimensional relativistic phase space (i.e., uniformly in rapidity, dz/z). The total energy and momentum must, however, be that of the jet ($\beta = d + 1$).

Although the forms of (2.16) and (2.18b) are simple and easy to interpret, they do not agree with the assumption we used in FF1, that the probability of finding mesons in dz approaches a constant as $z \to 1$. Eq. (2.5) has the property that $F(z)$ approaches $f(1-z)$ as z becomes large. The mesons at large z almost surely contain the original quark and at $z = 1$, the meson must contain the original quark. Thus in order for $F(z)$ to approach a constant as $z \to 1$, we must require that $f(\eta)$ approach a constant as $\eta \to 0$, something that the form (2.16) does not do. A function that leads to results similar to those found in FF1 is simply (2.16) plus a small constant. We take

$$f(\eta) = 1 - a + 3a\eta^2 \,, \tag{2.21}$$

where the parameter a and the power, $d = 2$, are chosen by comparing $F(z)$ to experiment. (Actually we compare the charged-particle distribution, $D_q^{h^+}(z) + D_q^{h^-}(z)$, to data.) The form (2.21) gives

$$C(r) = (1-a)/(r+1) + 3a/(r+3) \,, \tag{2.22a}$$

so that using (2.8c) we find

$$g(\eta) = (3/\eta + 4a(1-a)\eta^{2-2a})/(3-2a) \,, \tag{2.22b}$$

and (2.8d) yields

$$\begin{aligned} zF(z) = {} & 3/(3-2a) \\ & + 3az^2/(2a-1) \\ & + 2a(2a^2 - 3a - 2)z^{3-2a}/((3-2a)(2a-1)) \,. \end{aligned} \tag{2.23}$$

As we will see later after discussing flavor, if we interpret all primary mesons as

• Eq. (2.19b) with $f(\eta) = 2\eta$ was used by Bjorken in ref. [8] to estimate the large-$p_\perp$ $\pi^0\pi^0$ cross section in pp collisions and by Ellis, Jacob and Landshoff in ref. [9] to estimate same side large-$p_\perp$ correlations in pp collisions.

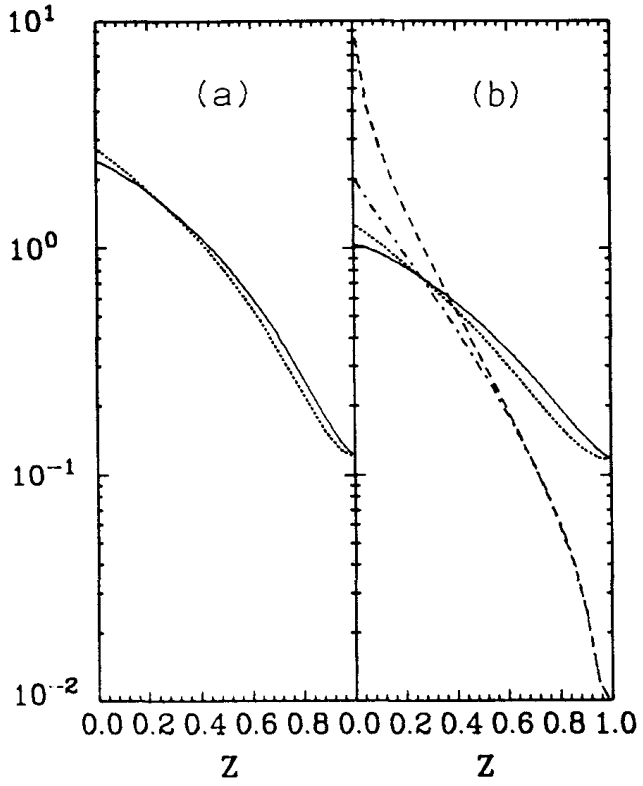


Fig. 2. (a) (left) The probability of finding the first-rank meson at z (in dz), $f(1-z)$ [---], and the probability of finding any hadron at z (times z), $zF(z)$ [——], given by eqs. (2.21) and (2.23), respectively, with $a = 0.88$, $\alpha_{ps} = 1$, $\alpha_v = 0$. (b) (right) The functions $\alpha_{ps} f(1-z)$ and $\alpha_{ps} zF(z)$ given by eqs. (2.21) and (2.23) with $a = 0.77$ together with the distribution of secondary mesons, $\alpha_v \hat{f}(1-z)$ and $\alpha_v z\hat{F}(z)$, from parents distributed according to (2.21) and (2.23), where $\alpha_{ps} = \alpha_v = 0.5$ and where we have used the parent-daughter relation in (2.54). —— $\alpha_{ps} zF(z)$; ..... $\alpha_{ps} f(1-z)$; —·—· $\alpha_v z\hat{F}(z)$; — — — $\alpha_v \hat{f}(1-z)$.

pseudoscalar mesons with no secondary decays then by comparing $D_q^{h^+}(z) + D_q^{h^-}(z)$ with data and with the results of FF1, we find that $a = 0.88$ gives a good fit. Fig. 2a shows $zF(z)$ and $f(1-z)$ for this case. We see that the relationship (2.18c) is still approximately valid.

2.5. *Including flavor*

Next we examine the question of the flavor u, d, s, $\bar{u}$, $\bar{d}$, and $\bar{s}$ of the quarks and hence the isospin and strangeness of the primary mesons. We call the isospin and strangeness properties the "flavor" of the primary mesons. For example, a "u$\bar{d}$" primary meson has the flavor of a π^+ or equally well a ρ^+. In this section, we will discuss everything in terms of the pseudoscalars and in the next section include the

possibility of forming vectors or other resonances as well. We assume, as in FF1, that new $q\bar{q}$ pairs are $u\bar{u}$ with probability γ_u, $d\bar{d}$ with equal probability γ_d and $s\bar{s}$ with probability γ_s. From isospin symmetry $\gamma_u = \gamma_d = \gamma$, say, so $\gamma_s = 1 - 2\gamma$. As in FF1, we suppose that strange $s\bar{s}$ pairs are half as likely as unstrange $u\bar{u}$ pairs so that *

$$\gamma = 0.4 \,. \tag{2.24}$$

Let the primary meson states be defined by quark labels "$a\bar{b}$", etc., where a is u, d, or s and likewise for b. A sum on such a label is a sum on u, d, and s. Our assumptions about flavor mean that if the ranks of successive primary mesons are 1, 2, 3, ..., then their flavors must be of the type of a chain $a\bar{b}$ for the first with some $\bar{b}$, then $b\bar{c}$ for the second, $c\bar{d}$ for the third, etc., as shown in fig. 1. The probability of finding them with various momenta is still (2.3) but multiplied by the factor $\gamma_b, \gamma_c, \gamma_d, \ldots$ giving the chance that the correct new pairs $\bar{b}b$, $\bar{c}c$, $\bar{d}d$, ... were indeed formed. It is now easy to modify what we did earlier in subsect. 2.2 for flavor.

For example, for a quark of type q, the mean number of primary meson states of type "$a\bar{b}$" at z is, in analogy to eq. (2.5),

$$P_q^{a\bar{b}}(z) = \delta_{qa}\gamma_b f(1-z) + \int f(\eta) \sum_c \gamma_c P_c^{a\bar{b}}(z/\eta) \,. \tag{2.25}$$

The first term arises because the "$a\bar{b}$" primary meson state might be of first rank (only if a = q, of course, hence the delta function δ_{aq}) with probability $f(1-z)$ times the chance, γ_b, that the first new pair is of the required type b. The second term occurs if the "$a\bar{b}$" primary meson is not of first rank in hierarchy. The first pair might be $c\bar{c}$ (with probability γ_c) and leave a momentum η to the cascade of quark c (with probability $f(\eta)$) in which cascade we find an "$a\bar{b}$" state with probability $P_c^{a\bar{b}}(z/\eta)\, dz/\eta$.

It is seen that the flavor distribution of the 2nd and higher rank primary mesons are independent of what quark started the cascade because we have assumed the new pairs are made with flavors independent of the quark flavor that makes the cascade. Defining a "mean quark" flavor to be $\langle q \rangle$, and equal to u, d, and s with probability γ, γ, and $(1 - 2\gamma)$, respectively, we write

$$P_{\langle q \rangle}^{a\bar{b}}(z) = \sum_c \gamma_c P_c^{a\bar{b}}(z) \,. \tag{2.26}$$

* It is not unreasonable that $s\bar{s}$ pairs are formed less often than $u\bar{u}$ and $d\bar{d}$, for s quarks may have a larger mass than u or d. For example, in a vector force field (e.g. $F = eE$ where E is an electric field and e is the charge on a particle) constant in space and time, the rate of production of pairs of particles of rest mass m and transverse momentum $\boldsymbol{k}_\perp$ ($\boldsymbol{k}_\perp$ is a two-dimensional vector transverse to the direction of the field, the particle has transverse momentum $\boldsymbol{k}_\perp$, the antiparticle $-\boldsymbol{k}_\perp$) is, according to the Dirac equation,

$$(F/2\pi) \exp(-\pi(m^2 + k_\perp^2)/F)\, d^2k_\perp/(2\pi)^2$$

per unit volume per second. Thus it is more difficult to make pairs of mass m by a factor $\exp(-\pi m^2/F)$.

The quantity $P_{\langle q \rangle}$ appears on the right hand side of eq. (2.25) so multiplying by γ_q and summing on q leaves

$$P^{a\bar{b}}_{\langle q \rangle}(z) = \gamma_a \gamma_b f(1-z) + \int f(\eta) P^{a\bar{b}}_{\langle q \rangle}(z/\eta)\, d\eta/\eta \,, \tag{2.27}$$

which, after comparing with (2.5), yields

$$P^{a\bar{b}}_{\langle q \rangle}(z) = \gamma_a \gamma_b F(z) \,. \tag{2.28}$$

Thus if we know the distribution of primary mesons $F(z)$ from quarks disregarding flavor, then those of flavor "$a\bar{b}$" occur in the "mean quark" cascade with probabilities, $\gamma_a \gamma_b$, that the first and the second quark of the pair "$a\bar{b}$" have the appropriate flavors. Only the contribution of the first-rank quark differs from the average. Substituting back into eq. (2.25) and integrating, one finds in general that

$$P^{a\bar{b}}_{q}(z) = \delta_{qa} \gamma_b f(1-z) + \gamma_a \gamma_b \bar{F}(z) \,, \tag{2.29a}$$

where

$$\bar{F}(z) = F(z) - f(1-z) \tag{2.29b}$$

is the probability of finding a primary meson at z of rank higher than one, and $F(z)$ and $f(1-z)$ are the functions given in (2.21) and (2.23).

To be more explicit, the distributions of primary meson states of flavor h from a quark q are given by

$$D^h_q(z) = A^h_q f(1-z) + B^h \bar{F}(z) \,, \tag{2.30}$$

where A^h_q and $B^h = \Sigma_q \gamma_q A^h_q$ have the values given in table 1. For example, the quark content of an eta meson is given by $\eta = S \cos\theta_M + N \sin\theta_M$, where $N = \sqrt{\frac{1}{2}}(u\bar{u} + d\bar{d})$ and $S = s\bar{s}$. Thus, the quantity A^η_u is given by $\frac{1}{2}\gamma \sin^2\theta_M$, where γ is the probability of producing a $\bar{u}$ quark ($\bar{u}u$ pair) to combine with the initial u and $\frac{1}{2}\sin^2\theta_M$ is the chance that this $\bar{u}u$ pair actually forms the desired η.

For now we shall assume all primary mesons are pseudoscalar mesons with no decays. To compare with FF1, we find from table 1 with $\gamma = 0.4$

$$D(z) \equiv D^{\pi^+}_u(z) + D^{\pi^-}_u(z) = 0.32\bar{F}(z) + 0.40 f(1-z) \,, \tag{2.31a}$$

$$K_u(z) \equiv D^{K^+}_u(z) + D^{K^-}_u(z) = 0.16\bar{F}(z) + 0.20 f(1-z) \,, \tag{2.31b}$$

$$K_s(z) \equiv D^{K^+}_s(z) + D^{K^-}_s(z) = 0.16\bar{F}(z) + 0.40 f(1-z) \,; \tag{2.31c}$$

in addition we have

$$D^{\pi^+}_u(z) - D^{\pi^-}_u(z) = 0.4 f(1-z) \,, \tag{2.32a}$$

$$D^{K^+}_u(z) - D^{K^-}_u(z) = 0.2 f(1-z) \,, \tag{2.32b}$$

Table 1
Values of the constants A^i_q, B^i, where q is the quark flavor and i is the primary meson flavor

	$\pi^+(\rho^+)$	$\pi^0(\rho^0)$	$\pi^-(\rho^-)$	$K^+(K^{*+})$	$K^0(K^{*0})$	$K^-(K^{*-})$	$\bar{K}^0(\bar{K}^{*0})$	$\eta(\omega)$	$\eta'(\phi)$
A^i_u	γ	$\frac{1}{2}\gamma$	0	$(1-2\gamma)$	0	0	0	$\frac{1}{2}\gamma \sin^2\theta_M$	$\frac{1}{2}\gamma \cos^2\theta_M$
A^i_d	0	$\frac{1}{2}\gamma$	γ	0	$(1-2\gamma)$	0	0	$\frac{1}{2}\gamma \sin^2\theta_M$	$\frac{1}{2}\gamma \cos^2\theta_M$
A^i_s	0	0	0	0	0	γ	γ	$(1-2\gamma)\cos^2\theta_M$	$(1-2\gamma)\sin^2\theta_M$
B^i	γ^2	γ^2	γ^2	$\gamma(1-2\gamma)$	$\gamma(1-2\gamma)$	$\gamma(1-2\gamma)$	$\gamma(1-2\gamma)$	$\gamma^2 \sin^2\theta_M$ $+(1-2\gamma)^2\cos^2\theta_M$	$\cos^2\theta_M\gamma^2$ $+(1-2\gamma)^2\sin^2\theta_M$

The quantity θ_M is the mixing angle defined by $\eta(\omega) = S\cos\theta_M + N\sin\theta_M$, $\eta'(\phi) = -S\sin\theta_M + N\cos\theta_M$, where $S = s\bar{s}$ and $N = \sqrt{\frac{1}{2}}(u\bar{u} + d\bar{d})$ and where we use $\theta_M = \theta_{ps} = 45^\circ$ and $\theta_M = \theta_v = 90^\circ$ for the pseudoscalar and vector nonets, respectively.

$$D_s^{K^-}(z) - D_s^{K^+}(z) = 0.4\, f(1-z)\,, \tag{2.32c}$$

$$D_u^{K^+}(z) = 0.5\, D_u^{\pi^+}(z)\,, \tag{2.32d}$$

$$D_u^{K^-}(z) = 0.5\, D_u^{\pi^-}(z)\,. \tag{2.32e}$$

These last two equations imply

$$D_u^{K^-}(z)/D_u^{K^+}(z) = D_u^{\pi^-}(z)/D_u^{\pi^+}(z)\,, \tag{2.33}$$

which is what we assumed in eq. (3.7) of FF1. In addition, we have

$$D_u^{\pi^-}(z)/D_u^{\pi^+}(z) = \frac{\gamma^2 \bar{F}(z)}{\gamma f(1-z) + \gamma^2 \bar{F}(z)}\,, \tag{2.34}$$

which approaches zero as z becomes large. This was another of our assumptions in FF1 and here it is a consequence of the property that at $z = 1$, the primary meson must contain the original quark (i.e., $F(z) = f(1-z)$ at $z = 1$ so $\bar{F}(z) \to 0$).

For this case, where all the primary mesons are pseudoscalars with no subsequent decay, the best choice for parameter a in (2.21) and (2.23) is $a = 0.88$. This is determined in fig. 3 by requiring that the model (dashed curve) agree with the lepton data on the charged-particle distribution $D_q^{h^+}(z) + D_q^{h^-}(z)$ and with the earlier FF1 choice (solid curve). The value a is chosen so that $D_q^{h^+}(z = 1) + D_q^{h^-}(z = 1) = 0.6(1 - a)$ agrees with the FF1 curve at $z = 1$.

The fraction of the total momentum carried by each type of primary meson is given by

$$\int z P_q^{a\bar{b}}(z)\, dz = \delta_{qa}\gamma_b \bar{z} + \gamma_a \gamma_b (1 - \bar{z})\,, \tag{2.35}$$

where

$$\bar{z} = \int z f(1-z)\, dz \tag{2.36a}$$

is the mean momentum of the first hierarchy meson and is given by

$$\bar{z} = \tfrac{1}{2} - \tfrac{1}{4}a, \tag{2.36b}$$

for $f(\eta)$ as in (2.21). The momentum carried by the various hadrons with $a = 0.88$ is given in table 2. In FF1 we did not include the η and η' explicitly and, as can be seen from table 2, if we allow the η and η' to decay then the model reproduces the results of FF1 quite closely.

It is of interest to ask how the charges of the primary mesons are distributed along the direction of the initial quark. Suppose we have some quantity such as charge Q, (or third component of isospin I_3, or hypercharge Y) that can be calculated for a primary meson state additively from the quark content of that meson

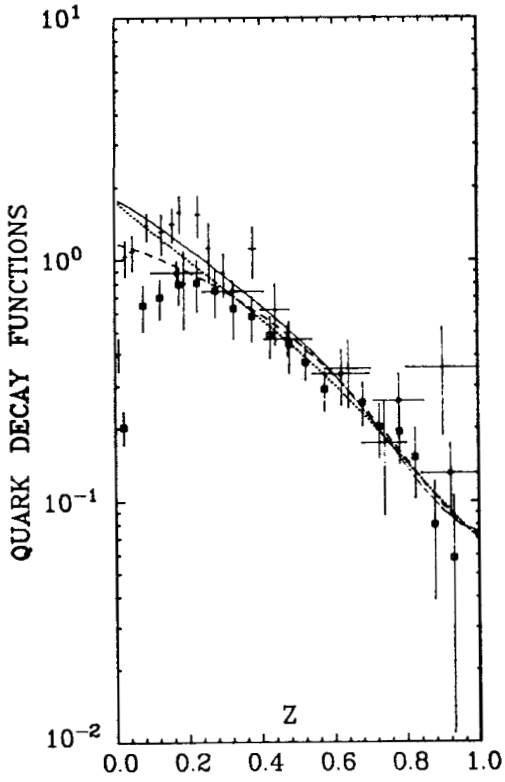


Fig. 3. Comparison of the charged particle distribution $D_u^{h^+}(z) + D_u^{h^-}(z)$ from the quark-jet model with $a = 0.88$, $\alpha_{ps} = 1$, $\alpha_v = 0$ (dashed curve) and with $a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$ (dotted curve) with the lepton data from fig. 6 of FF1 and with the distribution used in FF1 (solid line). (There is an error in the captions to fig. 6 and fig. 7a in FF1. The ep data from Daken et al. has been averaged over the *two* Q^2 bins $1.0 \leqslant Q^2 \leqslant 2.0$ GeV^2 *and* $2.0 \leqslant Q^2 \leqslant 3.0$ GeV^2 with $12 < s < 30$ GeV^2.) □ $zN(e^+e^- \to h^\pm)$; ◇ $zN(ep \to h^\pm)$; + $zN(\nu p \to h^\pm)$.

such that each quark flavor, "a", carries e_a and each antiquark $\bar{a}$ contributes $-e_a$. Then if we weigh each primary meson, $a\bar{b}$, by its charge $e_a - e_b$, we obtain the net mean charge distribution of a quark jet. Namely,

$$\langle Q_q(z)\rangle = \sum_{a,b} (e_a - e_b) P_q^{a\bar{b}}(z) = (e_q - e_{\langle q\rangle}) f(1-z)\,, \tag{2.37}$$

from (2.29), where

$$e_{\langle q\rangle} = \sum_a \gamma_a e_a \tag{2.38}$$

is the charge of a "mean" quark. Thus the hadrons in each jet carry a total mean charge, $e_q - e_{\langle q\rangle}$, equal to the charge on the quark producing the jet *plus* a constant error, $-e_{\langle q\rangle}$, proportional to the deviation of the probability of production of new

Table 2
Fraction of the total momentum of a u-, d- and s-quark carried by the primary mesons (before decay, called direct) resulting from our jet model with $a = 0.88$, $\alpha_{ps} = 1.0$, and $\alpha_v = 0.0$. In addition, the results are shown for the total mesons (direct + indirect) after the η and η' are allowed to decay. For comparison, the values from table 2 of FF1 are also given (see also table 6).

		$a = 0.88$, $\alpha_{ps} = 1$ results			FF1 results		
		u	d	s	u	d	s
direct	π^+	0.23	0.12	0.12			
	π^0	0.17	0.17	0.12			
	π^-	0.12	0.23	0.12			
	total π	0.51	0.51	0.35			
	$\eta + \eta'$	0.20	0.20	0.20			
direct + indirect	π^+	0.26	0.15	0.15	0.27	0.15	0.15
	π^0	0.24	0.24	0.18	0.21	0.21	0.15
	π^-	0.15	0.26	0.15	0.15	0.27	0.15
	total π	0.65	0.65	0.48	0.63	0.63	0.45
	K^+	0.11	0.06	0.06	0.13	0.08	0.08
	K^0	0.06	0.11	0.06	0.08	0.13	0.08
	K^-	0.06	0.06	0.17	0.08	0.08	0.19
	$\bar{K}^0$	0.06	0.06	0.17	0.08	0.08	0.19
	total K	0.29	0.29	0.45	0.37	0.37	0.54
	γ	0.06	0.06	0.06			

pairs from the SU(3) value of $\frac{1}{3}$. For the case of electric charge, we have $e_{\langle q \rangle} = \gamma - \frac{1}{3}$ and thus the mean total charge of a quark jet is

$$\langle Q_u \rangle = 1 - \gamma = 0.60 , \tag{2.39}$$

$$\langle Q_d \rangle = -\gamma = -0.40 , \tag{2.40}$$

$$\langle Q_s \rangle = -\gamma = -0.40 , \tag{2.41}$$

which is in close agreement with the empirically fitted values of 0.59, −0.40, and −0.39, respectively, found in FF1. The mean total I_3 is just the I_3 of the quark, independent of γ since I_3 of the mean quark is zero (u and d coming with equal weight). Hypercharge, Y is $2(Q - I_3)$, so its mean for u- and d-quark jets is 0.2 and for s-quark jets is −0.8.

The charge on the jet of a "mean" quark is $\Sigma_q \gamma_q (e_q - e_{\langle q \rangle}) = 0$ so that the second and higher hierarchy primary mesons carry no charge on the average. Thus, were it not for complications from secondary disintegrations, which we discuss later, the experimental determination of the net mean charge, $\Sigma_h e_h D_q^h(z)$, where e_h is the charge of the hadron h, would give a direct measure of the distribution of the first-rank primary meson, $f(1 - z)$.

At first sight one might wonder why the total average charge of the hadrons is not that of the original quark, because after all, the new quark pairs $q\bar{q}$ are each neutral. However, as pointed out by Rosner and Farrar [11] and by Cahn and Colglasier [12], for a jet starting from a quark, each new pair has its antiquark going into a hadron of one lower rank in the hierarchy than does its quark. The new antiquark has, on the average, the higher momentum. Thus, when we consider all the hadrons higher than some very small momentum z_0 (small enough to ensure that the mean number of positive and negative hadrons in dz are practically the same), we are counting more antiquarks than quarks. This makes no difference for isospin, since we have $\bar{u}$ and $\bar{d}$ antiquarks equally likely, but the average electric charge of antiquarks is not zero but is given by $e_{\langle\bar{q}\rangle} = (-\frac{2}{3} + \frac{1}{3})\gamma + \frac{1}{3}(1 - 2\gamma)$. The $\bar{s}$ quarks are not in their full number, $\frac{1}{3}$, needed to cancel the average $\bar{u} + \bar{d}$ charge. Since, in counting charge, only mesons (we leave out the complication and slight numerical modification due to baryon production) are counted, the total number of quarks and antiquarks counted must be exactly the same. Therefore, we must succeed in counting an excess of exactly one antiquark over quarks from the new pairs, to compensate the extra quark from which the jet started. Thus, the charges of all hadrons above z_0 exceed that of the original quark e_q by the charge of one mean antiquark $e_{\langle\bar{q}\rangle} = -e_{\langle q\rangle}$ (about $-\frac{1}{15}$ charge units). For this reason, the net charge of a "mean" quark jet is zero. This last result can also be seen in the following way. In a "mean" quark jet, all the mesons are formed out of pairs of "mean" quarks and "mean" antiquarks. Any tendency of the antiquark of one new pair to combine with a quark of lower rank to make a meson has no effect, for that lower-rank quark is also a "mean" quark with the same probability of flavors as the antiquark. In the general case, charge imbalance can be seen, in the mean, only in the hadron containing the original quark as indicated by (2.37).

We now turn to the problem of generalizing the double-decay function in (2.12) to include the effects of correlations in flavor. We define

$$P_q^{a\bar{b},\,c\bar{d}}(z_1, z_2)\,dz_1\,dz_2 = \text{probability of finding a primary meson of flavor ``}a\bar{b}\text{'' at } z_1 \textit{ and} \text{ one of flavor ``}c\bar{d}\text{'' at } z_2 \text{ in a jet originated by a quark of flavor q when the primary meson at } z_2 \text{ has a larger rank in hierarchy than the one at } z_1. \tag{2.42}$$

This probability is the sum of the following four pieces.

(i) The probability that the primary meson at z_1 is of rank 1 and the primary meson at z_2 is of rank 2; given by

$$\delta_{qa}\gamma_b\delta_{bc}\gamma_d f(1 - z_1)\,dz_1 f(1 - z_2/(1 - z_1))\,\frac{dz_2}{1 - z_1}\,. \tag{2.43a}$$

For this term q = a and b is arbitrary with probability γ_b, but then c = b and d comes with probability γ_d. The probability that z_1 is rank-1 is $f(1 - z_1)\,dz_1$, and

then since $1 - z_1$ momentum is left, the probability that z_2 is second-rank (i.e., rank-1 in the remaining hierarchy) is $f(1 - z_2/(1 - z_1))\,dz_2/(1 - z_1)$.

(ii) The probability that z_1 is of rank one, but z_2 is of rank higher than 2:

$$\delta_{qa}\gamma_b\gamma_c\gamma_d f(1 - z_1)\,dz_1 \bar{F}(z_2/(1 - z_1))\frac{dz_2}{1 - z_1}\,. \tag{2.43b}$$

For this case, the chance to get $c\bar{d}$ is now independent of q, a, and b and is $\gamma_c\gamma_d$.

(iii) The probability that the primary meson at z_1 is not first in rank but higher, but the primary meson at z_2 is directly of next rank to the one at z_1:

$$\gamma_a\gamma_b\delta_{bc}\gamma_d \int_{z_1+z_2}^{1} g(\eta)\,d\eta f(1 - z_1/\eta)\frac{dz_1}{\eta} f(1 - z_2/(\eta - z_1))\frac{dz_2}{\eta - z_1}\,, \tag{2.43c}$$

where the early primary mesons leave momentum η (with probability $g(\eta)\,d\eta$) and the next two primary mesons come as rank one and two. Here the flavor of q has no affect, a and b come with probability $\gamma_a\gamma_b$, but c must equal b. We must, of course, integrate over all remaining momenta η.

(iv) Neither the primary meson at z_1 or z_2 is first in rank, nor are they adjacent:

$$\gamma_a\gamma_b\gamma_c\gamma_d \int_{z_1+z_2}^{1} g(\eta)\,d\eta f(1 - z_1/\eta)\frac{dz_1}{\eta}\bar{F}(z_2/(\eta - z_1))\frac{dz_2}{\eta - z_1}\,. \tag{2.43d}$$

The complete double-fragmentation function for producing two hadrons of flavor $h_1 = a\bar{b}$ and $h_2 = cd$ is given by symmetrizing (2.43a–d) with respect to z_1 and z_2. Namely,

$$\tilde{P}_q^{a\bar{b},\,c\bar{d}}(z_1, z_2) = P_q^{a\bar{b},\,c\bar{d}}(z_1, z_2) + P_q^{c\bar{d},\,a\bar{b}}(z_2, z_1)\,. \tag{2.44}$$

2.6. The production of resonances

Finally, we must decide in what nonet a primary meson is formed. We suppose that the pair of quarks $q\bar{q}$ is in the low pseudoscalar 0^- configuration with the probability α_{ps}, that it is a vector meson 1^- with probability α_v, a tensor meson 2^+ with probability α_t, etc. Then these objects are allowed to disintegrate as we know they should from the particle tables.

We suppose that the choices of the probabilities α are dynamic and do not depend on the flavor. Although attempts have been made to determine from hadron experiments the production rate of, say, ρ^0 compared to π^0 at larger $p_\perp$, the experiments are difficult to interpret. Indications are that ρ^0 is roughly as likely as π^0 at high $p_\perp$ (see references in table 16) and that higher resonances are less likely [13]. For definiteness for the present, we shall choose

$$\alpha_{ps} = \alpha_v = 0.5\,, \tag{2.45}$$

with no higher resonances * ($\alpha_t = 0$, etc.), which yields $D_u^{o^0}(z)/D_u^{\pi^0}(z) \to 1$ as $z \to 1$. Future experiments may, however, indicate a better choice.

2.7. The Monte Carlo method

2.7.1. Recursive scheme

The method of Monte Carlo generation of a complete quark cascade is clear. Suppose one starts with a quark of flavor q and momentum W_0. Then

(i) One generates a value of $\eta_1 = 1 - z_1$ at random with probability given by $f(\eta)$ as in eq. (2.21).

(ii) One generates a quark pair $u\bar{u}$, $d\bar{d}$, or $s\bar{s}$ with probability γ, γ, and $(1-2\gamma)$ (i.e., 0.4, 0.4, and 0.2), respectively. The first primary meson state is then of type $q\bar{u}$, $q\bar{d}$, or $q\bar{s}$ depending on the pair chosen.

(iii) One decides on the spin-parity of the primary meson, according to (2.45), (i.e., pseudoscalar or vector with equal probabilities.

The first primary meson is now of momentum $(1-\eta_1)W_0$ and of type $q\bar{u}$, $q\bar{d}$ or $q\bar{s}$ depending on the choice made in (ii) and of spin depending on the choice made in (iii). This leaves, for the next step, a quark of type q_2 = u, d, or s (depending on the first quark pair chosen) with momentum $W_1 = \eta_1 W_0$. The cycle beginning with (i)–(iii) is then repeated for this quark q_2 with momentum W_1. Another η value is found and a new quark pair produced. This procedure is then repeated over and over until a desired point, to be discussed later (sect. 3 below), is reached.

Finally, we add transverse momentum according to subsect. 2.7.2 and then let the vector mesons decay, each with its known characteristic kinematics and branching ratios, as given in the particle tables. In addition, we allow the η and η' mesons to decay.

As discussed earlier, in choosing the spin-parity of the primary mesons (iii), we assume that α_{ps} and α_v are independent of flavor. For example, a $u\bar{u}$ primary meson state is a π^0 with probability $\frac{1}{2}\alpha_{ps}$, an η with probability $\frac{1}{2}\alpha_{ps}\sin^2\theta_{ps}$, an η' with probability $\frac{1}{2}\alpha_{ps}\cos^2\theta_{ps}$, a ρ^0 with probability $\frac{1}{2}\alpha_v$, a ω^0 with probability $\frac{1}{2}\alpha_v\sin^2\theta_v$ and a ϕ^0 with probability $\frac{1}{2}\alpha_v\cos^2\theta_v$ in agreement with table 1. The pseudoscalar and vector mixing angles θ_{ps} and θ_v are chosen to be 45° and 90°, respectively.

2.7.2. Including transverse momentum

The hadrons arising from the cascade or fragmentation of a quark do not travel in precisely the same direction as the initiating quark (i.e., z direction). We expect that the transverse momentum of these hadrons remains limited as the quark's momentum becomes large; however, we really have no knowledge as to the precise manner in which the various hadrons share the transverse momentum. One can imagine several ways to distribute transverse momentum among the primary mesons

* The theoretical idea that the quark spins combine randomly would recommend, rather, that $\alpha_v = 3\alpha_{ps}$.

in our jets. We will choose a particular way, not because we are sure it is nature's way, but for definiteness. Very few of the results in this paper depend on this particular choice.

We will incorporate transverse momentum into our model by assuming that the quark-antiquark pairs $q_i\bar{q}_i$ which are produced to discharge the color field conserve transverse momentum in a pairwise fashion and have no net transverse momentum. The quark q_i in the ith pair is assigned a transverse momentum $\boldsymbol{q}_{\perp_i}$ and the antiquark the balancing momentum $-\boldsymbol{q}_{\perp_i}$. The $\boldsymbol{q}_{\perp_i}$ are distributed according to the Gaussian distribution

$$\exp(-q_{\perp_i}^2/2\sigma_q^2)\, d^2q_\perp . \tag{2.46}$$

Except for the first-rank primary meson, all primary mesons are given a transverse momentum that is the vector sum of the transverse momenta of the two quarks which form it. The first-rank primary meson is assigned a transverse momentum given by

$$\boldsymbol{k}_\perp(1) = \boldsymbol{q}_{\perp_1} - \boldsymbol{q}_{\perp_0} , \tag{2.47a}$$

the second-rank primary meson

$$\boldsymbol{k}_\perp(2) = \boldsymbol{q}_{\perp_2} - \boldsymbol{q}_{\perp_1} , \tag{2.47b}$$

and the rth-rank primary meson

$$\boldsymbol{k}_\perp(r) = \boldsymbol{q}_{\perp_r} - \boldsymbol{q}_{\perp_{r-1}} : \tag{2.47c}$$

The initial $\boldsymbol{q}_{\perp_0}$ is also generated according to (2.46). This is to insure that the first primary meson has the same mean square transverse momentum as the others.

The net result is to produce a cascade of primary mesons all of which have the same distribution of transverse momentum at fixed z, a Gaussian with a mean $\langle k_\perp^2\rangle = 2\sigma^2$ and

$$\langle k_\perp\rangle_{\text{primary mesons}} = \sqrt{\tfrac{1}{2}\pi}\sigma , \tag{2.48a}$$

with

$$\sigma = \sqrt{2}\,\sigma_q , \tag{2.48b}$$

since two quarks contribute to each primary meson. This method introduces a correlation between primary mesons of adjacent rank, so that they tend to go oppositely, the mean of $(\boldsymbol{k}_{\perp_1} \cdot \boldsymbol{k}_{\perp_2})$ is $-\sigma^2$. There is no correlation between primary mesons whose rank is not adjacent. The physical ideas are reasonable enough, but the particular choice of $\langle(\boldsymbol{k}_{\perp_1} \cdot \boldsymbol{k}_{\perp_2})\rangle = -\sigma^2$ is a pure guess. It is made so that the total perpendicular momentum of the jet $\boldsymbol{k}_{\perp_{\text{jet}}} = \boldsymbol{k}_{\perp_1} + \boldsymbol{k}_{\perp_2} + \ldots + \boldsymbol{k}_{\perp_N}$ has a mean square which does not rise with N.

In general, we could also contemplate another parameter giving the center of mass of the new pairs a Gaussian distribution. This has the affect of decreasing the negative correlation coefficient $\langle \boldsymbol{k}_{\perp_1} \cdot \boldsymbol{k}_{\perp_2}\rangle / \sqrt{\langle k_{\perp_1}^2\rangle \langle k_{\perp_2}^2\rangle}$ between adjacent-rank

primary mesons from the value $-\frac{1}{2}$ to any value minus or plus up to $+\frac{1}{2}$ depending on parameters. In this case, extra rules would have to be added to keep the sum $k_{\perp_{\text{jet}}}$ within bounds, and rather than risk complications, we do not do this.

Our main aim for adding transverse momentum is to study the differences between the true rapidity Y and $Y_z = -\ln z$ and to investigate the effect of resonance decays on the distributions of transverse momentum. In addition, it will be interesting to see if there is any experimental indication that primary mesons adjacent in rank have negative transverse momentum correlations. Unfortunately, as we will discover in sect. 5, although the correlation is close in hierarchy, it is spread over a wide range in rapidity due to the mixing of rank in hierarchy and order in rapidity, and is further obscured by correlations between mesons coming from the decay of the same primary.

The value of the deviation σ in (2.48a) is chosen so that the finally observed charged pions (those produced directly as primary mesons *plus* those resulting from secondary decays of primary vector mesons) have a mean $k_\perp$ of

$$\langle k_\perp \rangle_{\pi^\pm} = 323 \text{ MeV} . \tag{2.49a}$$

This requires a choice of

$$\sigma = 350 \text{ MeV} \tag{2.49b}$$

and results in

$$\langle k_\perp \rangle_{\text{primary mesons}} = 439 \text{ MeV} . \tag{2.49c}$$

The difference between the mean transverse momentum of the primary mesons and the resulting observed pions is due to resonance decay and the significance of this will be discussed in detail in sect. 5.

2.7.3. Finite-momentum jets

When discussing the Monte Carlo method of generating jets, (i)–(iii) in subsect. 2.7.1, we avoided mentioning at what point one terminates the iterations. The procedure, as outlined, would produce a jet of infinite momentum, an infinite number of particles, and an infinitely long rapidity plateau. For a jet of very large momentum P in the z direction, there is no problem. One merely continues to produce particles until all the available momentum P is used up. For large P, for any z not too small, $p_z = zP$ is large enough that questions such as the difference between energy $E = \sqrt{p_z^2 + m^2 + P_\perp^2}$ and p_z are not important. But what shall we do for real jets of finite momentum P? Even if P is a few GeV so that we can take it equal to the energy, for smaller z (say, $z < 0.2$), one has ambiguities. In the plateau, we know how to resolve such ambiguities. The $dz/z = dp_z/p_z$ behavior in the plateau comes from relativistic phase-space factors multiplying matrix elements which are slowly varying so that the dz/z should be replaced by dp_z/E. But this is exactly $d(E + p_z)/(E + p_z)$. We shall thus resolve all ambiguities by interpreting all variables ξ, W, etc., to refer not to momentum but rather to the quantity $E + p_z = p_z + \sqrt{p_z^2 + m^2 + P_\perp^2}$. The

variable z in $F(z)$ and $f(1-z)$ in eqs. (2.23) and (2.21) with $\eta = 1 - z$ and similarly in all our equations refers to

$$z = (E + p_z)/(E_0 + p_{z_0}) , \tag{2.50}$$

where $E_0 + p_{z_0}$ is for the initial quark. Since the initial quark will always have fairly large momenta, we can replace $E_0 + p_{z_0}$ by $2P_0$, where P_0 is the initial quark momentum. Eq. (2.23) should be reasonable as long as p_z is not too large negative.

For energies large enough that a plateau region is developed, experience with inelastic hadron collisions shows us that scaling and limiting fragmentation hold approximately. This implies that a plot of hadron distributions *versus* rapidity

$$Y = \tfrac{1}{2} \ln((E + p_z)/(E - p_z)) = \ln(E + p_z) - \ln m_\perp , \tag{2.51a}$$

where

$$m_\perp^2 = m^2 + P_\perp^2 , \tag{2.51b}$$

has a simple property. The distribution of particles moving to the right in, say, the c.m.s., are the same for different energy collisions if only the origin of the rapidity Y axis is shifted by Y_0, the rapidity of the original right-moving beam. This is equivalent to the principle that the distribution of particles moving near one end and into the plateau, but not near the other end of the distribution, for one energy can be obtained from that of another energy simply by a Lorentz transformation in the $\hat{z}$ direction. This is because such a Lorentz transformation by a velocity v simply multiplies all $E + p_z$ by a common factor, $\sqrt{(1-v)/(1+v)}$, and does not affect rapidity differences or $E + p_z$ ratios. We shall assume that the same is true of our quark jets.

Jets of finite momentum are produced by first generating "master" jets of very large momentum P_0. Each primary meson has a mass m_i, a perpendicular momentum $\boldsymbol{P}_{\perp_i}$ generated according to subsect. 2.7.2, and an $E_i + p_{z_i}$ determined from

$$E_i + P_{z_i} = z_i(E_0 + P_0) , \tag{2.52}$$

where z_i is generated by the procedures of subsect. 2.7.1. The primary mesons are then allowed to decay according to the rates in the particle tables and jets of finite quark momentum P_q are produced by computing a new or scaled $(E + p_z)_{new}$ by

$$(E + p_z)_{new} = (E + p_z)_{old}(E_q + P_q)/(E_0 + P_0) \tag{2.53a}$$

and keeping all final mesons such that

$$(p_z)_{new} \geqslant 0 . \tag{2.53b}$$

Resonance decays are relatively simple to handle under this scheme. The primary mesons determined according to (2.52) are allowed to decay and the Lorentz transformations of z_i are easy as already mentioned. The sum of the z's of the decay products are equal to the z of the parent. However, occasionally parents with $p_{z_i} \leqslant 0$ will produce a decay product with $p_z \geqslant 0$. We are defining our jets to have only forward-moving particles and thus include such decay products with $p_z \geqslant 0$

Table 3
Mean multiplicity of all particles $\langle N \rangle$, charged particles $\langle N_{ch} \rangle$, positive particles $\langle N_+ \rangle$ and negative particles $\langle N_- \rangle$ for u-quark jets of various energies P_q resulting from the jet model with $a = 0.77$ and $\alpha_{ps} = \alpha_v = 0.5$. Also shown in the "correlation moment" $f_2 = \langle N(N-1) \rangle - \langle N \rangle^2$.

P_q (GeV)	3.5	10	50	500
$\langle N_{tot} \rangle$	4.6	8.2	14.1	22.6
f_2^{tot}	−1.3	2.2	12.9	33.4
$\langle N_{ch} \rangle$	2.7	4.8	8.2	13.0
f_2^{ch}	−0.8	0.2	3.3	9.5
$\langle N_+ \rangle$	1.6	2.7	4.4	6.8
f_2^+	−0.8	−1.1	1.2	0.9
$\langle N_- \rangle$	1.1	2.1	3.8	6.2
f_2^-	−0.4	−0.6	−0.6	−0.3

even though the parent was moving backward. Likewise, we exclude backward-moving secondaries even if the parent was moving forward.

Real quark jets cannot have both the energy and momentum of a free quark of definite mass *. However, it must be remembered that all quark jets occur in pairs. In e^+e^- collisions, two jets result from the $q\bar{q}$ pair produced. In ep reactions, one jet results from the quark "knocked out" of the proton by the neutrino and the other from the "hole" that was left behind. In hadron-hadron collisions at high $p_\perp$, according to the model in FF1, we may think of a quark as being knocked out of the proton by another quark and "generating" a color field as the quark pulls away from the "hole" it left behind. This results in a four-jet structure as discussed in ref. [5] (hereafter called FFF). Energy and momentum are conserved in the two-jet system *not* for the single jet. Hence, any quantity like energy or $k_\perp$ that is not conserved in the single jet is balanced by its oppositely moving partner.

Our quark jets do not precisely conserve energy or momentum. For example, a $P_q = 10$ GeV quark jet produced in the above manner has mean values $\langle E_{tot} \rangle = 9.8$ GeV and $\langle p_{z_{tot}} \rangle = 8.6$ GeV, where E_{tot} and $p_{z_{tot}}$ are the total energy and p_z of all the hadrons in the jet. The mean value of $E_{tot} + p_{z_{tot}}$ does approach $2P_q$ at high momentum. Rather than compensate for the fact that our cut-off procedure yields $E_{tot} + p_{z_{tot}}$ less than $2p_z$ at finite P_q, we can interpret our jets as having an *energy* equal to P_q since this comes out closely the same without any adjustment. For all P_q, the mean value of $E_{tot} - p_{z_{tot}}$ is about 1.2 GeV, and $\langle k_{\perp_{jet}} \rangle$ is about 610 MeV, where $k_{\perp_{jet}}$ is the total perpendicular momentum of all the hadrons in the jet.

Finally, in table 3, we show the particle multiplicities resulting using the above

* If one selects $E + p_z$ of all hadrons in the jet to equal the $E + p_z$ of a large momentum initial quark, then $E - p_z$ is equal to the mean density of particles in the plateau times the $m_\perp$ of the particles. See the discussion on p. 246 of ref. [14].

cut-off prescription. The table gives the mean multiplicity of all particles, charged particles, positive particles, and negative particles for a u-quark with energy $P_q = 3.5$, 10.0, 50.0, and 500.0 GeV, arrived at with $a = 0.77$ and $\alpha_{ps} = \alpha_v = 0.5$.

2.8. The analytic approximation

Although the Monte Carlo method is straightforward, some jet observables require the generation of a large number of typical jets in order to get sufficient statistical accuracy. In addition, it is very handy to have analytic formulas for easy analysis and understanding and so the reader can, without writing a Monte Carlo program, reproduce most of our results. For these reasons, we present approximate analytic solutions for those observables where analytic methods are not too difficult. Details of individual resonance decay are too tedious to handle exactly analytically. What we have done is to assume all products of vector meson decay are distributed as they would be in a two-body isotropic decay of massless products; that is spread uniformly in z from 0 to the z of the parent. Thus for all vector decays, we assume the daughter-parent relationship

$$\hat{G}(z) = 2\int_z^1 G(z/y)\,dy/y\,, \tag{2.54}$$

where $\hat{G}(z)$ is the distribution in z of a daughter of parents distributed in $G(z)$. The integral of (2.54) is easy for $G(z) = F(z)$ or $G(z) = f(1-z)$, since if $G(z) = z^n$ then $\hat{G}(z) = 2(1-z^n)/n$ for $n \neq 0$ and $-2\ln z$ for $n = 0$.

The total distribution of hadrons is

$$F_{\text{total}}(z) = \alpha_{ps}F(z) + \alpha_v\hat{F}(z)\,, \tag{2.55}$$

and the distribution of the first-rank meson is

$$f_{\text{total}}(1-z) = \alpha_{ps}f(1-z) + \alpha_v\hat{f}(1-z)\,, \tag{2.56}$$

where we assign all the decay products the rank in hierarchy of their parent primary meson and where α_{ps} and α_v, given by (2.45), are the relative strengths of the pseudoscalar and vector meson component. The values of $\alpha_{ps}zF(z)$, $\alpha_{ps}f(1-z)$, $\alpha_v z\hat{F}(z)$, and $\alpha_v\hat{f}(1-z)$ are given in fig. 2b with $a = 0.77$. As can be seen in fig. 3, the choice $a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$ produces a good fit to the lepton data on the charged particle z distribution and compares well with the FF1 result.

The distribution of hadrons of flavor h from a quark q is given by

$$D_q^h(z) = \alpha_{ps}[A_q^h f(1-z) + B^h\bar{F}(z)] + \alpha_v[\hat{A}_q^h\hat{f}(1-z) + \hat{B}^h\hat{\bar{F}}(z)]\,, \tag{2.57}$$

where $\bar{F}(z)$ is the probability that the meson at z has any rank greater than one and is defined by (2.29b). (Similarly, $\hat{\bar{F}}(z) = \hat{F}(z) - \hat{f}(1-z)$.) The constants $\hat{A}_q^h$ and $\hat{B}^h$

are given by

$$\hat{A}_q^h = \sum_{i=1}^{9} \beta_{i\to h} A_q^i , \tag{2.58a}$$

$$\hat{B}^h = \sum_{i=1}^{9} \beta_{i\to h} B^i , \tag{2.58b}$$

where the sum is over all nine vector mesons and $\beta_{i\to h}$ is the probability that one finds a daughter of type h as a decay product of a parent of type i. The constants A_q^h and B^h are as in (2.30) and are given in table 1 (for vector production one uses $\theta_M = \theta_v$ and for pseudoscalar production one uses $\theta_M = \theta_{ps}$). The ρ and K^* resonances decay 100% of the time to $\pi\pi$ and $K\pi$, respectively, and the $\beta_{i\to h}$ are easily calculable from isospin symmetry. The values used for $\beta_{i\to h}$ for ω and ϕ are given in table 4, where we approximate three-body decays by the two-body decay formula (2.54) but with $\beta_{i\to h}$ adjusted to give the correct frequency into the various mesons. For example, we take $\omega \to \pi^+\pi^-\pi^0$ and $\omega \to \pi^0\gamma$ with a 90% and 10% rate, respectively, so that $\beta_{\omega\to\pi^\pm} = \frac{1}{3}(0.90)$ and $\beta_{\omega\to\pi^0} = \frac{1}{3}(0.90) + \frac{1}{2}(0.10)$.

In addition, to agree better with the Monte Carlo results, we have also allowed the η and η' to decay. These decays are handled in the same manner (2.54) as the vector meson ω decay but with the $\beta_{i\to h}$ given in table 4.

Fig. 4 shows a comparison of the analytic approximation and the more precise Monte Carlo method. Agreement is good except for the pions resulting from the ω^0 decay (similarly for three body η and η' decay modes) which have been approximated by the two-body decay formula. The approximation for $K^{*0} \to K^+$ does not agree as well as $\rho^0 \to \pi^+$ due to the heavier mass of the K.

The inclusion of resonance decays into the analytic formula for the double-decay distribution $P_q^{a\bar{b},c\bar{d}}(z_1, z_2)$ given by (2.42) and (2.43) is straightforward. One

Table 4
Probability $\beta_{i\to h}$ to find a daughter meson of type h among all the decay products of a parent meson of type i

Daughter \ Parent	η	η'	ω	ϕ
π^+	0.09	0.27	**0.30**	**0.06**
π^0	0.40	0.24	**0.35**	**0.06**
π^-	0.09	0.27	**0.30**	**0.06**
K^+	0	0	**0**	**0.24**
K^0	**0**	0	**0**	**0.17**
K^-	**0**	0	**0**	**0.24**
$\bar{K}^0$	**0**	0	**0**	**0.17**

For $i = \rho$ and K^*, $\beta_{i\to h}$ is given by isospin symmetry assuming $\rho \to \pi\pi$ and $K^* \to K\pi$ 100% of the time.

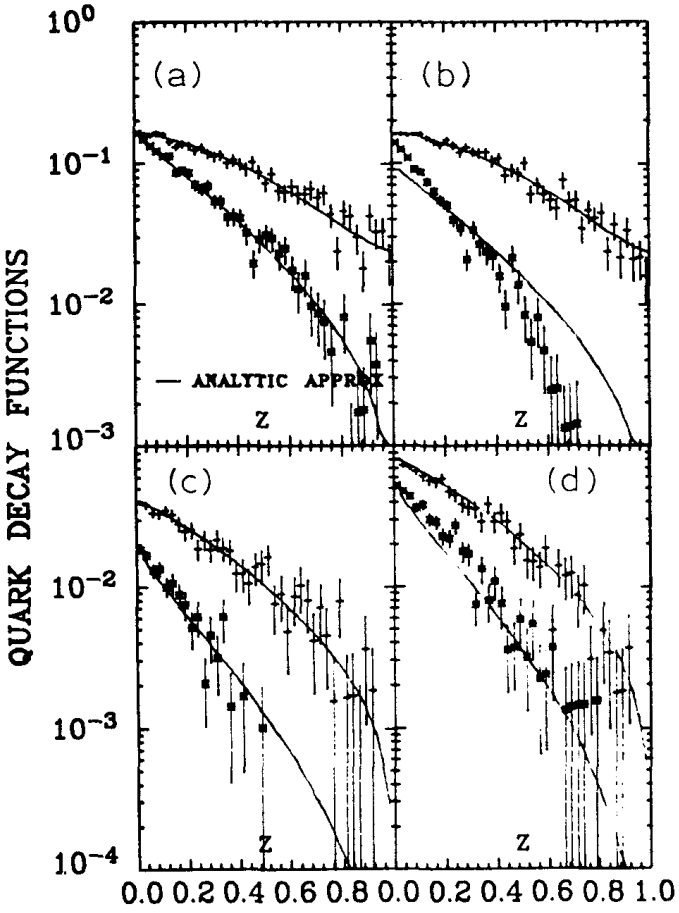


Fig. 4. Comparison of the Monte Carlo results ($a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$) for some of the quark-decay functions, $D_q^h(z)$, with the analytic approximation (solid curve). The number of ρ^0, ω^0, ϕ, and K^{*0} from a u-quark are shown together with the number of π^+ that have decayed from ρ^0 and ω^0 and the number K^+ from ϕ and K^{*0}. (a) + $zD(u \to \rho^0)$; × $zD(u \to \rho^0 \to \pi^+)$; (b) + $zD(u \to \omega^0)$; × $zD(u \to \omega^0 \to \pi^+)$; (c) + $zD(u \to \phi)$; × $zD(u \to \phi \to K^+)$; (d) + $zD(u \to K^{*0})$; × $zD(u \to K^{+0} \to K^+)$.

merely replaces the $F(z)$ and $f(1-z)$ functions in (2.43) by the $F_{\text{total}}(z)$ and $f_{\text{total}}(1-z)$ functions in (2.55) and (2.56). From this $D_q^{h_1 h_2}(z_1, z_2)$ is computed in a manner analogous to (2.57). In addition, however, one must add the contribution from the case where both the mesons at z_1 and z_2 came from the decay of the same resonance (type h_v) given by

$$B_{h_v \to h_1, h_2} D_q^{h_v}(z_1 + z_2)/(z_1 + z_2)\,, \tag{2.59}$$

where $D_q^{h_v}(z)$ is the single-particle distribution of the parent h_v and $B_{h_v \to h_1, h_2}$ in its branching ratio into $h_1 = a\bar{b}$ and $h_2 = c\bar{d}$.

Table 5
Description of the parameters used in the model of quark jets and an explanation of how they are determined

Parameter	Description	Determination
a	Parameter used to describe the function $f(\eta) = 1 - a + 3a\eta^2$, which is the probability that the first-rank primary meson leaves a fraction of momentum η to the remaining cascade.	Require that resulting charged distribution $D_q^{h^+}(z)$ fit the data shown in fig. 3.
γ	The probability of producing a $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ pair is given by γ, γ, and $(1-2\gamma)$, respectively.	Use the fact that at large $p_\perp$ in pp collisions $K^+/\pi^+ = \frac{1}{2}$. Requiring $D_q^{K^+}(z)/D_q^{\pi^+}(z) = \frac{1}{2}$ as $z \to 1$ then implies $\gamma = 0.4$.
α_{ps}, α_v	Relative probability of producing a pseudoscalar meson or a vector meson $(\alpha_{ps} + \alpha_v = 1)$.	Experiments indicate that $\rho^0/\pi^0 \simeq 1$ at large $p_\perp$ in pp collisions so we choose $\alpha_{ps} = \alpha_v = 0.5$.
σ	Determines the $\langle k_\perp \rangle$ of the produced primary mesons.	Require $\langle k_\perp \rangle_\pi$ near 330. Means $\langle k_\perp \rangle_{primary} = 439 = \sqrt{\frac{1}{2}\pi}\,\sigma$.
θ_{ps}, θ_v	Pseudoscalar and vector nonet mixing angles where $\eta(\omega) = S\cos\theta + N\sin\theta$, $\eta'(\phi) = -S\sin\theta + N\cos\theta$, with $S = s\bar{s}$, $N = \sqrt{\frac{1}{2}}(u\bar{u} + d\bar{d})$.	Use $\theta_{ps} = 45°$ and $\theta_v = 90°$ as simple numbers close to experiment.
$\beta_{i\to h}$	Probability to find a daughter meson of type h among all the decay products of a parent vector meson of type i.	Use branching ratios in particle tables.

Our model of quark jets is thus completely determined by the function $f(\eta)$, which we have parametrized by a parabola with one parameter, a, and the three additional parameters; the degree that SU(3) is broken in the formation of new $q\bar{q}$ pairs (γ in (2.24)), the spin-parity nature of the primary mesons (α_{ps}, α_v in (2.45)) and the mean transverse momentum of the primary mesons (σ in (2.48a)). From this, all the properties of quark jets follow. The parameters of the model together with a review of how they were determined is given in table 5. We now proceed to examine the results of the model in detail.

3. Properties of the "end" of the quark jet

3.1. Single-particle distribution $D_q^h(z)$

Figs. 5, 6, and 7 show the single-particle quark-decay functions, $D_q^h(z)$, resulting from our new jet model with $a = 0.77$ and $\alpha_{ps} = \alpha_v = 0.5$ together with the analytic

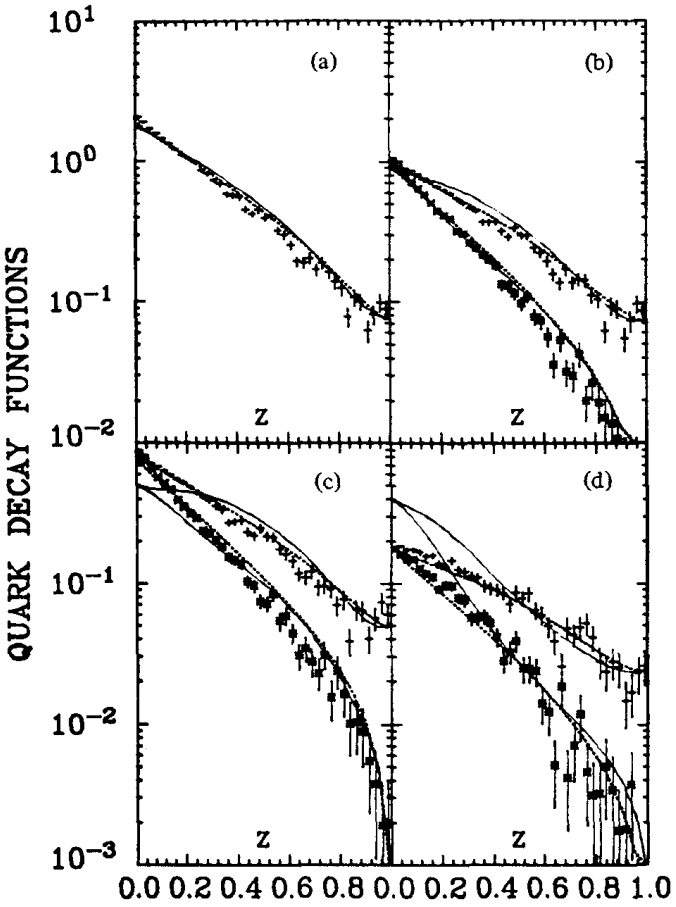


Fig. 5. The quark-decay functions (a) $zD(u \to h^+ + h^-, z)$; (b) $zD(u \to h^\pm, z)$, + $zD(u \to h^+)$, × $zD(u - h^-)$; (c) $zD(u \to \pi^\pm, z)$, + $zD(u \to \pi^+)$, × $zD(u \to \pi^-)$; and (d) $zD(u \to K^\pm, z)$, + $zD(u \to K^+)$, × $zD(u \to K^-)$, resulting from the quark-jet model with $a = 0.77$ and $\alpha_{ps} = \alpha_v = 0.5$ (points). Also shown is the analytic approximation (dotted curve) and the results obtained in FF1 (solid curves).

approximation (dotted curve). For comparison, the results arrived at in FF1 are also displayed (solid curves). The differences between the analytic calculations and the Monte Carlo results are because in the analytic approximation, we have considered all decays as massless two-body decays. For instance, as seen in fig. 4 the $\omega \to 3\pi$ decay is not treated correctly in the analytic approximation. Nevertheless, the two methods agree quite closely.

The decay functions in figs. 5, 6, and 7 agree well for the most part with the results from FF1. There are, however, the following notable differences.

(i) The new decay functions have

$$D_u^{K^\pm}(z)/D_u^{\pi^\pm}(z) \underset{z\to 0}{\to} 0.2\,, \tag{3.1}$$

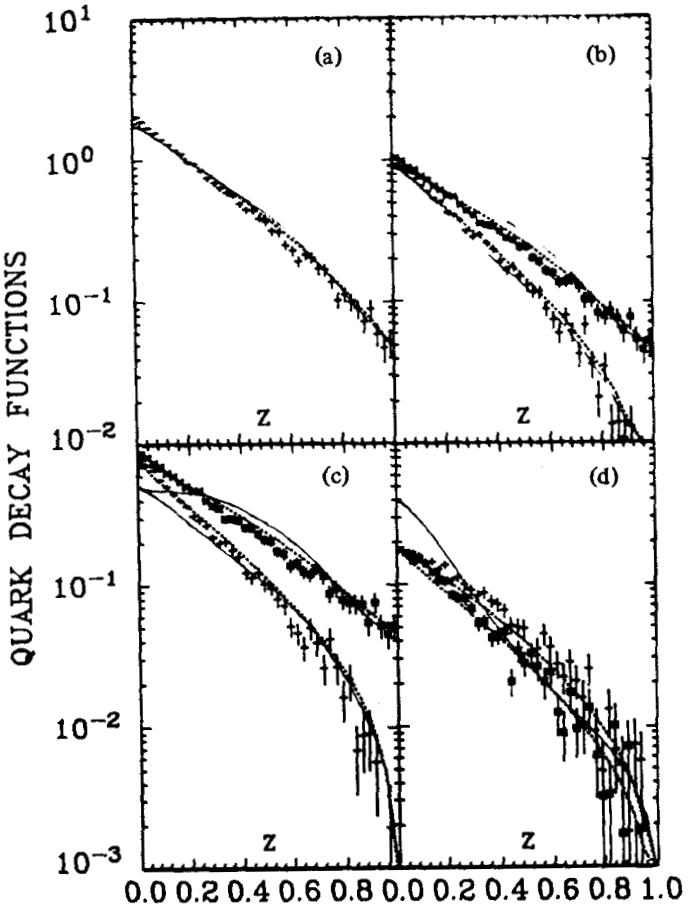


Fig. 6. Same as fig. 5 but for a d-quark.

whereas the FF1 results yield a value of 0.8. Both have

$$D_u^{K^\pm}(z)/D_u^{\pi^\pm}(z) \underset{z\to 1}{\to} 0.5, \tag{3.2}$$

by construction, since $\gamma = 0.4$. The difference at small z is due to the many K^* resonances that decay into kaons *and* pions. The new results are probably more reasonable than the FF1 values.

(ii) The new approach yields

$$D_d^{K^+}(z) > D_d^{K^-}(z), \tag{3.3}$$

whereas in FF1 we assumed equality for simplicity. However, as discussed in FF1, (3.3) is expected.

Table 6 gives the total fraction of momentum carried by the direct (primary)

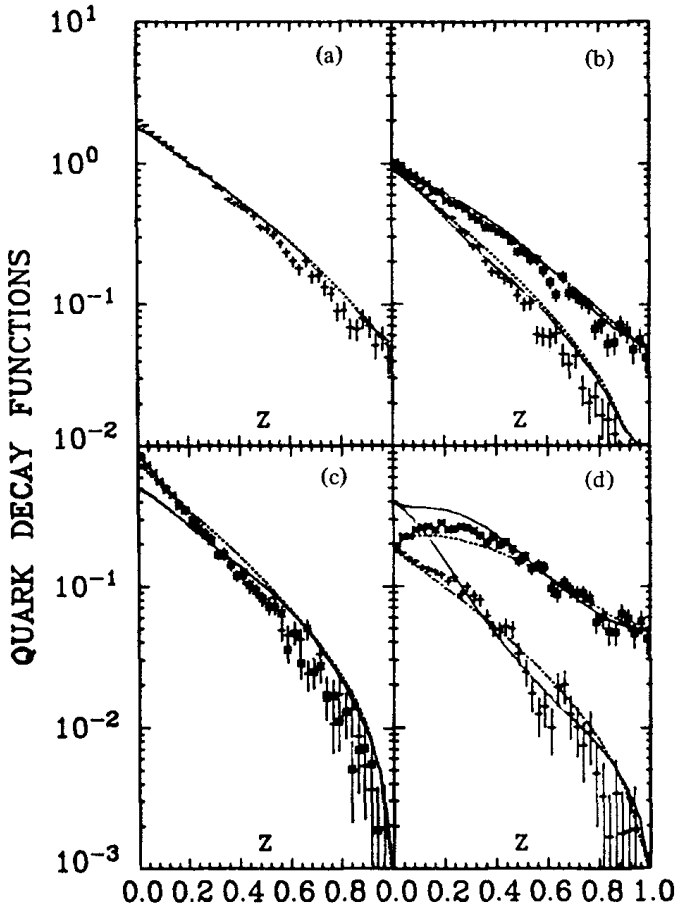


Fig. 7. Same as fig. 5 but for an s-quark.

mesons and by the final hadrons (direct + decay) for a u-, d-, and s-quark. These results can be compared to the FF1 values shown in table 2.

In spite of the differences at small z between our new results and the FF1 results, they are not very different for large z ($z \geqslant 0.6$). Since the large-$p_\perp$ single-particle predictions made in FF1 and most of the two-particle correlation results obtained in FFF (ref. [5]) are only sensitive to the large-z region of the quark decay functions, most of the predictions made in FF1 and FFF remain unchanged. Changes and improvements of the results of FF1 and FFF will be discussed in sect. 6.

3.2. Approach to the rapidity plateau

Figs. 8, 9, and 10 show the number of hadrons, charged hadrons, positive hadrons and negative hadrons per 0.1 unit of Y_z for a u-, d- and s-quark, respectively. We

Table 6
Fraction of total momentum carried by the direct-primary (before decay) mesons and the direct-plus-indirect (from a decay) mesons resulting from a u-, d- and s-quark in our jet model with $a = 0.77$ and $\alpha_{ps} = \alpha_v = 0.5$

	Particle	u	d	s
Direct	$\pi^+ = \rho^+$	0.12	0.06	0.06
	$\pi^0 = \rho^0$	0.09	0.09	0.06
	$\pi^- = \rho^-$	0.06	0.12	0.06
	$K^+ = K^{*+}$	0.06	0.03	0.03
	$K^0 = K^{*0}$	0.03	0.06	0.03
	$K^- = K^{*-}$	0.03	0.03	0.09
	$\bar{K}^0 = \bar{K}^{*0}$	0.03	0.03	0.09
	η	0.05	0.05	0.05
	η'	0.05	0.05	0.05
	ω	0.09	0.09	0.06
	ϕ	0.01	0.01	0.04
total = direct + indirect	π^+	0.29	0.19	0.19
	π^0	0.26	0.26	0.20
	π^-	0.19	0.29	0.19
	K^+	0.08	0.06	0.06
	K^0	0.06	0.08	0.06
	K^-	0.04	0.04	0.13
	$\bar{K}^0$	0.04	0.04	0.13
	γ	0.04	0.04	0.04
total	π	0.74	0.74	0.58
total	K	0.22	0.22	0.38

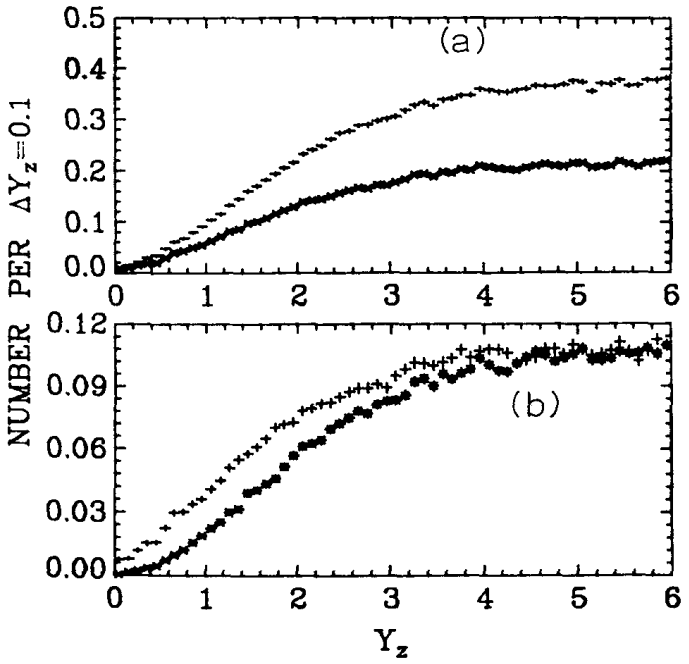


Fig. 8. The number of particles, charged particles, positive particles and negative particles per $\Delta Y_z = 0.1$ resulting from a u-quark jet with $a = 0.77$ and $\alpha_{ps} = \alpha_v = 0.5$, where $Y_z = -\ln z$. (a) + all particles; x charged particles; (b) + positive particles, x negative particles.

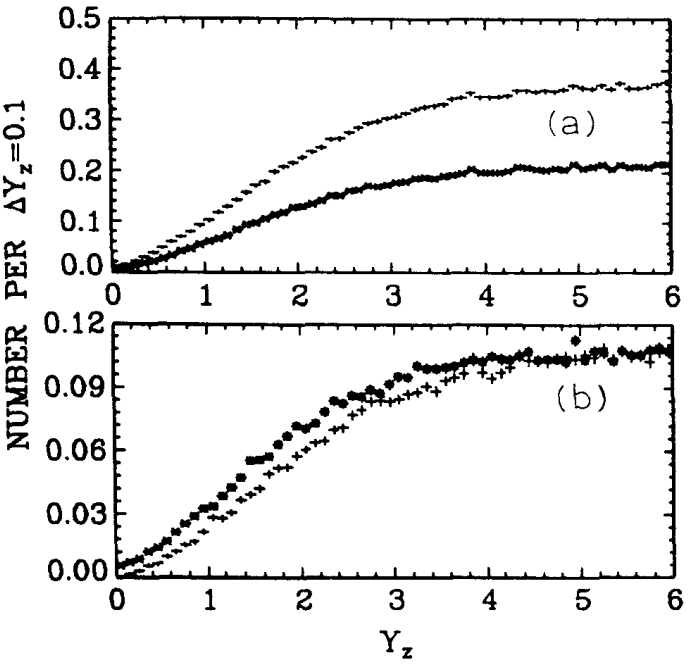


Fig. 9. Same as fig. 8 but for a d-quark jet.

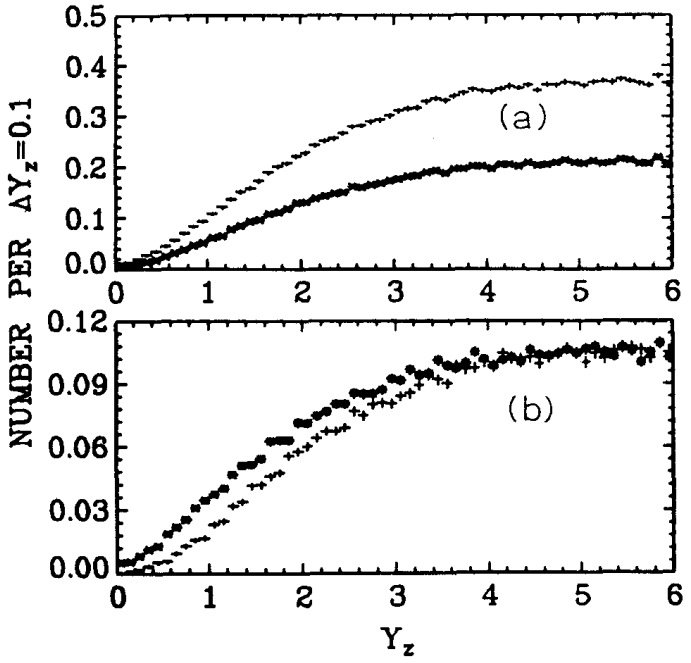


Fig. 10. Same as fig. 8 but for an s-quark jet.

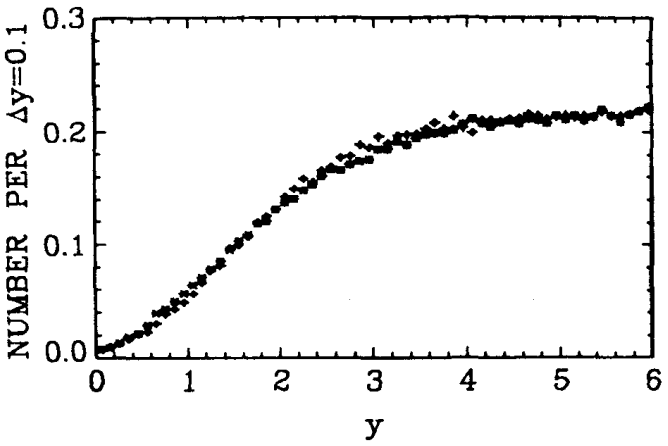


Fig. 11. Comparison of the number of charged particles per $\Delta Y_z = 0.1$ *versus* Y_z with the number of charged particles per $\Delta Y = 0.1$ *versus* $Y_{end} - Y$ for a u-quark jet, where $Y_z = -\ln z$ and Y is the true rapidity (2.51) and Y_{end} is given by (3.6b) with $m_{\perp_0} = 380$ MeV. $a = 0.77$, α_{ps} $\alpha_v = 0.5$, $\diamond\, y = Y_{end} - Y$, $\times\, y = Y_z$.

have defined the "z rapidity", Y_z, by

$$Y_z = -\ln z = -\ln((E + p_z)/2P_q) \,. \tag{3.4}$$

This is a convenient variable for us since it does not depend on the generated value of the perpendicular mass, $m_\perp$, in equation (2.51b). The usual rapidity, Y, is given by (2.51a) and hence

$$Y = -Y_z + \ln(2P_q/m_\perp) \,, \tag{3.5}$$

so that particles at the same Y_z may be found spread a bit in Y because of differences of mass and transverse momenta. It is usual to measure Y from some Y_{end} at the high-momentum end of the jet. For this one might expect to take $\ln(2P_q/m_q)$, where m_q is the mass of the original quark, but this has no clear meaning. Alternatively, data could be plotted using

$$Y_{max} = \ln(2P_q/m_\pi) \tag{3.6a}$$

for Y_{end} since this is the maximum rapidity a pion of the jet could have. These differences are, of course, only shifts of scale of Y for convenience in plotting. In fig. 11, we compare the $Y_{end} - Y$ and Y_z distributions, where we have chosen

$$Y_{end} = Y_{max} - \ln(m_{\perp_0}/m_\pi) \,, \tag{3.6b}$$

with $m_{\perp_0}$ an arbitrary number chosen in an attempt to make the two distributions agree. The distributions are nearly identical with $m_{\perp_0} = 380$ MeV, the $Y_{end} - Y$ distribution being slightly steeper. (This value is close to the mean $m_\perp$.)

The model yields about 2.2 charged particles per unit rapidity in the plateau with equal numbers of positive and negative particles. (Some multiplicities are given in

table 3.) Towards the "end" of the quark jet ($Y_z \lesssim 3$), on the other hand, the number of positive and negative hadrons is not the same and one can hope to discriminate among the quark flavors by observations in this region of Y_z.

3.3. Distribution of charge

Any single-quark jet has an integral number of hadrons and an integral charge. The average total charge of a u-, d- and s-quark jet of infinite momentum is nearly the charge of the quark (for $\gamma = 0.4$ we have 0.6, –0.4, and –0.4 as shown in (2.39)). Table 7 shows the number of times various total jet charges Q occur for 10 000 jets with energy $P_q = 10$ GeV. Jets initiated by u-quarks are more likely to have positive total charge than those initiated by d-quarks. For example, at $P_q = 10$ GeV, $Q = +2$ jets occur $\simeq 4$ times more often for u-jets than for d-jets and $Q = +4$ jets occur

Table 7
Number of times various total quark-jet charge Q occurred for 10 000 quark jets with energy $P_q = 10$ GeV

Charge Q	u	d	s
–4	7	39	42
–3	42	231	227
–2	307	1094	1115
–1	1291	2972	3009
0	3118	3675	3661
1	3460	1576	1483
2	1398	340	390
3	299	62	60
4	64	6	7
$\langle Q \rangle$ [a]	0.54	–0.36	–0.36

Same as above but observe only those hadrons with $z \geq 0.1$

Charge Q	u	d	s
–4	2	0	6
–3	20	88	74
–2	258	710	781
–1	1442	2969	3172
0	3560	4123	4091
1	3639	1742	1606
2	967	342	255
3	106	23	15
4	6	3	0
$\langle Q \rangle$	0.39	–0.21	–0.28

[a] For infinite momentum quark jets, the mean charge, $\langle Q \rangle$, is 0.60, –0.40, –0.40 for a u-, d- and s-jet, respectively.

about 10 times more frequently. The average total charge, $\langle Q \rangle$, for a P_q = 10 GeV u- and d-jet is 0.54 and −0.36, respectively. Because of the finite total momentum available (P_q = 10 GeV) summing only $z > 0$ does not yet yield the full total mean charge 0.6 and −0.4 expected. There is some overlap of charge into negative z. An interesting point is that the mean total jet charge $\langle Q \rangle$ is *not* precisely the same for each fixed multiplicity. As table 8a shows, the odd charged multiplicity jets (N_{ch} = 1, 3, etc.) have a mean charge $\langle Q \rangle$ that is somewhat larger than the jets of even charge multiplicity (N_{ch} = 2, 4, etc.). For an extreme example, for u-quark jets having only one charged particle (N_{ch} = 1), that particle is positive about 13 times more often than it is negative ($\langle Q \rangle$ = 0.86).

Unfortunately, total jet charge is often a difficult quantity to measure experimen-

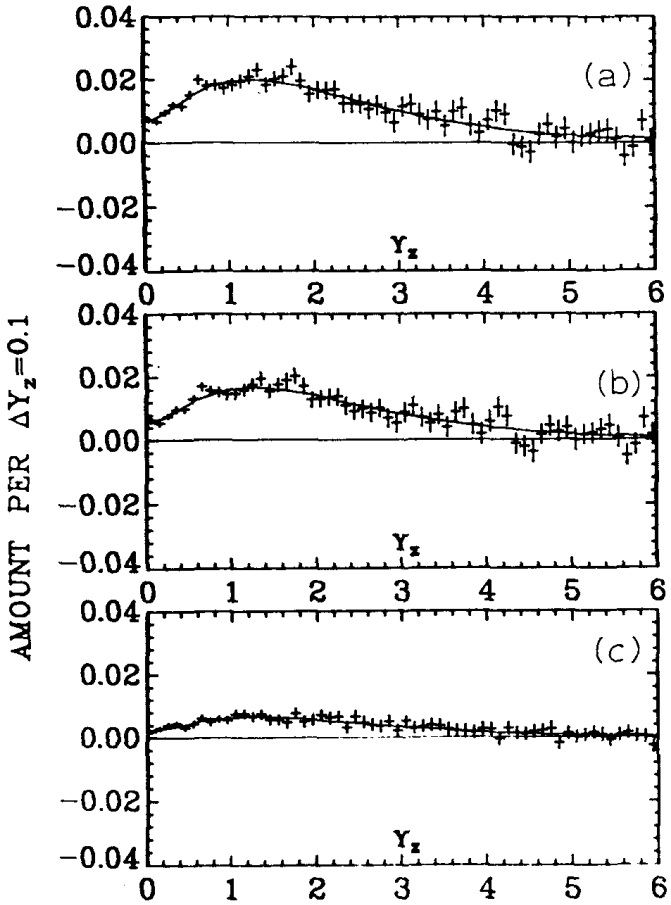


Fig. 12. Distribution of (a) charge Q, total Q = 0.6, (b) third component of isospin I_3, total I_3 = 0.5, and (c) strangeness S, total S = 0.2, along the Y_z ($Y_z = -\ln z$) axis for a u-quark jet. Also shown are the analytic results (solid curves). a = 0.77, $\alpha_{ps} = \alpha_v$ = 0.5.

Table 8a
Number of times various total quark-jet charge Q occurred for 10 000 jets with momentum P_q = 10 GeV, where N_{ch} is the charged multiplicity

		u	d	s
all jets	$Q > 0$	5232	1984	1941
	$Q < 0$	1650	4341	4398
	$Q = 0$	3118	3675	3661
	$\langle Q \rangle$ [a]	0.54	−0.36	−0.36
$N_{ch} = 1$ jets	$Q > 0$	420	102	81
	$Q < 0$	31	287	341
	$\langle Q \rangle$	0.86	−0.48	−0.62
$N_{ch} = 2$ jets	$Q > 0$	175	25	37
	$Q < 0$	13	118	147
	$Q = 0$	591	799	839
	$\langle Q \rangle$	0.42	−0.20	−0.22
$N_{ch} = 3$ jets	$Q > 0$	1029	423	427
	$Q < 0$	266	883	937
	$\langle Q \rangle$	0.64	−0.38	−0.41
N_{ch} = even jets ($N_{ch} \neq 0$)	$Q > 0$	1462	346	397
	$Q < 0$	314	1133	1157
	$Q = 0$	3052	3535	3512
	$\langle Q \rangle$	0.50	−0.33	−0.31
N_{ch} = odd jets	$Q > 0$	3770	1638	1543
	$Q < 0$	1336	3208	3241
	$\langle Q \rangle$	0.58	−0.40	−0.43

[a] For infinite-momentum jets $\langle Q \rangle$ = 0.60, −0.40 and −0.40 for a u-, d- and s-quark, respectively.

tally. One can never be sure that all the jet particles have been included and that one has not included extra background particles. A more tractable quantity experimentally is the total charge of all those hadrons in a jet whose z value exceeds some threshold. As shown in tables 7 and 8b, charges defined for P_q = 10 GeV and $z \geq 0.1$ have the same characteristics as the complete jet but the magnitudes are smaller. The average total charges ($z \geq 0.1$) are 0.39 and −0.21 for a u- and d-jet, respectively.

Figs. 12, 13, and 14 show the distribution of charge Q, third component of isospin I_3, and strangeness S (of course, $S = 2(Q - I_3)$) along the Y_z axis for a u-, d- and s-quark, respectively. As discussed in subsect. 2.5 and given by (2.37), these quantities are related to the distribution of the first-rank primary meson, $f(1-z)$, and the integrated values are given by (2.39), (2.40), and (2.41) (with $S = Y$ since we have no baryons among the resulting hadrons). Also shown in these figures is the analytic approximation. As can be seen, the charge, isospin, and strangeness are distributed

Table 8b
Same as table 8a except only hadrons with $z \geq 0.1$ are observed

		u	d	s
all jets	$Q > 0$	4718	2110	1876
	$Q < 0$	1722	3767	4033
	$Q = 0$	3560	4123	4091
	$\langle Q \rangle$	0.39	−0.21	−0.28
$N_{ch} = 1$ **jets**	$Q > 0$	2624	1205	1090
	$Q < 0$	936	2120	2272
	$\langle Q \rangle$	0.47	−0.28	−0.35
$N_{ch} = 2$ jets	$Q > 0$	871	309	240
	$Q < 0$	229	651	713
	$Q = 0$	2383	2557	2402
	$\langle Q \rangle$	0.37	−0.19	−0.28
$N_{ch} = 3$ jets	$Q > 0$	1101	551	522
	$Q < 0$	512	918	949
	$\langle Q \rangle$	0.47	−0.33	−0.37
N_{ch} = even jets ($N_{ch} \neq 0$)	$Q > 0$	973	345	255
	$Q < 0$	260	710	787
	$Q = 0$	2607	2792	2643
	$\langle Q \rangle$	0.37	−0.19	−0.29
N_{ch} = odd **jets**	$Q > 0$	3745	1765	1621
	$Q < 0$	1462	3057	3246
	$\langle Q \rangle$	0.47	−0.29	−0.36

over a considerable range in Y_z. The quark quantum numbers are spread over almost 4 units of Y_z, which will make it difficult to determine them experimentally.

One of the most important experimental questions about high-$p_\perp$ hadron collisions is whether the jets really come from quark cascades as we have supposed in FF1 and FFF, or possibly from other types of objects such as gluons or di-quarks. We have calculated the flavor of the quarks to be expected under various circumstances (see fig. 25 of FFF). When the characteristics of the quark jets of definite flavor in lepton experiments are known, the details can be checked against the jets observed in hadron experiments. But what should we measure to most readily identify the flavor of a quark jet: the total charge, the charge of the fastest hadron, of the fastest two? We have used our model for a "standard" jet as a kind of laboratory of typical jets to test various ideas. The following sections on the flavor properties of the jets, therefore, contain very detailed information from these studies on various quantities which have been selected because they are easy to measure experimentally or because they are expected to differ as much as possible for u- and d-flavor jets.

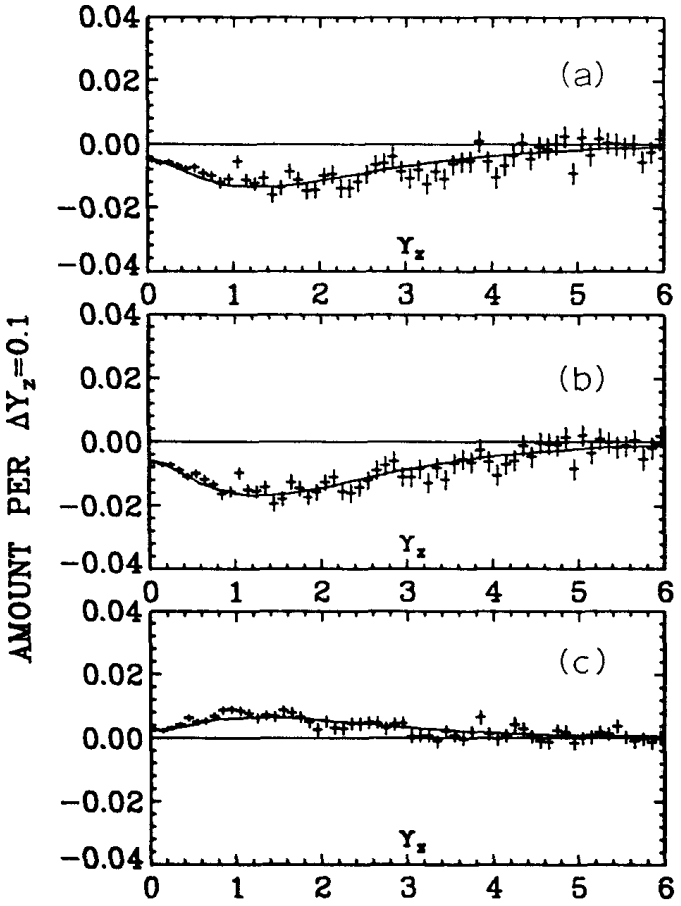


Fig. 13. Same as fig. 12 but for a d-quark jet. (a) Total $Q = -0.4$, (b) total $I_3 = -0.5$, (c) total $S = 0.2$.

3.4. Fastest-particle analysis

It is interesting to ask about the fraction of momentum, z, carried by the fastest hadron (largest z) when a quark fragments into hadrons in an unbiased fashion. Figs. 15 and 16 show the z distributions of various hadrons fragmenting from a u-quark of energy 10 GeV resulting from the model where, by first and second, we mean fastest (largest z) and second-fastest (next-largest z). These figures and table 9 show that for a u-quark jet, the fastest hadron carries on the average only 39% of the quark momentum and the fastest charged hadron only 30%. Fig. 17 and table 10 show that for a u-quark, the fastest charged hadron carries 54% of the total charged momentum with the second fastest charged particle taking about 21%. (All charged particles carry on the average about 57% of the u-quark momentum.) The fastest two

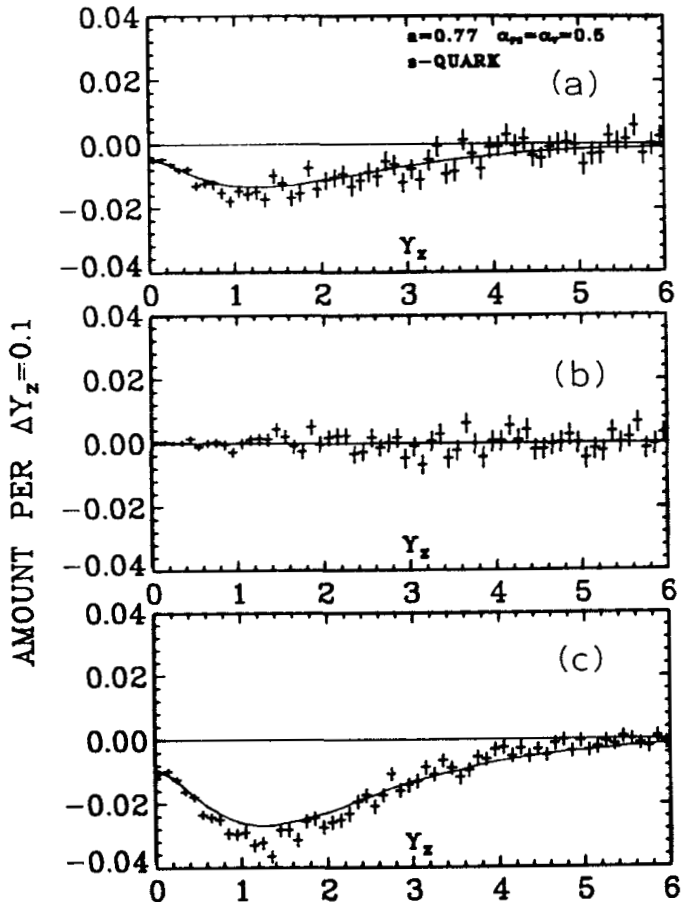


Fig. 14. Same as fig. 12 but for an s-quark jet. (a) Total $Q = -0.4$, (b) total $I_3 = 0.0$, (c) total $S = -0.8$.

Table 9
Mean values of z for hadrons fragmenting from a u-, d- and s-quark of energy $P_q = 10$ GeV

	u	d	s
fastest hadron	0.39	0.39	0.38
2nd-fastest hadron	0.18	0.18	0.19
3rd-fastest hadron	0.11	0.11	0.11
fastest charged	0.30	0.28	0.27
2nd-fastest charged	0.12	0.12	0.12
3rd-fastest charged	0.06	0.06	0.06
all charged	0.57	0.54	0.52
fastest positive	0.25	0.17	0.15
2nd-fastest positive	0.07	0.04	0.04
3rd-fastest positive	0.02	0.01	0.01
fastest negative	0.15	0.22	0.22
2nd-fastest negative	0.04	0.06	0.06
3rd-fastest negative	0.01	0.02	0.02

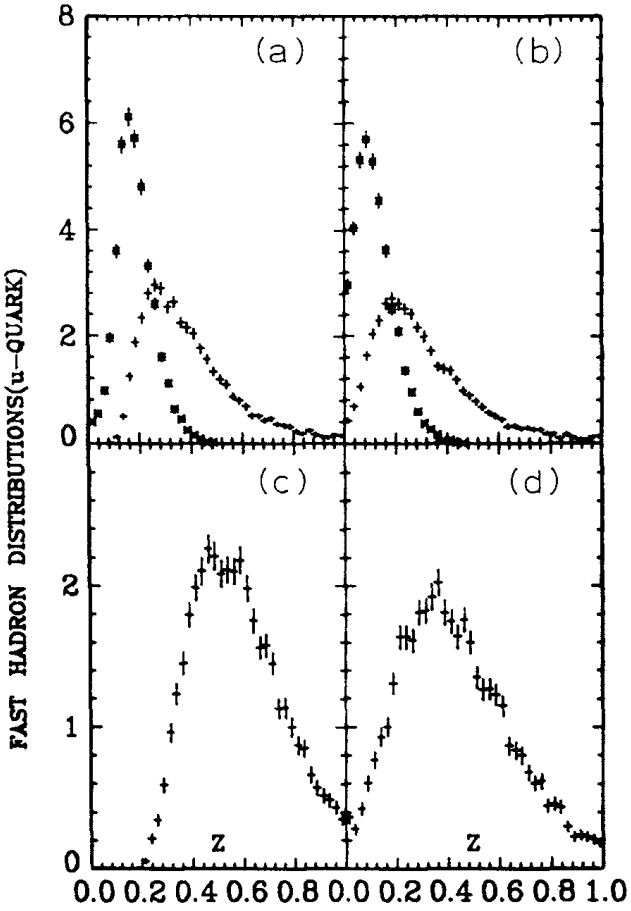


Fig. 15. The predicted z distributions of various hadrons in a u-quark jet with P_q = 10 GeV, where first and second refer to the *fastest* (largest-z) and second *fastest* (next largest-z), *not* to rank in hierarchy. a = 0.77, $\alpha_{ps} = \alpha_v$ = 0.5. (a) + First particle ($\langle z \rangle$ = 0.39), × second particle ($\langle z \rangle$ = 0.18); (b) + first charged ($\langle z \rangle$ = 0.30), × second charged ($\langle z \rangle$ = 0.12), (c) first two particles ($\langle z \rangle$ = 0.57), (d) first two charged ($\langle z \rangle$ = 0.42).

Table 10
Mean values of z_R for hadrons fragmenting from a u-, d- and s-quark of energy P_q = 10 GeV, where $z_R = z/z_{ch}$ and z_{ch} is the sum of the z of all charged hadrons

	u	d	s
Fraction of jets with at least one charged particle	0.993	0.986	0.985
fastest charged	0.54	0.53	0.53
2nd-fastest charged	0.21	0.22	0.22
Fraction of jets with at least one positive particle	0.988	0.943	0.932
fastest positive	0.45	0.31	0.30
2nd-fastest positive	0.12	0.08	0.08
Fraction of jets with at least one negative particle	0.929	0.973	0.972
fastest negative	0.27	0.41	0.42
2nd-fastest negative	0.07	0.11	0.11

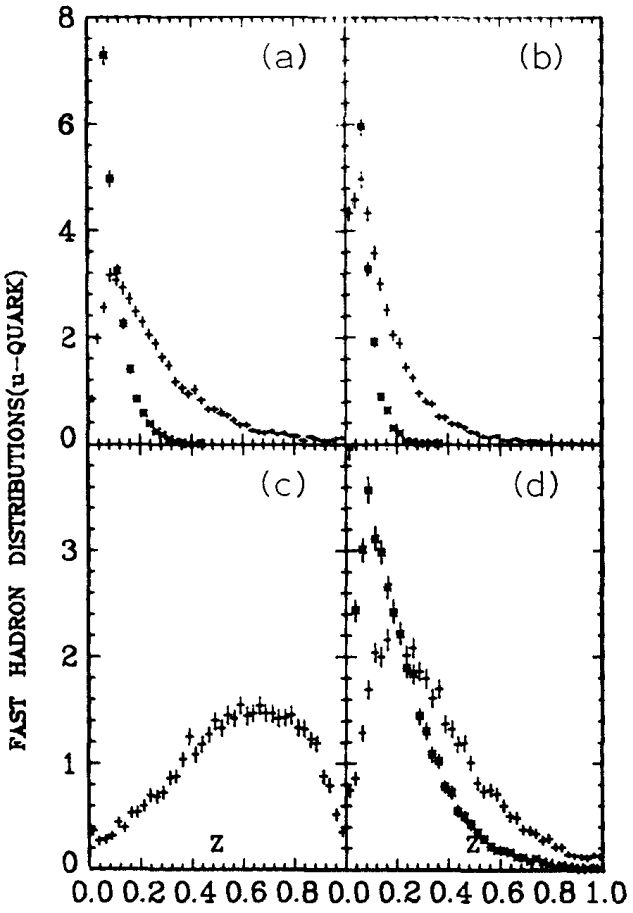


Fig. 16. Same as fig. 15, continued. (a) + First positive ($\langle z \rangle = 0.25$), × second positive ($\langle z \rangle = 0.07$); (b) + First negative ($\langle z \rangle = 0.15$), × second negative ($\langle z \rangle = 0.04$), (c) all charged ($\langle z \rangle = 0.57$), (d) + First two positive ($\langle z \rangle = 0.32$), × first two negative ($\langle z \rangle = 0.19$).

charged hadrons in a u-quark jet carry about 75% of the total charged momentum. Roughly, the fastest charged hadron takes, on the average, half the charged momentum; the second fastest charged hadron takes half the remaining charged momentum, etc. This feature of the model is in agreement with experimental observations of jets produced in νp interactions [15] as seen in fig. 19.

We might have expected, from the point of view from which we began, that the primary meson containing the original quark would be likely to have a large fraction of that quark's momentum. Yet the function $f(1-z)$ in eq. (2.21) shows this primary meson to be more likely to have less than half of the initial quark's momentum. In fact, from (2.36) the average is less than a third of the jet momentum (for $a = 0.77$). Because of this, there is a considerable difference between rank in hierarchy and order

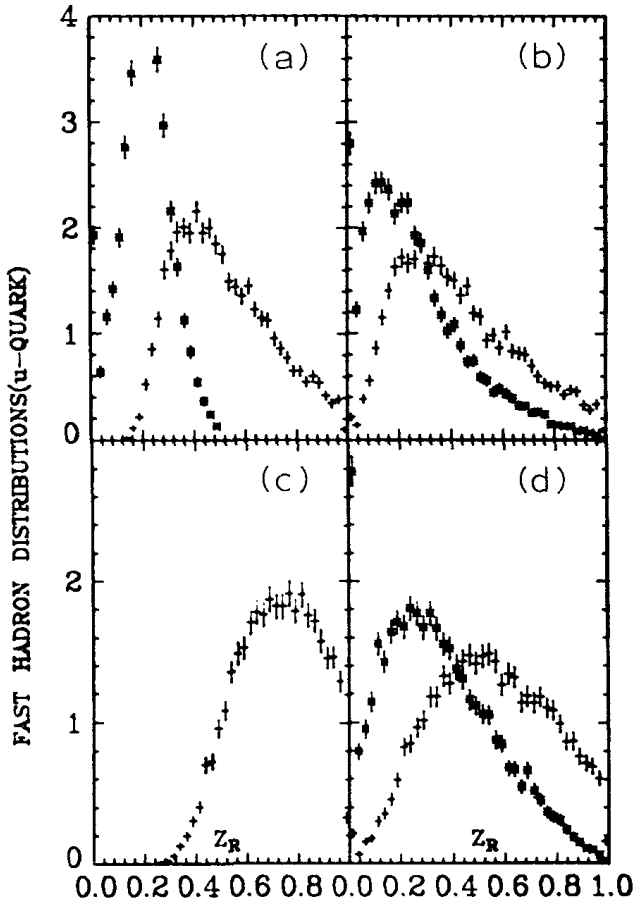


Fig. 17. The predicted z_R distributions of various hadrons in a u-quark jet with P_q = 10 GeV, where z_R is the relative fraction of charged momentum defined by $z_R = z/z_{ch}$ where z_{ch} is the sum of the z's of all charged particles in the jet. As in figs. 15 and 16, first and second refer to order in z *not* to rank in hierarchy. a = 0.77, $\alpha_{ps} = \alpha_v$ = 0.5. (a) + first charged ($\langle z_R \rangle$ = 0.54), X second charged ($\langle z_R \rangle$ = 0.21), (b) first positive ($\langle z_R \rangle$ = 0.45), X first negative ($\langle z_R \rangle$ = 0.27), (c) first two charged ($\langle z_R \rangle$ = 0.75), (d) + first two positive ($\langle z_R \rangle$ = 0.57), X first two negative ($\langle z_R \rangle$ = 0.34).

Table 11
Number of times the fastest (largest-z) hadron occurred with rank r (r = 1, ..., 5) for 10 000 u-quark jets

Rank r	Unbiased	$z \geq 0.5$	$z \geq 0.75$
1	4267	1345	340
2	2566	564	97
3	1506	197	23
4	795	56	3
5	436	17	0
Mean rank	2.3	1.7	1.4

All the decay products of a particular primary meson are assigned the rank of that meson.

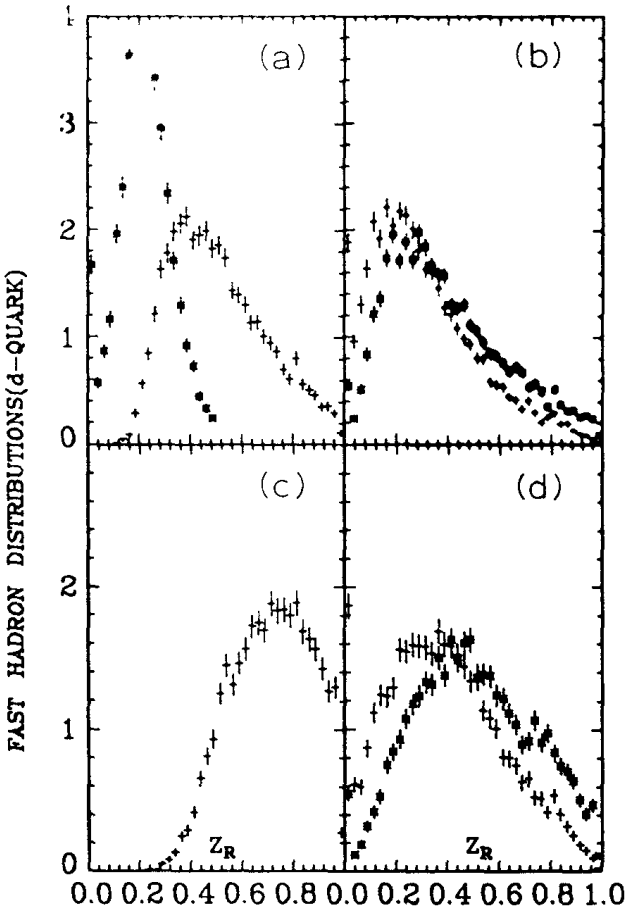


Fig. 18. Same as fig. 17 but for a d-quark jet. (a) + first charged ($\langle z_R \rangle = 0.53$), × second charged ($\langle z_R \rangle = 0.22$), (b) + first positive ($\langle z_R \rangle = 0.31$), × first negative ($\langle z_R \rangle = 0.41$), (c) first two charged ($\langle z_R \rangle = 0.75$), (d) + first two positive ($\langle z_R \rangle = 0.39$), × first two negative ($\langle z_R \rangle = 0.52$).

in rapidity (or z). It is relatively easy for the first-rank primary meson to leave the remaining cascade with a large momentum and thus to end up with a momentum less than some subsequent (higher-rank) meson. Table 11 shows that the fastest hadron from an unbiased u-quark jet is the first-rank meson only 43% of the time. However, as one selects on events where the z of the fastest hadron is large, it becomes more likely that it is of rank-one in hierarchy. For $z \geqslant 0.75$, this probability increases to 73% becoming 100% at $z = 1$.

An interesting property of a jet, that has not yet been tested experimentally, is how the fraction of momentum carried by the fastest hadron in a jet depends on the charge of that hadron. In our model, for a u-quark, the first-rank primary meson can be positive or neutral, *not* negative. A negative hadron can arise only further down the hierarchy chain (rank greater than one) *or* from the decay of a

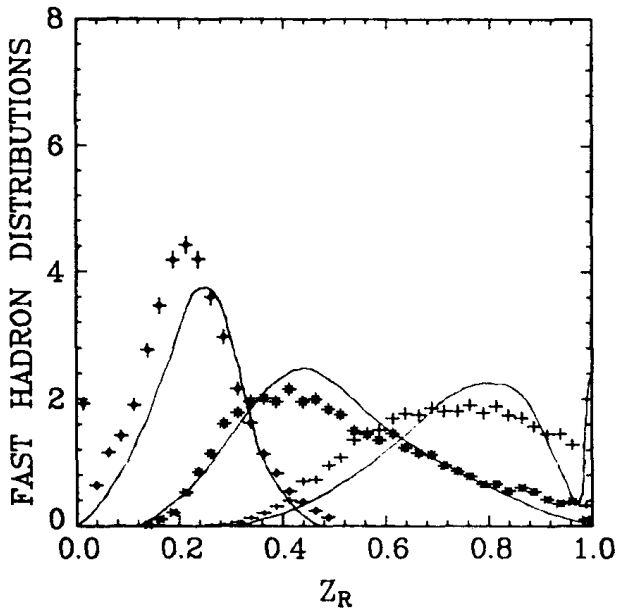


Fig. 19. Comparison of the predicted distributions of z_R for the fastest, second-fastest and fastest two charged hadrons in a u-quark jet (points) with data from $\nu p \to \mu^- +$ jet + X (ref. [15], solid curve), where z_R is the relative fraction of charged momentum as in fig. 17. (Note that in this figure, the theory are the points with errors and the data are represented by the solid curves.) $a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$, $P_q = 10$ GeV. × first charged, ◇ second charged, + first two charged, —— neutrino data.

rank-one neutral resonance. Because of this, the fastest positive hadron in a u-quark jet carries, on the average, a larger portion of the total momentum (45%) than does the fastest negative hadron (27%). Results of this type are tabulated in table 10 and illustrated in figs. 17 and 18. The fastest two positives in a u-quark jet carry 57% of the charged momentum, whereas only 34% is carried by the fastest two negatives. For a d-quark jet, the situation is reversed (see fig. 18); the fastest two negatives carry 52% of the charged momentum compared to 39% for the fastest two positives. These predictions should be easy to check by comparing the jets produced in $\nu p \to \mu^-$ + jet which are predominantly u-quark jets to those produced in $\bar{\nu} p \to \mu^+$ + jet which are predominantly d-quark jets.

A similar interesting and easily measured observable is the flavor of the fastest (largest-z) hadron. As table 12 shows, for a u-quark jet the fastest hadron is positive 42% of the time and negative only 20% of the time; the rest of the time, it is neutral. For a d-quark, on the other hand, the fastest hadron is positive 24% of the time and negative 34% of the time. The average charge of the fastest hadron in a u- and d-quark jet is about 0.23 and –0.10, respectively. The fastest hadron does "remember" to some extent the flavor of the quark that initiated the jet. Table 12 also shows the relatively large portion of strange particles predicted in our jets. For

Table 12
Number of times the various hadrons occured as the fastest (largest-z) particle for 10 000 quark jets with energy $P_q = 10$ GeV. Also shown are the average charge Q, third component I_3, and strangeness S of the fastest hadron

Fastest hadron	u	d	s
π^+	3194	1645	1447
π^0	2411	2462	1536
π^-	1511	2888	1356
K^+	1055	707	528
K^0	600	1062	517
K^-	466	473	2165
$\bar{K}^0$	458	458	2174
γ	305	305	277
h^+	4249	2352	1975
h^-	1977	3361	3521
h^0	3774	4287	4504
h^+/h^-	2.15	0.70	0.56
ave Q	0.227	−0.101	−0.155
ave I_3	0.191	−0.143	0.010
ave S	0.073	0.084	−0.329

a u-quark jet, the fastest hadron is a strange particle ($K^+, K^0, K^-, \bar{K}^0$) about 26% of the time.

3.5. Correlations between the fastest two charged hadrons

Although there is a considerable difference between rank in hierarchy and order in rapidity (or z), the hadrons in our jets do not completely "forget" the underlying hierarchy structure shown in fig. 1. The charge correlations between the fastest two charged hadrons, where h_1 is the fastest charged hadron and h_2 is the second fastest charged hadron and where neutrals are ignored, shown in tables 13a and 13b are a result of the underlying hierarchy structure and resonance decay effects. The combination $h_1 = +, h_2 = -$ occurs more often than $h_1 = +, h_2 = +$ and $h_1 = -, h_2 = +$ occurs more often than $h_1 = -, h_2 = -$. In fact, for a u-quark jet, the unlike sign configuration occurs almost twice as frequently as the like sign configuration. It is very unlikely for two negatives (positives) to occur as the fastest two charged particles in a u-quark (d-quark) jet (−− = 7% for a u-quark and ++ = 10% for a d-quark).

3.6. Determining the flavor of the quark

As we have seen, there are differences in the distribution of charge, total charge, fastest hadron, and correlations between the two fastest charged hadrons depending

Table 13a
Number of times various charge combinations occurred for the fastest (largest-z) two-charged hadron h_1, h_2 (h_1 = fastest charged hadron, h_2 = second fastest charged hadron) that fragmented from 10 000 quark jets with P_q = 10 GeV

h_1	h_2	u	d	s
+	+	2594	965	845
+	−	3608	3024	2743
+	none	420	102	81
+	arbitrary	6622	4091	3669
−	+	2550	3515	3648
−	−	731	1967	2193
−	none	31	287	341
−	arbitrary	3312	5769	6182
none	none	66	140	149

Same as above but $z_{h_1} > 0.5$

h_1	h_2	u	d	s
+	+	395	61	48
+	−	588	368	233
+	none	151	11	4
+	arbitrary	1134	440	285
−	+	264	520	474
−	−	36	226	264
−	none	4	99	84
−	arbitrary	304	845	822
none	none	8562	8715	8893

on whether the jet is initiated by a u-, d- or s-quark. The differences in the hadron charge distributions can only be seen if one has a large statistical sample of jets of one particular quark type. It is interesting to see if a criterion exists that allows one to determine the flavor of the quark jet on an event-by-event basis. Clearly, any criterion for finite-momentum jets will not work perfectly, so in order to compare the usefulness of various criteria, we define a reliability by

$$\text{reliability} \equiv (N_T - N_F)/(N_T + N_F) , \tag{3.7}$$

where N_T refers to the number of times the criterion correctly identifies the flavor of the jet from a sample containing an equal mixture of u- and d-quark jets and N_F is the number of times it is incorrect. If the criterion succeeds as often as it fails, the reliability is zero. The criterion may only apply to a sub-sample of the jets so we define an efficiency by

$$\text{efficiency} \equiv (N_T + N_F)/(\text{total number of jets}) . \tag{3.8}$$

Table 13b
Same as table 13a except that only hadrons with $z \geq 0.1$ are observed

h_1	h_2	u	d	s
•	+	1421	560	432
•	−	2006	1833	1610
•	none	2624	1205	1090
•	arbitrary	6051	3598	3132
	+	1618	1891	1978
	−	442	1060	1170
	none	936	2120	2272
	arbitrary	2996	5071	5420
none	none	953	1331	1448

Same as above but $z_{h_1} \geq 0.5$

h_1	h_2	u	d	s
•	+	158	33	12
•	−	253	225	140
•	none	723	182	133
•	arbitrary	1134	440	285
	+	174	217	220
	−	14	98	111
	none	116	530	491
	arbitrary	304	845	822
none	none	8562	8715	8893

For example, if we say that all jets where the fastest charged hadron is positive are u-quark jets and all jets where it is negative are d-quark jets, then for P_q = 10 GeV the reliability is about 0.25 and the efficiency is 99% (only 1% of the jets contain no charged particles). One can increase the reliability by lowering efficiency. Using the same criterion but applying it only to those jets that have one charged particle with $z \geq 0.5$ yields a reliability of 0.45 but an efficiency of only 14% *. In table 15, we list other criteria together with the reliability and efficiency for an equal mixture of u- and d-quark jets with P_q = 10 GeV.

* One disadvantage of the criterion which only can be applied to a small fraction of jets (low efficiency) is, of course, lower counting rate. But a much more serious difficulty is the bias introduced by not knowing experimentally the difference between a u-jet and a d-jet in the fraction of times the criterion can be applied. We could, of course, calculate these efficiency differences for our "standard" jet but we would not know whether they were right experimentally. The uncertainties are least if the efficiency is high.

Table 14
Number of times various weighted quark charge $Q_W(p)$ occured for 10 000 quark jets with energy $P_q = 10$ GeV, where $Q_W(p) = \Sigma_{i=1}^{N} z_i^p q_i$ with q_i and z_i the charge and z value of the ith hadron in a quark jet with N charged hadrons. Results are given for $p = 0.2$ and 0.5 and σ_{rms} is the root-mean-square deviation from the mean.

	Weighted charge	u	d	s
$p = 0.2$	$Q_W \geq 0$	7197	3499	3322
	$Q_W < 0$	2737	6361	6529
	$\langle Q_W \rangle$	0.39	−0.25	−0.27
	σ_{rms}	0.68	0.66	0.67
$p = 0.5$	$Q_W \geq 0$	7172	3601	3320
	$Q_W < 0$	2762	6259	6531
	$\langle Q_W \rangle$	0.26	−0.15	−0.18
	σ_{rms}	0.43	0.42	0.41
Same as above but observe only hadrons with $z \geq 0.1$				
$p = 0.2$	$Q_W \geq 0$	6172	3482	3045
	$Q_W < 0$	2875	5187	5507
	$\langle Q_W \rangle$	0.34	−0.19	−0.26
	σ_{rms}	0.76	0.77	0.75
$p = 0.5$	$Q_W \geq 0$	6168	3486	3039
	$Q_W < 0$	2879	5183	5513
	$\langle Q_W \rangle$	0.24	−0.14	−0.18
	σ_{rms}	0.52	0.53	0.51

From table 15, one sees that none of the first fourteen criteria listed gives a very high reliability while still maintaining a reasonable efficiency. This is because the rank in hierarchy and the order in momentum are considerably mixed up as discussed earlier and witnessed by table 11. The highest reliability with a large efficiency is obtained by assigning jets with total charge $Q > 0$ as u-quark jets and jets with $Q < 0$ as d-quark jets. The reliability is 0.45 with a 66% efficiency. Unfortunately, the total charge on a quark jet is not easily measured since it depends on including all the forward moving low-momentum hadrons which are difficult to measure without also including background from other sources. A quantity easy to measure experimentally is the total charge of all those hadrons in the jet which have $z \geq 0.1$. The reliability, however, now decreases to 0.38. If one wants to apply this criterion to all the events with $Q = 0$ jets assigned as d type, then the reliability decreases even further to 0.26 (number 8 in table 15).

The mean total charge of a large momentum u-jet is 0.60 and a d-jet is −0.40 differing by 1.0. This would seem to permit a clear separation of these jets. However, even with a jet of extreme momentum, the total charge Q of all hadrons of $P_z > 0$

Table 15
Various criterion for determining the flavor of the quark initiating the jet for an equal mixture of P_q = 10 GeV u- and d-quark jets together with the reliability = (true−false)/(true+false), where "true" refers to the number of times the criterion succeeded in selecting the correct flavor and "false" to the number of times the criterion failed. The efficiency is the number of jets (true+false) to which the criterion can be applied divided by the total number of jets generated.

Criterion	Reliability	Efficiency
1. fastest hadron positive ⇒ u-jet fastest hadron negative ⇒ d-jet	0.27	60%
2. fastest charged hadron positive ⇒ u-jet fastest charged hadron negative ⇒ d-jet	0.25	99%
3. charged hadron with $z \geq 0.5$ is positive ⇒ u-jet charged hadron with $z \geq 0.5$ is negative ⇒ d-jet	0.45	14%
4. fastest two charged hadrons are both positive ⇒ u-jet fastest two charged hadrons are both negative ⇒ d-jet	0.46	31%
5. total charge of jet $Q > 0$ ⇒ u-jet total charge of jet $Q < 0$ ⇒ d-jet	0.45	66%
6. same as 5 but only observe hadrons with $z \geq 0.1$	0.38	62%
7. total charge of jet $Q > 0$ ⇒ u-jet total charge of jet $Q \leq 0$ ⇒ d-jet	0.33	100%
8. same as 7 but only observe those hadrons with $z \geq 0.1$	0.26	100%
9. for N_{ch} = 1 jets $Q > 0$ ⇒ u-jet for N_{ch} = 1 jets $Q < 0$ ⇒ d-jet	0.68	4%
10. for N_{ch} = 3 jets $Q > 0$ ⇒ u-jet for N_{ch} = 3 jets $Q < 0$ ⇒ d-jet	0.47	13%
11. for N_{ch} = odd jets $Q > 0$ ⇒ u-jet for N_{ch} = odd jets $Q < 0$ ⇒ d-jet	0.40	50%
12. same as 9 but only observe hadrons with $z \geq 0.1$	0.38	34%
13. same as 10 but only observe hadrons with $z \geq 0.1$	0.31	15%
14. same as 11 but only observe hadrons with $z \geq 0.1$	0.36	50%
15. "weighted" charge $Q_W(p = 0.2) \geq 0$ ⇒ u-jet "weighted" charge $Q_W(p = 0.2) < 0$ ⇒ d-jet (see eq. (3.9))	0.37	99%
16. "weighted" charge $Q_W(p = 0.5) \geq 0$ ⇒ u-jet "weighted" charge $Q_W(p = 0.5) < 0$ ⇒ d-jet	0.36	99%
17. same as 15 but only observe hadrons with $z \geq 0.1$	0.28	89%
18. same as 16 but only observe hadrons with $z \geq 0.1$	0.28	89%
19. "weighted" charge $Q_W(p = 0.2) \geq 0.4$ ⇒ u-jet "weighted" charge $Q_W(p = 0.2) < -0.3$ ⇒ d-jet	0.51	63%
20. "weighted" charge $Q_W(p = 0.5) \geq 0.4$ ⇒ u-jet "weighted" charge $Q_W(p = 0.5) < -0.3$ ⇒ d-jet	0.58	46%
21. same as 19 but only observe hadrons with $z \geq 0.1$	0.38	62%
22. same as 20 but only observe hadrons with $z \geq 0.1$	0.40	55%

must be an integer and thus a random variable. There is an unavoidable noise depending on whether a particular charged particle in the plateau happens to have P_z greater or less than zero. Even though the plateau is neutral and all the difference of u- and d-quark jets lies far away at higher z, one is trying to sum a long series like +1−1+1+1−1 +1−1−1... not knowing where to stop, but knowing only that +1 and −1 become more and more equally likely to occur as we go further down the series (to lower z). The proper thing to do is, of course, the analogue of Abel summation, weigh the terms with a gradually decreasing weight as we go down the series. If the weight falls gradually enough from unity at the beginning, the excess charge there will be accurately picked up. However, the random ±1 far down where the weight has fallen toward zero will produce no fluctuations. That is, if particle i has "z rapidity" Y_{z_i} and charge q_i, we form the "weighted" charge

$$Q_W(p) = \sum_i q_i \exp(-pY_{z_i}) = \sum_i z_i^p q_i \,, \tag{3.9}$$

where p is a small number. This quantity will have a mean (close to $\langle Q \rangle$ as $p \to 0$) distinct for u- and d-quark jets. Furthermore, the "noise" or fluctuations expected from having to stop the sum below some finite $z_{\min}$ is $\pm z_{\min}^p$ which can be made small as long as $z_{\min}$ can be made small enough.

For a given experimental circumstance, however, $z_{\min}$ is fixed and the criteria that p be small and that $z_{\min}^p$ also be small are opposed. For sufficiently small $z_{\min}$ there is no problem, but because of the wide fluctuations in rapidity that the particles in our model suffer, we have found that in practice the method does not work as well as we hoped. For $z_{\min} = 0.1$, with $p = 0.5$, for example, the fluctuating uncertainty $z_{\min}^p$ is 0.3 times less than the gross sum $Q = \Sigma q_i$; but such a large p means that $Q(P)$ does not average as large as $\langle Q \rangle$. Even worse is that for such a large p the contributions of high-z particles depend so strongly on the precise z value they actually have.

Figs. 20 and 21 show the distribution of $Q_W(p)$ with $p = 0.2$ and 0.5, respectively, for a u- and d-quark jet of energy $P_q = 10$ GeV (including all hadrons with $P_z > 0$). The $p = 0.2$ distributions are considerably broader than the $p = 0.5$ case; however, the former has mean values $\langle Q_W \rangle$ that are more widely separated ($\langle Q_W \rangle_u - \langle Q_W \rangle_d = 0.64$ for $p = 0.2$ and only 0.41 for $p = 0.5$). In both cases, there is a clear separation of the u- and d-jets. By the use of table 14, we find a reliability of 0.37 if we assign jets with $Q_W \geq 0$ as u-quark type and those jets with $Q_W < 0$ as d-quark type with $p = 0.2$. The efficiency of this criterion is excellent (99% since we include only those jets with at least one charged hadron). One can obtain a higher reliability (but lower efficiency) by excluding from consideration those jets with Q_W values occurring in the overlap region of the u- and d-quark jet distributions. For example, table 15 shows that if we assign jets with $Q_W \geq 0.4$ as u-type and those with $Q_W < -0.3$ as d-type, then for $p = 0.5$ we get a 58% reliability with 46% efficiency. This "weighted" charge technique gives us better reliability factors than the

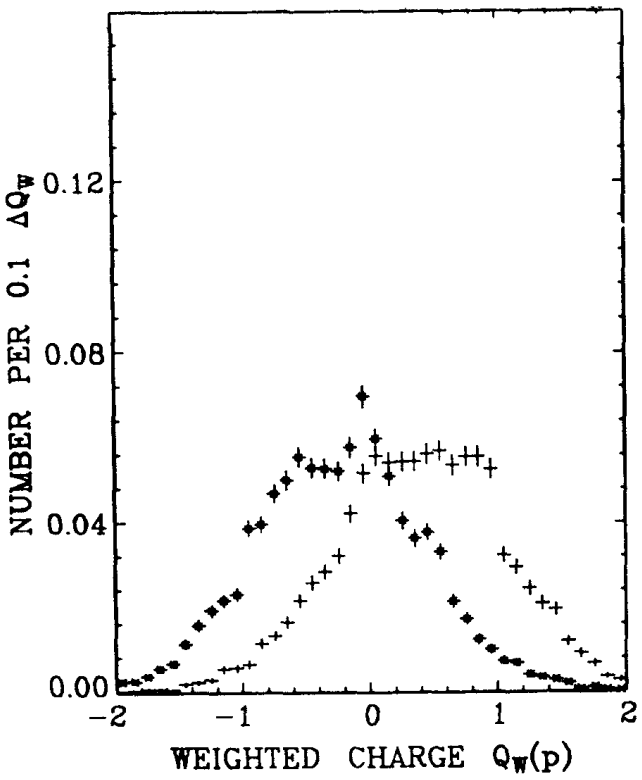


Fig. 20. Predicted distribution of weighted quark charge $Q_W(p)$ for a u- (+) and d-quark (X) jet with P_q = 10 GeV, where $Q_W(p) = \Sigma_{i=1}^{N} z_i^p q_i$ with z_i and q_i being the z value and charge of the ith hadron and where the sum is over all hadrons in the jet with $p_z \geq 0$. The value of the power p is taken to be 0.2. a = 0.77, $\alpha_{ps} = \alpha_v$ = 0.5. d-quark, $\langle Q_W \rangle$ = −0.25; u-quark, $\langle Q_W \rangle$ = 0.39.

other criterion in table 15. Nevertheless, even a reliability of 58% (about 4 out of 5 correct) is not too impressive.

In practice, for hadron-induced jets, it might be possible to go to lower z_{min}, perhaps even to p_z somewhat negative, allowing a smaller p value. The many "background" particles from the longitudinal jets may perhaps not contribute to much noise. But data from lepton jets should ultimately determine the best parameter p, z_{min} and criterion to use.

In view of the small reliability factors in table 15, we must conclude that our efforts to find a method of determining the quark flavor of a jet on an event-by-event basis have failed for jets of energy $P_q \lesssim 10$ GeV. The criterion can be made more reliable only with much higher momentum jets. (The method of "weighted" charge becomes an exact separation at very high momenta.) Nevertheless, table 15 is useful in summarizing the differences between u- and d-quark jets.

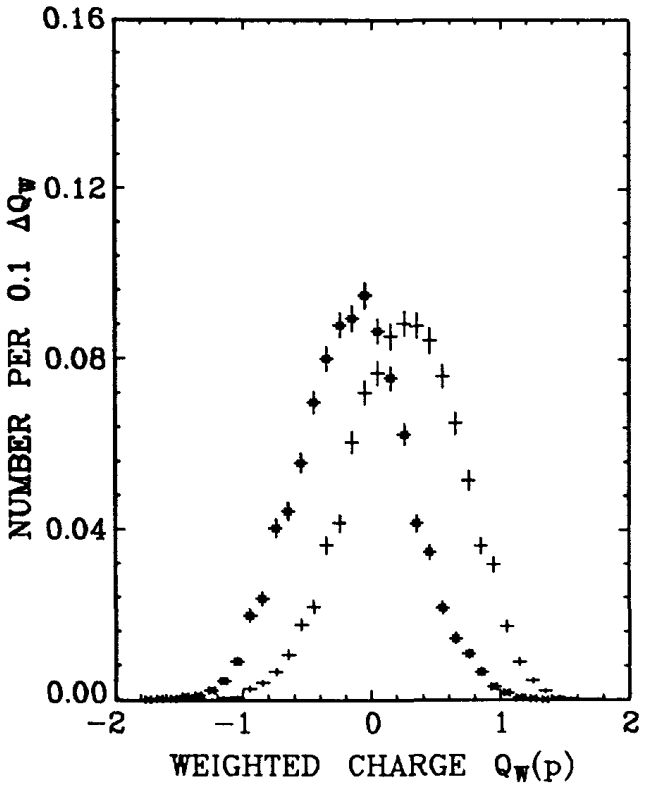


Fig. 21. Same as fig. 20 but where the power p is taken to be 0.5. d-quark, $\langle Q_W \rangle = -0.15$, u-quark, $\langle Q_W \rangle = 0.26$.

4. Properties of the quark rapidity plateau

4.1. Rapidity correlations

4.1.1. Correlations between adjacent-rank mesons

There are two sources of correlations in our model. Naturally, there is the correlation among secondary particles that are the decay products of the same primary meson. In addition, however, the primary mesons are not formed at random in rapidity. Primary mesons adjacent in rank are correlated in both flavor and rapidity since they each contain a quark (or antiquark) that came from the same $q\bar{q}$ pair. The two primary mesons of adjacent rank tend to occur near each other in rapidity, Y_z, as shown in fig. 22. The mean $|\Delta Y_z|$ between mesons adjacent in rank is about 1.8 units, where all the decay products of a particular primary meson are assigned the rank of that meson (see fig. 1). Fig. 22 also shows the distribution of $|\Delta Y_z|$ between mesons with the same rank ($\langle |\Delta Y_z| \rangle = 0.9$). All flavor correlations in the quark jets occur between primary mesons of adjacent rank. The flavor of a meson

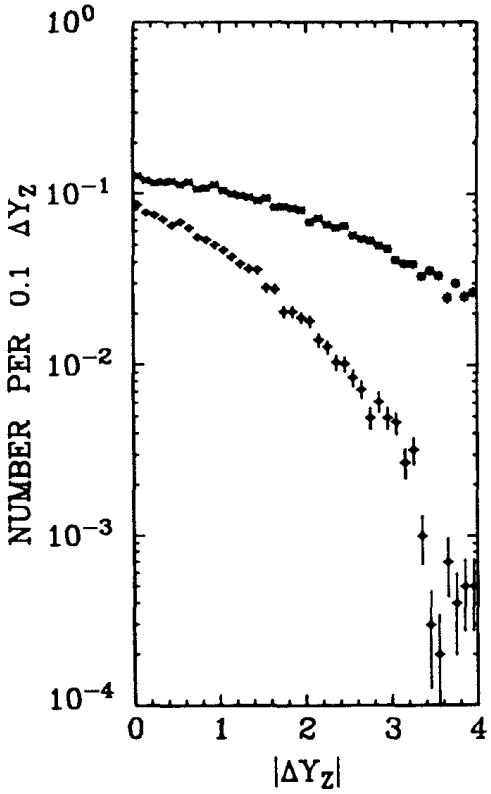


Fig. 22. The distribution of distances ΔY_z between the hadrons coming from one primary and those coming from another primary next in rank ×, $\langle |Y_z| \rangle = 1.8$. The distribution ΔY_z among the secondary hadrons which come from the decay of a single primary ◇, $\langle |Y_z| \rangle = 0.9$. $a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$.

of rank $r + 2$ is independent of the flavor of the meson of rank r. Rank is not a physical observable; however, the correlations shown in fig. 22 do produce effects in certain physical observables which we now proceed to examine.

4.1.2. Rapidity-gap distributions [16]

Figs. 23 and 24 show the distribution of Y_z gaps between all particles, charged particles, and positive particles (equals negative particles) before and after decay in the rapidity plateau *. Because two primary mesons of adjacent rank have net charge either ±1 or zero, small ΔY_z gaps occur more frequently between charge

* In order to avoid biasing against large Y_z gaps, the distributions in figs. 23, 24 and 25 are calculated by first ordering all particles in Y_z with $Y_{z_{i+1}} > Y_{z_i}$ and then considering all Y_{z_i} in the range $3.0 < Y_{z_i} < 6.0$ *but* when forming the gap length $\Delta Y_{z_i} = |Y_{z_i} - Y_{z_{i+1}}|$, $Y_{z_{i+1}}$ is allowed the range $3.0 < Y_{z_{i+1}} < 9.0$. Then only gaps $\Delta Y_{z_i} < 3.0$ are displayed. This method requires a rapidity interval *twice* as large as the maximum gap one decides to plot.

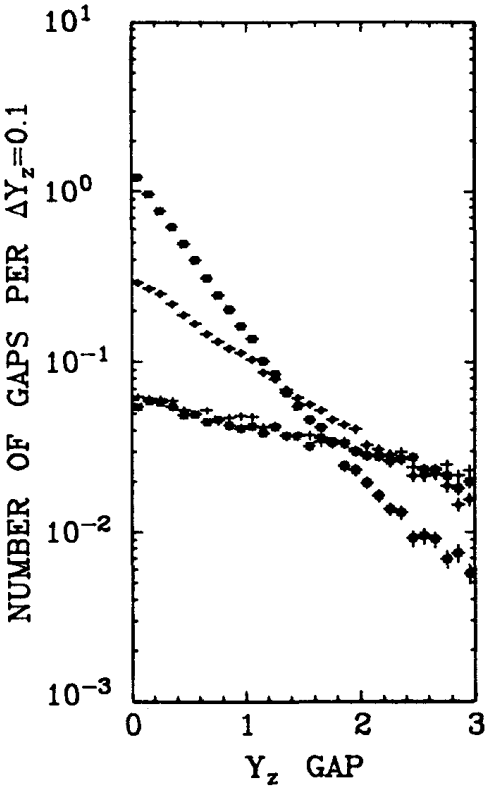


Fig. 23. Predicted number of times various Y_z gaps occur between all primary mesons, between charged primaries and between positive primaries (the same as for negative primaries) in the rapidity plateau of a u-quark jet *before* the primary mesons are allowed to decay. $a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$. □ all, ◇ charged, + positives, x negatives.

mesons (due to +–) than between negative mesons. For large Y_z gaps, the situation is reversed and there are more large Y_z gaps between the negatives than between the charged mesons. (Note that a gap between negatives can contain positives or neutrals whereas a gap between charged particles can only contain neutrals.)

The distribution of Y_z gaps carrying a specific charge is shown in fig. 25, where

$$Q_k = \sum_{i=1}^{k} q_i - q_0 \tag{4.1}$$

is the charge carried by the kth Y_z gap and q_i is the charge of the ith meson (ordered in Y_z). The quantity q_0 is the charge of the quark that initiated the jet (u-quark in fig. 25). If the mesons were ordered in hierarchy, only the net charge, $\Sigma_{i=1}^{k} q_i$, equal to one or zero could be obtained for a u-quark jet; Q_k must be the charge of an anti-quark. For ordering in Y_z, other values of $\Sigma_{i=1}^{k} q_i$ can occur, but only by mixing up hierarchy order and Y_z order. Because of this, for a u-jet, gaps of charge $-\frac{2}{3}$ and $\frac{1}{3}$

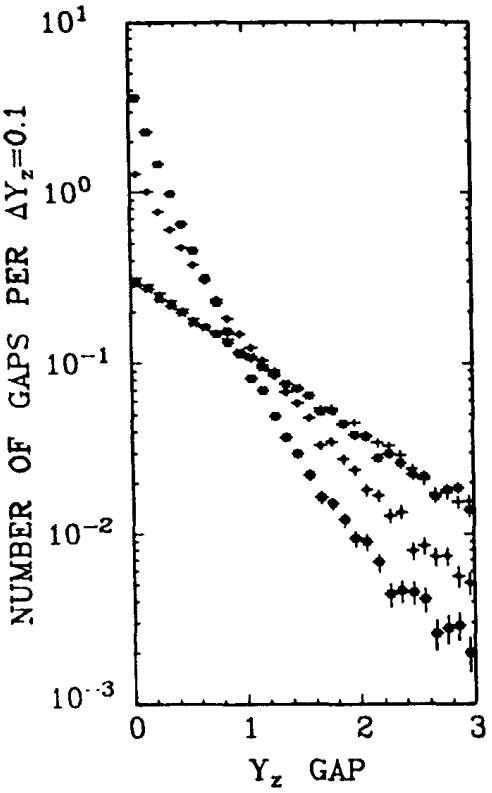


Fig. 24. The same as fig. 23 but *after* the vector mesons (and the η and η') have decayed according to the particle tables.

dominate at large gaps over the other gap charges *. (The distribution of $Q = -\frac{2}{3}$ and $Q = \frac{1}{3}$ gaps are equal, as are gaps of charge $Q = \frac{4}{3}$ and $Q = -\frac{1}{3}$, etc.)

4.1.3. Charged particle correlations

In analogy with the statistical-mechanical description of density fluctuations in a liquid, one defines a correlation function (see, for example, ref. [17])

$$C(Y_1, Y_2) = \frac{1}{\sigma}\frac{d^2\sigma}{dY_1\, dY_2} - \frac{1}{\sigma^2}\frac{d\sigma}{dY_1}\frac{d\sigma}{dY_2} . \tag{4.2}$$

The "correlation moment", f_2, is related to it by

$$f_2 = \int dY_1\, dY_2\, C(Y_1, Y_2) = \langle N(N-1)\rangle - \langle N\rangle^2 . \tag{4.3}$$

* In an exchange picture like that discussed by Pirilä, Thomas and Quigg in ref. [16], large gaps with charge $-\frac{2}{3}$ and $\frac{1}{3}$ correspond to the (allowed) exchange of a $\bar{u}$- and $\bar{d}$-quark, respectively, whereas the other gap charges correspond to lower-lying (exotic-type) exchanges. Hence the former, with a higher intercept ($\frac{1}{2}$), is expected to dominate at large gaps as it does.

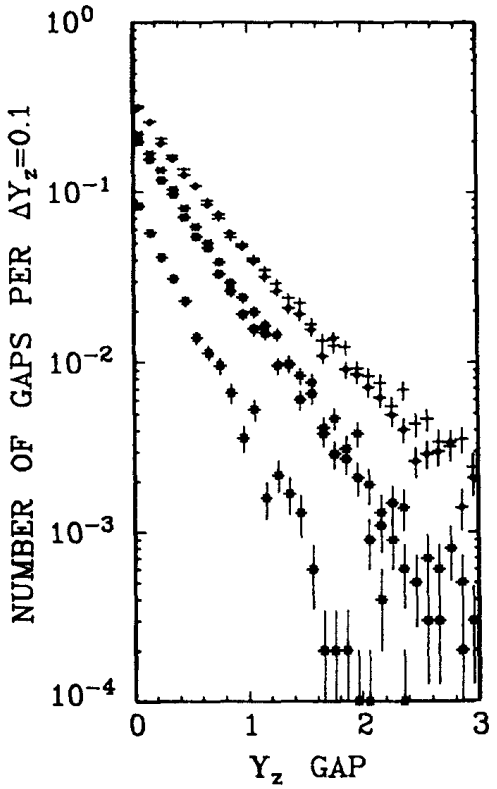


Fig. 25. Predicted number of times various *charged* Y_z gaps occur for a u-quark jet (after decay) where the charge of the kth gap is given by $Q_k = \Sigma_{i=1}^{k} q_i - q_0$, where q_i is the charge of the ith hadron (ordered in Y_z) and q_0 is the charge of the initiating quark (in this case $q_0 = \frac{2}{3}$). $a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$. $\diamond\ Q = -\frac{2}{3}$, $+\ Q = \frac{1}{3}$, $\times\ Q = \frac{4}{3}$, $\square\ Q = -\frac{1}{3}$, $*\ Q = -\frac{4}{3}$.

In addition, one can define a "normalized" correlation function by

$$R(Y_1, Y_2) = C(Y_1, Y_2) \Big/ \frac{1}{\sigma^2} \frac{d\sigma}{dY_1} \frac{d\sigma}{dY_2} . \tag{4.4}$$

In the absence of any dynamical correlations, the probability to observe in a single jet one hadron at rapidity Y_1 and a second hadron at rapidity Y_2, together with anything else, would be equal to the product of the probabilities to find hadrons at rapidity Y_1 and Y_2 in different jets and (4.2) would vanish. Figs. 26 and 27 show the value of $C(Y_{z_1}, Y_{z_2})$ and $R(Y_{z_1}, Y_{z_2})$, respectively, for a negative hadron at $Y_{z_1} = 4.0$ and a positive (upper) and negative (lower) hadron at Y_{z_2} *versus* $\Delta Y_z = Y_{z_1} - Y_{z_2}$. The region $-2 \leqslant \Delta Y_z \leqslant 2$ is the plateau region and $2 \leqslant \Delta Y_z \leqslant 4$ is the "end of the quark" region as can be seen from fig. 8. Also shown are the analytic results for the case where no particles decay (in this case, the analytic method is exact). Resonance decay clearly plays an important role in correlations; however, in

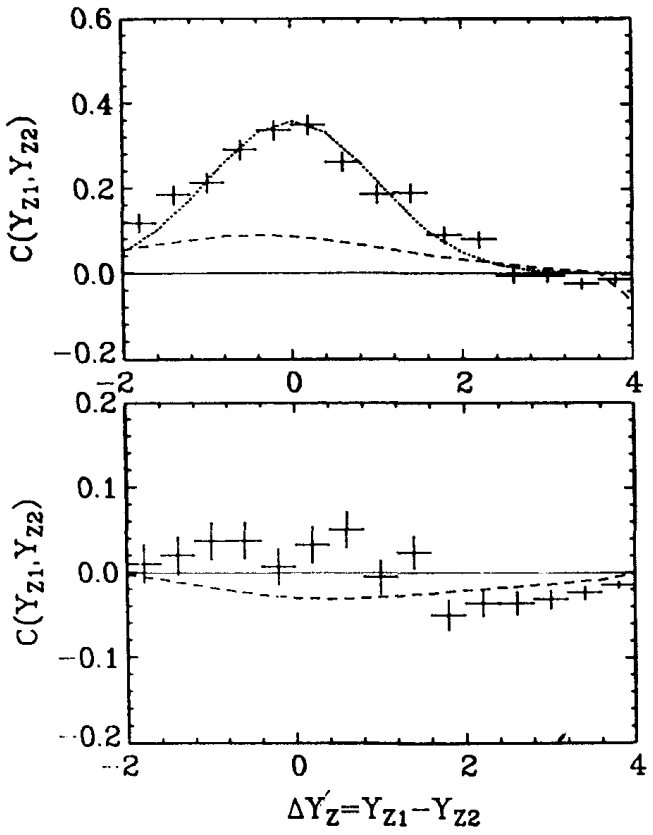


Fig. 26. Predicted behavior of the two-particle correlation function $C(Y_{z_1}, Y_{z_2})$ defined by (4.2) where $Y_{z_1} = 4.0$ and we have plotted *versus* $\Delta Y_z = Y_{z_1} - Y_{z_2}$. Results are given for h_1 negative and h_2 positive (upper) and h_1 negative and h_2 negative (lower) and are generated using Monte Carlo (points). The dashed curves are the results for $C(Y_{z_1}, Y_{z_2})$ *before* the primary mesons are allowed to decay. The dotted curve (to guide the eye) is $0.36\exp(-\frac{1}{2}(\Delta Y_z)^2)$. $a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$, $Y_{z_1} = (4.0, 4.4)$.

our model there is some clustering of the primary mesons before decay. The final correlation function (after decay) $C(Y_1, Y_2)$ for oppositely charged hadrons shows the characteristic short-range correlation behavior [16–19] roughly like $\exp(-(\Delta Y_z)^2/4\delta^2)$ with a correlation length, 2δ, of about 1.4. The correlation between two negatives is, on the other hand, quite small. The "correlation moment" f_2 in (4.3) after decay is given in table 3 and for ++ and −− it is negative.

4.1.4. Mass distribution

The distribution of two-particle mass for u-quark jets of $P_q = 10$ GeV resulting from the model is shown in figs. 28 and 29. One can see the ρ^0 and K^{*+} in the $\pi^+\pi^-$ and $K^+\pi^0$ combinations (our resonances have zero mass width). If our estimate of

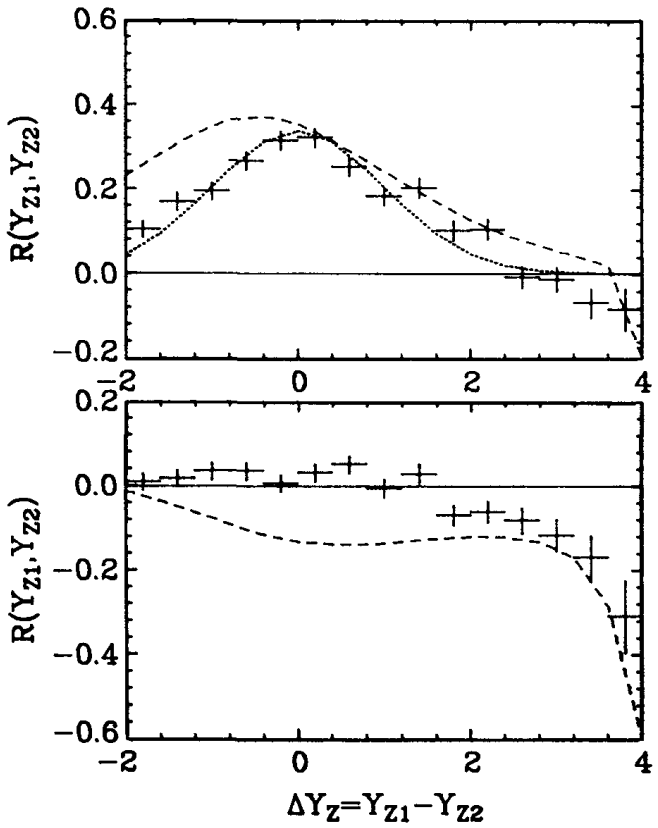


Fig. 27. Same as fig. 26 but for the "normalized" correlation function defined in (4.4). The dotted curve (to guide the eye) is 0.34 $\exp(-\frac{1}{2}(\Delta Y_z)^2)$.

the resonance contributions is roughly correct, then one would expect to see such resonances in the jets produced in lepton- or hadron-initiated processes.

4.2. Comparison with pp collisions

Many of the features predicted by our model for the behavior of the quark rapidity plateau are similar to those observed for the low-$p_\perp$ rapidity plateau generated in proton-proton collisions. The rapidity gap distributions, +− charged correlations, mass distributions, and charged multiplicities are quite similar to that seen in pp interactions. Similarities between the plateau behavior in lepton and low-$p_\perp$ hadron experiments have often been noted [20]. It will be interesting to see experimentally if they are really very much alike and to study theoretically why this may (or may not) be so.

In making this comparison, one must be careful of the contributions from diffrac-

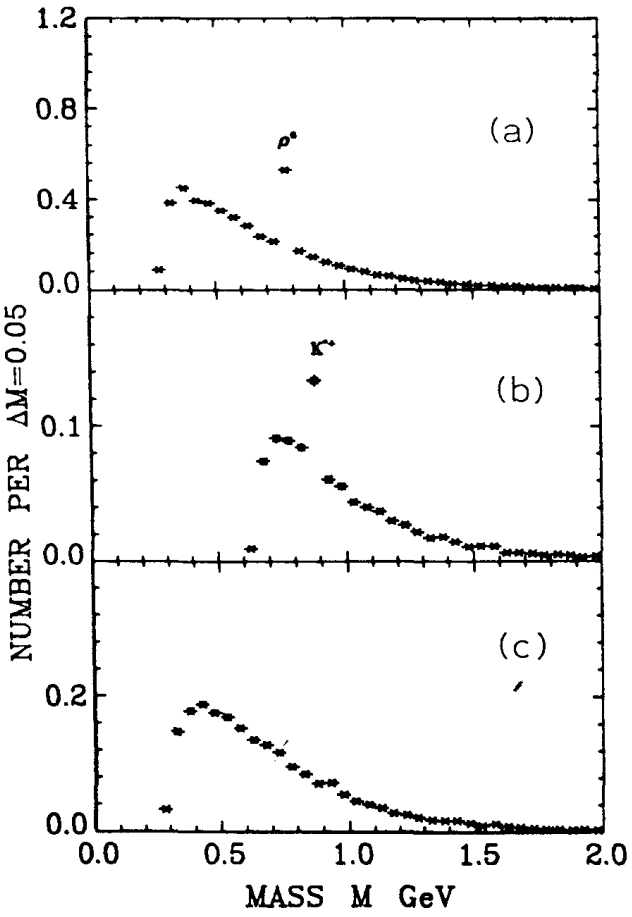


Fig. 28. Plot of the two-particle mass of (a) $\pi^+\pi^-$, (b) $K^+\pi^0$, and (c) $\pi^+\pi^+$ from a u-quark jet with P_q = 10 GeV from our jet model with a = 0.77, $\alpha_{ps} = \alpha_v$ = 0.5. (Note that our resonances have zero width.)

tive events in pp collisions which have no counterpart in the quark jets. The comparison should properly be made only with the inelastic non-diffractive part of pp collisions (if that can actually be defined and separated out). One place where this diffractive component plays an important role is in the behavior of f_2^-. In pp collisions, f_2^- becomes positive as the negative particle multiplicity $\langle N_- \rangle$ increases (i.e., as the energy of the collision is increased). On the other hand, our model for quark jets predicts a negative f_2^- for $P_q \lesssim 500$ GeV (see table 3) and recent data on νp interactions, where the model should apply, indicate that it is indeed negative [21]. For those $p\bar{p}$ annihilation events which contain no final baryons so diffraction plays no role f_2^- is also negative.

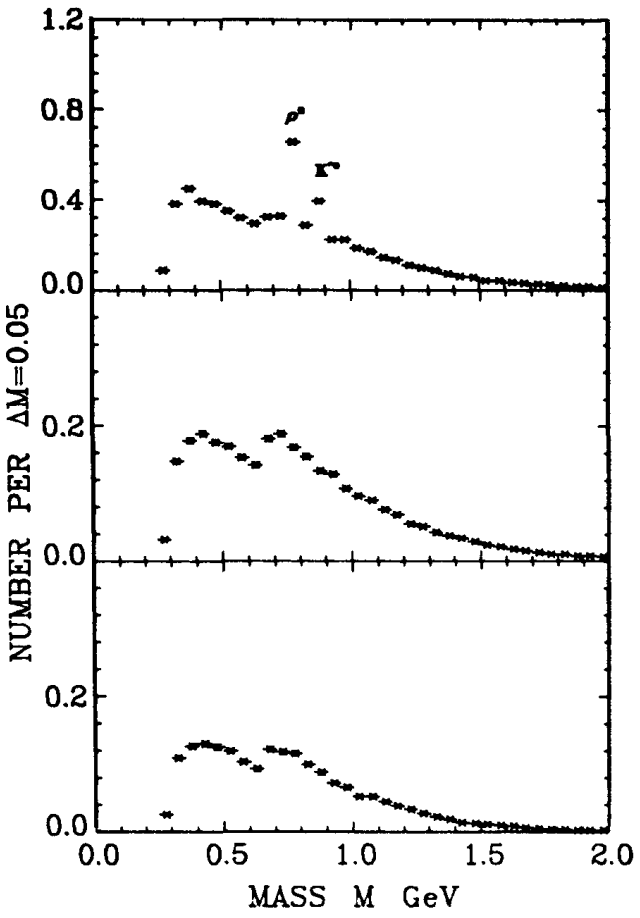


Fig. 29. Same as fig. 28 but for (a) h^+h^-, (b) h^+h^+, and (c) h^-h^-, where $h = \pi + K$.

4.3. Problems with the model

There is, of course, the obvious problem that the model contains no protons or other baryons and is thus incomplete. They might be included in the same framework if we were to consider that the field which makes the quark-antiquark pairs could, from time to time, make pairs of diquarks (two quarks whose color is in a $\bar{3}$ representation) and anti-diquarks. To implement this idea, however, requires inventing a few new functions and parameters which we have too little information at present to guess. Since baryons are probably not produced with great frequency in quark jets, neglecting them probably causes no serious modifications of the meson distributions. Of course, we cannot predict baryon distributions at all.

In addition, we have not included quantum effects. We have dealt solely with probabilities and not amplitudes. It would be interesting to imagine the recursive

principles to apply to amplitudes instead of probabilities. Instead of (2.3), one would write

$$\sqrt{\gamma_a}\sqrt{\gamma_b}\sqrt{\gamma_c}\ldots\prod_i \phi(\eta_i) \tag{4.5}$$

as the amplitude to find primary mesons with a set of flavors and momenta, where $\phi(\eta)$ is a (possibly complex) amplitude whose absolute square is $f(\eta)$. By adding amplitudes in the appropriate manner, allowance would then be made for the Bose nature of indistinguishable primary mesons.

The most serious problem with the model is that it suffers from a defect in principle. As Casher, Kogut and Susskind have pointed out [10], the correct way to look at the development of these jets is from the center out, not from each end in toward the center. That is, as the original quarks' rapidities separate in, say, an e^+e^- experiment, the first new quark pairs are made at relatively small momenta, then the quarks leaving the new pair generate a new pair between them, etc. We have not been able to develop a simple ansatz for both flavor and momentum distributions in accord with this idea. Instead, our ansatz thinks of the first pair forming near one end of the momentum chain, and then further ones follow generally down the momentum scale. This mechanism has been well refuted by Bjorken and Kogut and Susskind. Our first surprise from the "chain-decay" point of view is that the function, $zf(1-z)$, in (2.21) giving the momentum distribution of the first primary meson must be so peaked toward *low* momentum, z, to agree with experiment. It is not clear how the leading quark manages to find itself so far down in momentum.

We can also see that something is wrong in principle with the chain-decay ansatz by considering the nature of the plateau region for an energetic $q\bar{q}$ pair produced in an e^+e^- colliding beam experiment. We have argued that this plateau region could be analyzed by going far down in rapidity in the jet from the quark q; even so far down that the hadrons are, in fact, moving slightly backwards (i.e., in the direction of $\bar{q}$). If this is true, then all properties of this region would have to be the same if analyzed (in reverse order) as a property of the $\bar{q}$-jet. Measure rapidities in the c.m.s. so the quark q has rapidity Y_0 and the antiquark $-Y_0$. We first ask in the q-system, what is the probability that three primary mesons adjacent in hierarchy are located at rapidities Y_1, Y_2, and Y_3 (per dY_1, dY_2, dY_3). This function $K(Y_0 - Y_1, Y_0 - Y_2, Y_0 - Y_3)$ deep in the plateau region depends only on rapidity differences, and not on Y_0 (i.e., it equals a function $L(Y_2 - Y_1, Y_3 - Y_2)$). Now seen in the other direction as a $\bar{q}$ jet, the three primary mesons are in rank 3, 2, 1 and the rapidity distance from the end of $\bar{q}$ are $Y_0 + Y_3$, $Y_0 + Y_2$, and $Y_0 + Y_1$. Calculating in this way, we have $K(Y_0 + Y_3, Y_0 + Y_2, Y_0 + Y_1)$ and, therefore, $L(Y_3 - Y_2, Y_2 - Y_1)$. Choosing $Y_2 = 0$, $Y_1 = \ln \xi$, and $Y_3 = -\ln \eta$ so that the three particles are in the $E + p_z$ ratio ξ, 1, $1/\eta$ (or seen in the reverse order $E - p_z$ ratio η, 1, $1/\xi$), we should have

$$L(-\ln \xi, -\ln \eta) = L(-\ln \eta, -\ln \xi)\,, \tag{4.6}$$

if the theory were truly symmetrical.

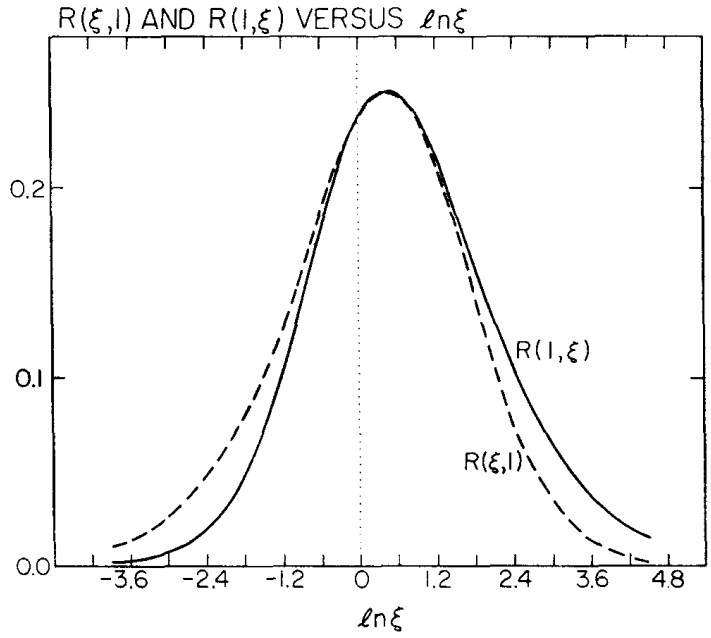


Fig. 30. Illustration of the asymmetry in our model of functions $R(\xi, \eta)$ describing correlations among three primaries adjacent in rank in the plateau (discussed in subsect. 4.3). The curves are $R(1, \xi)$ (solid) and $R(\xi, 1)$ (dashed) *versus* ln ξ. They were calculated using $f(\eta) = 3\eta^2$. Theoretically, $R(\xi, \eta)$ should be a symmetric function.

Direct calculation of the function $L(-\ln \xi, -\ln \eta)$, which we write as $\xi\eta R(\xi, \eta)$ so R is the number of particles per $d\xi\, d\eta$, shows it is *not* symmetrical. The difference between $R(\xi, \eta)$ and $R(\eta, \xi)$ is not great, yet that there is any difference at all shows our physical view is not entirely sound. The difference is illustrated in fig. 30, where we choose $\eta = 1$ and compare $R(\xi, 1)$ and $R(1, \xi)$ calculated for the simple case $f(\eta) = 3\eta^2$. The differences for other values of η are similar or not as large. Since rank in hierarchy is not physically observable, neither is $R(\xi, \eta)$. However, any amount of asymmetry of $R(\xi, \eta)$ will result in an observable (in principle) asymmetry in the rapidity plateau for three hadrons whose flavors are such that they *could* occur adjacent in rank (e.g., K^+, K^-, π^+). These asymmetries are quite small. None the less, they represent a defect of the model.

We have found the "chain-decay" ansatz so simple to formulate and to analyze arithmetically and analytically that we wanted to study it as a possible approximation by which we might get a rough idea of the general behavior of quark jets. It is important to have such a scheme to make some suggestions for the planning and design of "jet" experiments. It is delightful that so few assumptions will yield a jet model which is so complete in being able to describe a jet. (Only the transverse momentum correlation questions are unanswered.)

5. Transverse-momentum results

5.1. Transverse-momentum distributions versus z, Y, and M^2

The *mean* distribution of transverse momentum depends in no way on our assumed correlations between primary mesons. None of the details of our recursive scheme are of much concern; just the distribution in z, $F(z)$, of primary mesons and the assumption that all the *primary* mesons are distributed in transverse momentum in the same way (which we take to be a Gaussian as eq. (2.46) independent of z, flavor or spin). In our previous work (FFF), we assumed that all the *final* mesons (direct plus indirect) were distributed the same, independent of z. Our newer view, that it is the *primary* mesons which have the universal distribution in transverse momentum, produces a number of interesting effects which we will now discuss.

The transverse momentum of the final mesons has two components: the original momentum $P_\perp$ of the primary mesons plus the additional $Q_\perp$ received by the products of those primary mesons that decay. Assuming that half of the primary mesons are vectors that decay and letting the η and η' also decay, we can easily calculate the mean transverse momentum expected for various particles in the jet. As discussed in subsect. 2.7.2, we find that using $\sigma = 350$ MeV in (2.48) produces a mean transverse momentum of charged pions of $\langle k_\perp \rangle_\pi = 323$ MeV. The mean transverse momentum of the primaries (2.48) is then 439 MeV, considerably larger than that of the secondaries [*].

The reason for this can be seen in the coordinate system boosted so that the z component of the primary meson is zero. It does have a $p_\perp$, however. If it is of mass M and disintegrates into two mesons each of mass $\frac{1}{2}M$ (so that the Q value is zero), then these decay products have only $\frac{1}{2}p_\perp$. Thus disintegrations considerably reduce the effect of $p_\perp$. The mean value of $Q_\perp$ generated by a spherical distribution of maximum momentum Q if $p_\perp = 0$ is $\frac{1}{4}\pi Q$, but the values of Q from the particle tables are not so large as 400 MeV on the average. Of course, there are mutual effects when both $p_\perp$ and Q are non-zero but the qualitative effect is clear; the large $\langle p_\perp \rangle$ of the primaries is decreased by observing the mean $\langle k_\perp \rangle$ of the final mesons. Even though some primaries do not decay and thus contribute directly to $\langle k_\perp \rangle$, in averaging over all particles the indirect mesons (decay products) have more weight for they occur more frequently. In addition, because more of the decay products of the vector mesons (and $\eta + \eta'$) are pions than kaons, the transverse-momentum distribution of kaons resembles more closely the distribution of primaries than does the pion distribution. The latter distribution is more sharply peaked at small $k_\perp$, as shown in fig. 31. The mean $k_\perp$ for final pions is

$$\langle k_\perp \rangle_{\pi^\pm} = 323 \text{ MeV}, \tag{5.1a}$$

whereas for final kaons, it is

[*] This has been pointed out by Seiden in ref. [13].

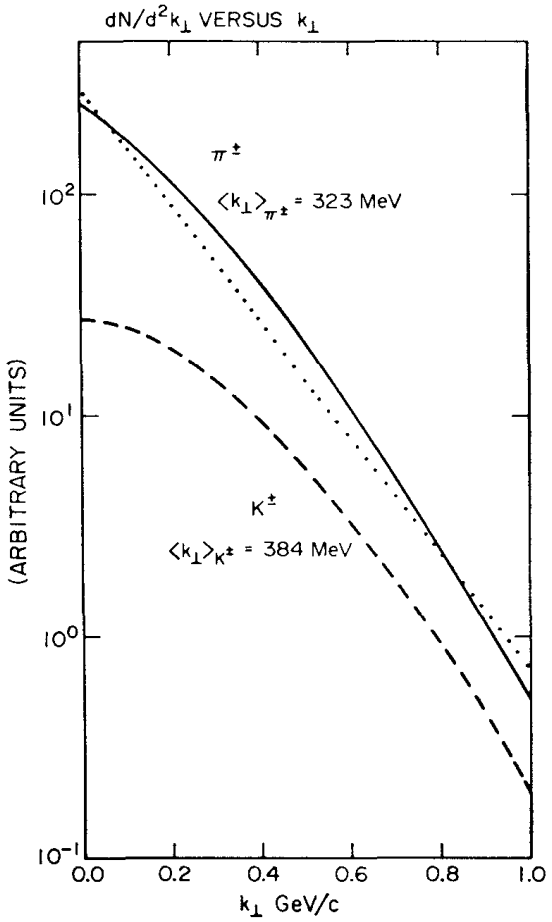


Fig. 31. Distribution of transverse momentum, $dN/d^2k_\perp$ *versus* $k_\perp$ for $\pi^\pm$ and $K^\pm$ resulting from a u-quark jet in our model with $a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$ and $\sigma = 350$ MeV. The dotted straight line is $\exp(-6k_\perp)$.

$$\langle k_\perp \rangle_{K^\pm} = 384 \text{ MeV} . \tag{5.1b}$$

Almost all the rho mesons are direct so that

$$\langle k_\perp \rangle_\rho = 439 \text{ MeV} , \tag{5.1c}$$

which is the same as the distribution of primary mesons. The mean $k_\perp$ of all the final particles is very close to that for pions since 75% of the final particles are pions. The exact shapes of the transverse-momentum distributions shown in fig. 31 depend on our arbitrary assumption of a Gaussian distribution for the primaries (2.46). The tendency for higher-mass particles like the K or ρ to have a flatter distribution in $k_\perp$ and a larger $\langle k_\perp \rangle$ than the lighter π mesons is observed in the plateau region of ordinary pp collisions. In our model, the reason is not directly

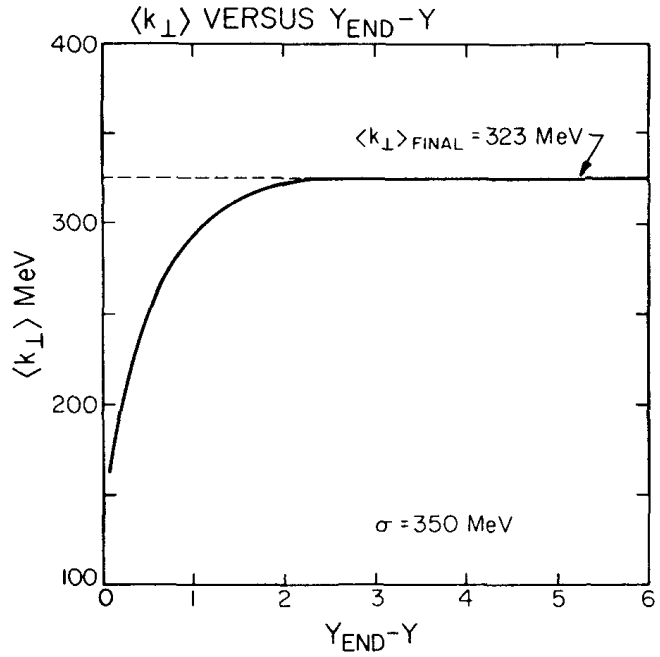


Fig. 32. The mean value of the transverse momentum of all hadrons, $\langle k_\perp \rangle$, *versus* $Y_{\text{end}} - Y$, where Y_{end} is as in fig. 11.

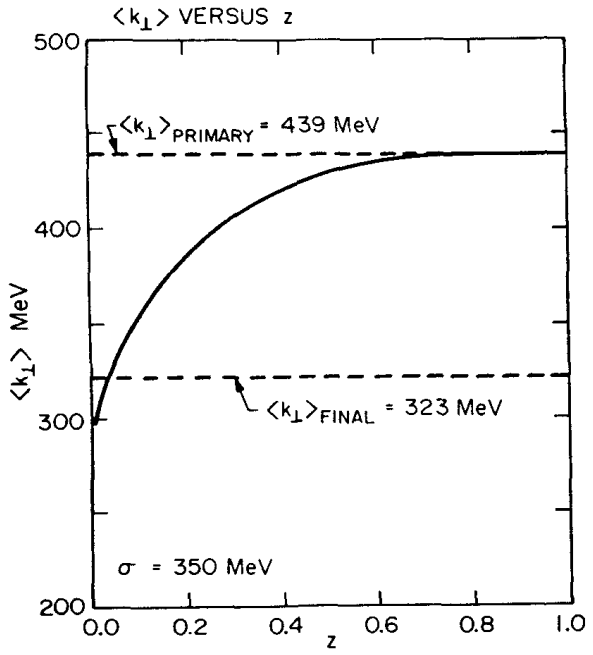


Fig. 33. The mean value of the transverse momentum of all hadrons, $\langle k_\perp \rangle$, *versus* z. Also shown by dashed lines are the total mean $k_\perp$ (summing over all z) of the primary mesons before decay (439 MeV) and after decay (323 MeV).

due to the K and ρ being more massive than the π, but because a smaller fraction of them are due to resonance decays than for the π.

How is the transverse momentum distributed in rapidity Y and in z? We see from fig. 32 that $\langle k_\perp \rangle$ for particles of a given rapidity Y, in the plateau region, and well toward the forward direction where the number of particles begins to fall off, have the mean $k_\perp$ of 323 MeV as in (5.1a). Only the few particles in the large-Y region (near Y_{end}) have $\langle k_\perp \rangle$ smaller than 323 MeV. On the other hand, particles with a given z have a mean $k_\perp$ which depends on z as shown in fig. 33. For $z \gtrsim 0.1$, $\langle k_\perp \rangle$ is larger than 323 MeV and approaches the mean of the primary mesons, 439 MeV, as z approaches one.

This latter behavior is obvious since only primary mesons can obtain a z value very near one. However, the reason $\langle k_\perp \rangle$ for particles with z in the range 0.2–0.6 is so large requires further discussion. Consider a primary meson generated at some z_0 (and true rapidity Y_0), but look at it in the frame for which it has zero longitudinal momentum p_z. First, suppose its original perpendicular momentum, $p_\perp$, is zero and suppose it decays into two mesons. Then the secondaries acquire with equal likelihood both positive and negative z component of momentum Q_z; that is, both positive and negative rapidity. Thus, back in the lab system (just add a constant Y), the secondaries are found symmetrically spread in rapidity about the original rapidity Y_0 of the parent. It is otherwise with z, the secondaries are widely spread in z but only over values below z_0 (uniformly from 0 to z_0 for massless products, for example).

Now to see what happens if the parent meson has non-zero $p_\perp$, imagine this $p_\perp$ to arise from a very small rotation of the original large p_z (by an angle $p_\perp/p_z$). It is clear that if a product ends up at a momentum p_{z_1}, its share of the parent $p_\perp$ is only $(p_{z_1}/p_z)p_\perp$, or fractionally z/z_0 of the original. That is, daughter particles carry a fraction z/z_0 of the originally large parent perpendicular momentum. Those that obtain a z near z_0 have a larger share than those which obtain a smaller z. It is this effect (combined with the fact that higher-z particles are more likely themselves to be primaries) which means that the higher-z particles ($z \gtrsim 0.1$) have a mean $k_\perp$ considerably larger than the average. (It is true that decay products with the very highest z/z_0 have received so much Q_z that the mean transverse decay momentum $Q_\perp$ must be smaller. But this depends on Q_z (or z/z_0) *via* a curve with a vertical tangent, an ellipse in fact, and the effect is not great and is overwhelmed by the $(z/z_0)p_\perp$ effect.)

The mean $k_\perp$ of all final particles is not much affected by the region $z \gtrsim 0.1$ because there are so many more particles of low z. Hence, the region $z \gtrsim 0.1$ can have a considerably higher $\langle k_\perp \rangle$ than the $\langle k_\perp \rangle$ of all particles. Far in the plateau of Y_z, one ultimately gets the same low mean of all particles. But, where the curve of the number of particles *versus* Y_z first begins to fall, the $\langle k_\perp \rangle$ begins to rise rapidly since each decay particle is sent to larger Y_z. For the true rapidity Y, however, decay sends particles equally up and down in Y so even a linear fall-off of the Y dis-

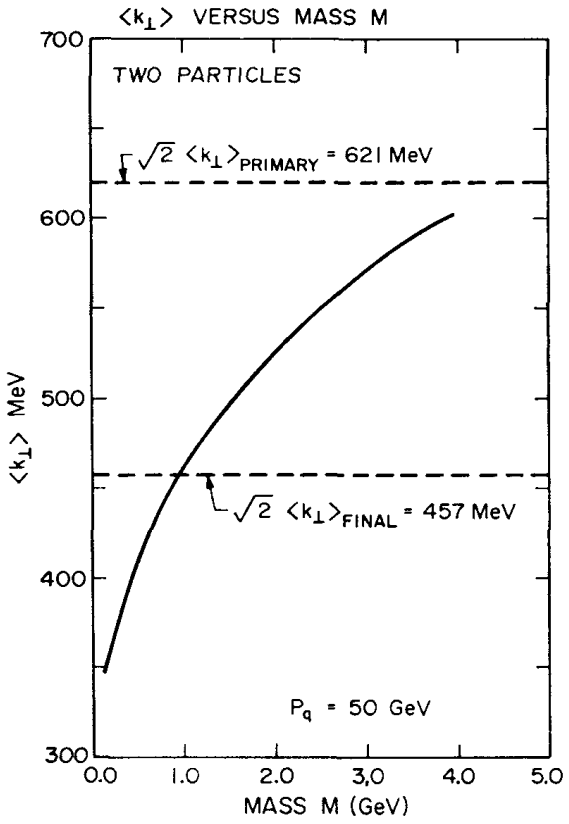


Fig. 34. The mean value of the transverse momentum, $\langle k_\perp \rangle$, of two particles *versus* the two-particle mass M. The results are for a u-quark of energy P_q = 50 GeV. Also shown, by dashed lines, is $\sqrt{2}$ times the $\langle k_\perp \rangle$ of primaries (621 MeV) and $\sqrt{2}$ times the $\langle k_\perp \rangle$ of the final mesons (457 MeV).

tribution would not distort the mean $k_\perp$. More radical variations, as those near the high Y_{end} do, however. The very highest Y values come from particles having abnormally low $m_\perp$ so the mean $k_\perp$ here is lower (see fig. 32).

Finally, for completeness, we show in fig. 34 the mean total $k_\perp$ of *two* particles whose mass M is given by

$$M^2 = (E_1 + E_2)^2 - (p_{z_1} + p_{z_2})^2 - k_\perp^2 \ , \tag{5.2a}$$

with

$$k_\perp^2 = (p_{x_1} + p_{x_2})^2 + (p_{y_1} + p_{y_2})^2 \ , \tag{5.2b}$$

where E_1 and E_2 are the energies of the two particles, respectively. As a function of the mass M, the two-particle $\langle k_\perp \rangle$ increases from a value of about 350 MeV at small mass M to a value equal to $\sqrt{2}$ times $\langle k_\perp \rangle_{primary}$ or about 621 MeV for M greater

than about 5.0 GeV. This sharp rise of $\langle k_\perp \rangle$ with mass M has also been observed in the plateau region of ordinary pp collisions [22].

We feel that the description of transverse momentum presented here is more satisfactory than the one used by us in FFF, in which all the final mesons in a quark jet had the same $\langle k_\perp \rangle_{q \to h} = 330$ MeV independent of z. We must change this to read that all *primary* mesons have the same $\langle k_\perp \rangle = 439$ MeV independent of z. Because the large-$p_\perp$ hadron experiments which we analyzed in FFF are sensitive predominantly to the large-z region of the quark decay functions so that $\langle k_\perp \rangle_{q \to h} \approx 439$ MeV, a number of our predictions in FFF will be changed •. For example, the mean values of P_{out} appearing in table 3 of FFF are all increased coming closer to the experimental findings. Also predictions made for experiments with small aperture acceptance will have to be reduced because of the wider spread in angle of the large-z particles. We are now computing corrections to FFF based on our new jet model. The results will be discussed elsewhere.

5.2. Correlations in transverse momentum

Our assumptions in subsect. 2.7.2 that each produced quark-antiquark pair conserves transverse momentum with no net $k_\perp$ and that the transverse momentum of a primary meson is the sum of the $k_\perp$ of two quarks leads to correlations in transverse momentum of the hadrons in the quark jet. As discussed in that subsection, primary mesons of adjacent rank tend to go oppositely with $\langle \boldsymbol{k}_{\perp_1} \cdot \boldsymbol{k}_{\perp_2} \rangle = -\sigma^2$. However, as we have already learned from fig. 22, adjacent-rank mesons can find themselves considerably spread apart in rapidity Y_z so that these correlations are of long range. Fig. 35 shows the behavior of the asymmetry

$$\Sigma(Y_1, Y_2) = [N_D(Y_2) - N_U(Y_2)] / [N_U(Y_2) + N_D(Y_2)], \tag{5.3}$$

where $N_U(Y_2)$ and $N_D(Y_2)$ are the number of hadrons at Y_2 with $|\phi_{12}| < \frac{1}{2}\pi$ and $|\phi_{12}| > \frac{1}{2}\pi$, respectively, and where ϕ_{12} is the angle between $\boldsymbol{k}_{\perp_1}$ and $\boldsymbol{k}_{\perp_2}$. This

• After the completion of FFF, in which we assumed $\langle k_\perp \rangle_{q \to h} = 330$ MeV independent of z, several people pointed out to us that there were indications that $\langle k_\perp \rangle$ was, in fact, larger than 330 MeV for large z both from hadron and lepton experiments. They further noted that we would improve our predictions of large-$p_\perp$ correlations in pp collisions if we took a larger $\langle k_\perp \rangle_{q \to h}$. In particular, we would predict a larger P-out in better agreement with experiment. It is now clear from the analysis of Seiden. [13] and from data on νp collisions [21,23] that $\langle k_\perp \rangle_{g \to h}$ *is* considerably larger than 330 MeV at large z. We have previously been reluctant to change (stubborn) because we did not want to "fiddle" our high-$p_\perp$ model in order to agree with the very experiments we were trying to predict. Now that we have added the physically natural assumption that some of the pions are secondary to resonances, we feel it is natural to replace the naive simple rule that the pions have a constant (with z) mean transverse momentum by the equally naive and simple rule that it is the *primaries* that have the constant transverse momenta. This, without further complication, leads to a natural explanation of many of the experimental observations. We shall adopt it hereafter. We are grateful to J. Vander Velde and Knud Hansen for discussions concerning this point.

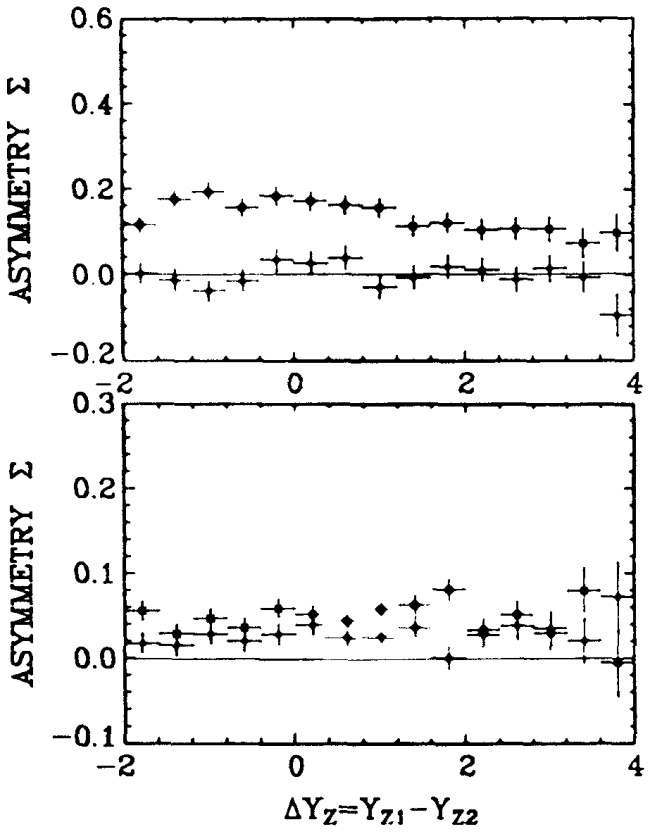


Fig. 35. Predicted behavior of the asymmetry $\Sigma(Y_1, Y_2) = (N_D(Y_2) - N_U(Y_2))/(N_U(Y_2) + N_D(Y_2))$ before (upper) and after (lower) decay of the primary mesons. The crosses (diamonds) are for the unlike (like) charge combinations and where N_U and N_D are the number of hadrons at Y_2 with $|\phi_{12}| < \frac{1}{2}\pi$ and $|\phi_{12}| > \frac{1}{2}\pi$, respectively, and ϕ_{12} is the angle between the transverse momentum vectors $k_{\perp_1}$ and $k_{\perp_2}$. The results are for particle h_1 at $Y_{z_1} = 4.0$ and are plotted *versus* $\Delta Y_z = Y_{z_1} - Y_{z_2}$. $a = 0.77$, $\alpha_{ps} = \alpha_v = 0.5$, $Y_{z_1} = (4.0, 4.4)$.

figure shows $\Sigma(Y_1, Y_2)$ for like and unlike charge combinations both before and after decay where particle one is at $Y_z = 4.0$ *versus* $\Delta Y_z = Y_{z_1} - Y_{z_2}$. (With this definition, Σ is positive for $k_{\perp_1} \cdot k_{\perp_2}$ negative.) Before decay, we see no correlation between same-sign mesons (since they cannot occur adjacent in rank) and a strong wide-range correlation between oppositely charged hadrons. Resonance decay plays a large role in determining the behavior of Σ. After decay, Σ is reduced for oppositely charged pairs and becomes non-zero and positive for same-sign hadrons. Similar transverse-momentum correlations have been observed in low-$p_\perp$ hadron-hadron collisions [24].

6. Some applications in large-$p_\perp$ hadron-hadron collisions

6.1. Large-$p_\perp$ particle ratios: comparison with FF1

Fig. 36 shows particle ratios for $\theta_{cm} = 90°$ pp collisions at large $p_\perp$ predicted using the quark-scattering model presented in FF1 but with our new quark-decay functions (analytic approximation). The quark-scattering model predicts that, except for the small smearing effects discussed in FFF, these ratios are functions only of $x_\perp = 2p_\perp/\sqrt{s}$ at fixed θ_{cm}. The new results are almost identical to those in FF1 except for the K^+/K^- ratio which is now somewhat larger (a factor of ≈ 1.8 at $x_\perp = 0.6$) due in part to (3.3). The new model, however, also allows us to investigate resonance production at large $p_\perp$. In fig. 37 predictions for ρ^0 (equal to ω^0) K^{*0}, $\bar{K}^{*0}$ and ϕ production are presented. (The ρ^0/π^0 ratio was constructed to approach one at large $x_\perp$ by our assumption that $\alpha_{ps} = \alpha_v$.) In addition, we calcu-

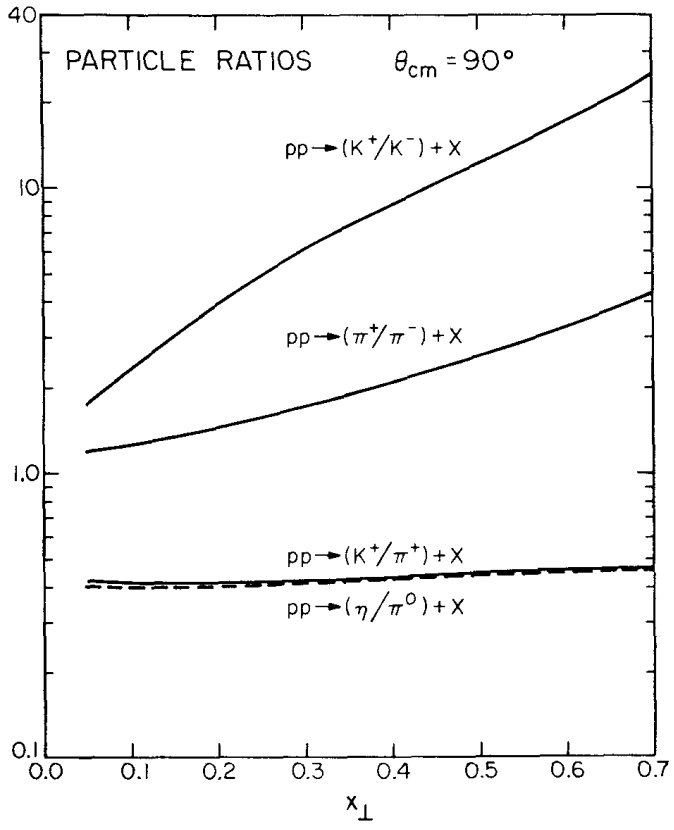


Fig. 36. Particle ratios *versus* $x_\perp = 2p_\perp/\sqrt{s}$ for $\theta_{cm} = 90°$ pp collisions at large $p_\perp$ predicted from the quark-scattering model of FF1 but using our new quark-decay functions.

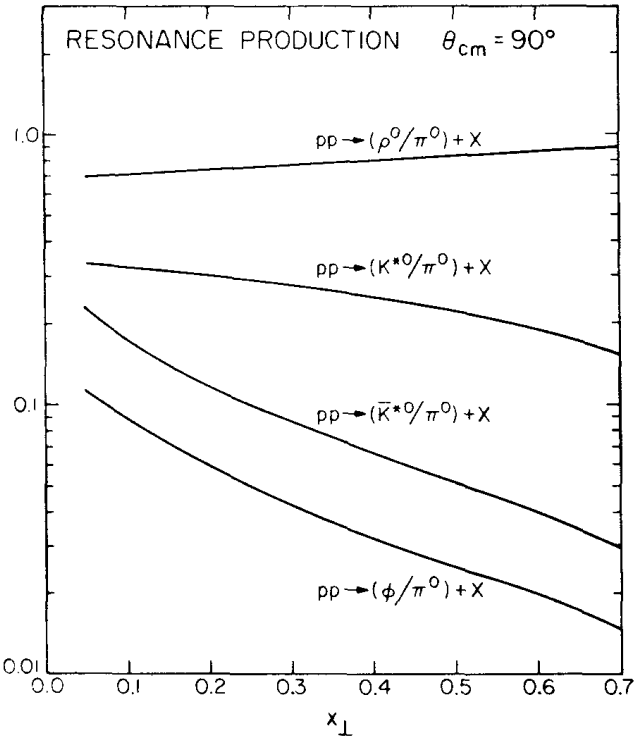


Fig. 37. Predicted ratios of resonance to π^0 production for $\theta_{cm} = 90°$ pp collisions *versus* $x_\perp = 2p_\perp/\sqrt{s}$ from the quark-scattering model of FF1 but using our new quark-decay functions.

Table 16
Comparison of large-$p_\perp$ particle ratios in pp collisions at $\theta_{cm} = 90°$ with the quark-scattering model of FF1 [4] using our new quark-decay functions, $D_q^h(z)$. Also shown are the contributions to the total π signal from ρ production and the contributions to the total K^+ signal from K^{*0} and ϕ production

Ratio	Experimental group	$x_\perp = 2p_\perp/\sqrt{s}$	Data	Predictions
π^+/π^-	C-P [28]	0.6	≃3.0	3.2
K^+/π^+	C-P	0.55	≃0.5	0.46
η/π^0	CCRS [29]	0.15	≃0.5	0.40
K^+/K^-	C-P	0.51	≃18	10.3
ρ^0/π^0	R-412 [30]	0.2	≃1.0	0.7
ϕ/π^0	C-F [31]	0.2	⩽0.06	0.06
$\rho^0 \to \pi^+/\pi^+$	R-410/13 [2,32]	0.1	≃0.05	0.08
$\rho^\pm \to \pi^0/\pi^0$	R-412	0.1–0.2	0.16	0.16
$K^{*0} \to K^+/K^+$	R-410/13	0.1–0.2	≃0.05	0.06
$\phi \to K^+/K^+$	R-410/13	0.1–0.2	0.005	0.01

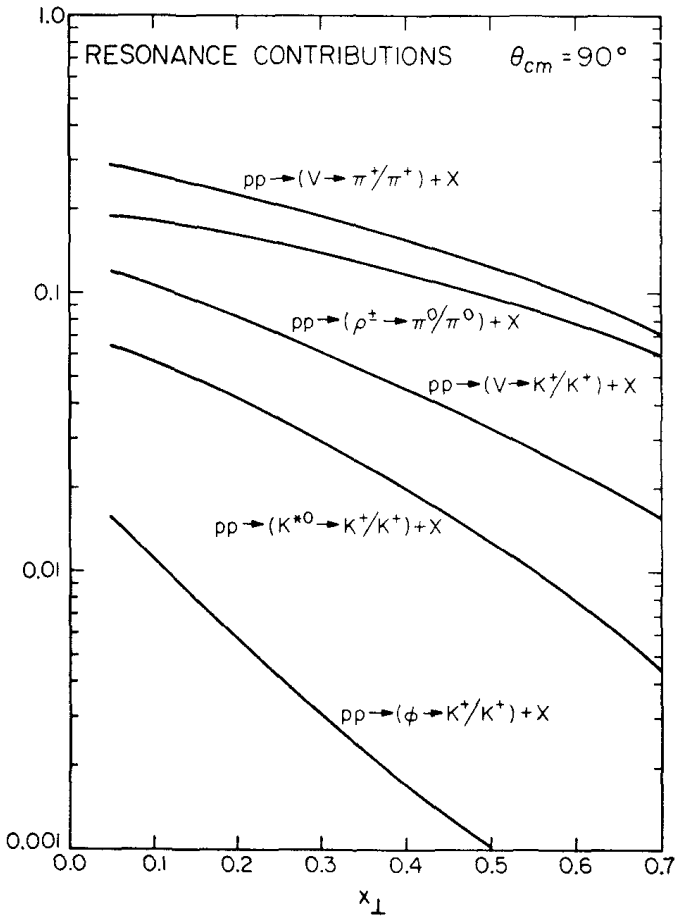


Fig. 38. Predicted contribution to the total large-$p_\perp$ meson signals for $\theta_{cm} = 90°$ pp collisions from resonance decays. The symbol V refers to the sum over all nine vector mesons and $K^{*0} \to K^+/K^+$ means the ratio of K^+'s due to K^{*0} decay to the *total* K^+ signal, etc.

late the fraction of the total large-$p_\perp$ signal that arises from the decay of various resonances. For example, fig. 38 shows that at ISR ($x_\perp = 0.1$) 27% of the total π^+ and 11% of the total K^+ signal comes from resonance decays. As $p_\perp$ increases, the contributions from resonance decays decrease. Table 16 gives a comparison of the expected values of various ratios with existing experimental data.

6.2. Same-side two-particle correlations: comparison with FFF ansatz

A distinctive feature of the quark-scattering model of FF1 is that high-$p_\perp$ particles in hadron-hadron collisions are not isolated but members of a cluster (jet) of particles representing the fragmentation of the quark. In FFF, we estimated the

number of particles to be found in the same direction as a large-$p_\perp$ trigger ("same-side" correlations) by assuming that the two-particle decay functions could be approximated in terms of the single-particle decay function by (FFF ansatz)

$$D_q^{h_1 h_2}(z_1, z_2) = D_q^{h_1}(z_1)\, D_{q_2}^{h_2}(z_2/(1-z_1))/(1-z_1)\,, \tag{6.1}$$

where h_1 is the trigger hadron so that $z_1 >> z_2$. The flavor of the quark q_2 was determined in the following manner. For what we called the "unambiguous" case where $h_1 = c\bar{a}$ contains the quark q, say q = c, then q_2 was taken to have the flavor "a". For example, if q is a u-quark and $h_1 = \pi^+$ or $h_1 = K^+$, then q_2 = d or q_2 = s, respectively. For this case, we find the FFF ansatz to agree reasonably well with the results of our new jet model ($a = 0.77$ $\alpha_{ps} = \alpha_v = 0.5$).

On the other hand, we stated in FFF that (6.1) was not appropriate for the "ambiguous" case where h_1 does not contain the quark q as a valence quark. We find that indeed our new jet model disagrees quite substantially with the FFF ansatz for the "ambiguous" case. Since the predictions in FFF are dominated by the unambiguous case, we feel that they are reliable.

However, we can now use the new jet model to estimate the number of K^+'s produced in association with a large-$p_\perp$ K^- trigger in pp collisions, which we did not attempt in FFF because this is dominated by the "ambiguous" case, and to improve the other same side correlation estimates in FFF. Before this is done, however, we must also change our handling of the transverse momentum of the particles within the quark jet to agree with the results of subsect. 5.1. We expect to discuss this elsewhere.

7. Summary and conclusions

In summary, we have found that the recursive model gives a convenient and easy way to compute properties of a quark jet in terms of only a few parameters. Although the model does not include the possibility of baryon emission and cannot represent a true physical theory of jets; the resulting structure seems very reasonable and generally consistent with observations available so far. We recommend it as a "standard" jet to help design experiments and to which experiments may be compared and contrasted.

Aside from the mean number of particles (*versus z*), to which the model is fitted anyway, the main interesting feature is the distribution of charge, $D_q^{h^+}(z) - D_q^{h^-}(z)$, or of any other property by which jets originating from u-quarks can be distinguished from those from d-quarks. We can describe this as the distribution of the hadron containing the original quark. In the model, this is widely distributed. It is far from true that in an average (unbiased) jet the hadron of largest momentum always contains the original quark (see table 11). For jets of limited momentum, individual d-quark jet events can often look just like u-quark events and a reliable way to distinguish them

event-by-event seems unavailable. They can, of course, be distinguished by their average properties. But our model, just like experiment, finds large statistical variations from event to event so that good averages require many events. (This can be seen most clearly on many of the Monte Carlo plots. Even for 40 000 jets, there are uncomfortably wide fluctuations for many quantities of interest.) One of the first things to determine experimentally is whether $D_q^{h^+}(z) - D_q^{h^-}(z)$ is really as widely distributed (see figs. 5, 8, and 12) as our model (or the previous parametrization of FF1) suggests. Preliminary experimental indications are that it is [15,21,23].

The hadrons observed in quark jets can be secondaries from the decay of higher resonances (we have included only the 1^- vector resonances). Although the model makes interesting special predictions for how these *primaries* may be correlated, we find the correlations are spread widely in rapidity, and are further considerably obscured by the correlations induced by resonance decay. This makes it hard to check the specifies of the recursive idea, but it helps us in our original purpose; to generate a standard jet that ought to look much like nature.

From our experience, we think that different jet models which have the right mean distributions of the various hadrons and which include the production of resonances in a way roughly consistent with experiment (our precise choice $\alpha_v = \alpha_{ps} = 0.5$ may require later modification) will be very difficult to distinguish experimentally. The recursive scheme is as good as any for our purpose •.

This is well illustrated by transverse-momentum distributions. Many interesting variations which have been observed (*versus* z or the two-particle mass M) in lepton and hadron experiments appear to be merely consequences of the fact that hadrons are often secondaries of primaries with much simpler properties. In fact, the primaries might all have a uniform transverse-momentum distribution.

For these reasons, we think of our jet model, not as an interesting theory to be checked by experiment, but rather as a possibly reliable guide as to what general properties might be expected experimentally. In particular, it can assist in the program of comparing hadron high-$p_\perp$ jets to lepton-generated jets.

On the other hand, quark jets are often investigated not just to be used as a tool to investigate hadron collisions, but rather as a subject of interest in itself. How are quark jets actually generated? From what we have learned, we think it will tax the ingenuity of experimenters to see behind those properties which are overshadowed by the effects of resonance decay in order to study effects more intrinsically related to the process of jet formation.

In this connection, the most fundamental experimental question is whether lepton-induced jets really have a quark origin at all. Even here we find difficulties. Properties averaged over many jets are unconvincing, for a *mean* charge, say, of $\frac{2}{3}$ can be achieved by averaging over events each of which is associated with an integral charge. The most promising method would seem to be the momentum weighted charge $Q_W(p)$ in (3.9). In principle, at least, for a jet of sufficiently large momentum, this

• For examples of other jet models see refs. [13,25–27].

is a characteristic of the object generating the jet and one which can be checked for each event separately. Unfortunately, in practice, it will be difficult to use for even 10 GeV jets (see fig. 21), if the charge is spread as widely as in our model. The method improves, but only slowly, as the jet energy increases.

We are grateful to C. Bromberg, G. Fox, and C. Quigg for numerous informative discussions. Also, we thank T. Ferbel, K. Hanson and J. Vander Velde for useful correspondence.

Note added in proof

It has been suggested to us that perhaps the simple form for $f(\eta)$ in (2.21) that we have used results in a distribution of quark charge, $D_q^{h^+}(z) - D_q^{h^-}(z)$, that is broader than actually implied by the fit to $D_u^{h^+}(z) + D_u^{h^-}(z)$ in fig. 3. For example, a form like $f(\eta) = A(1-\eta)^n\,(1 - a + 3a\eta^2)$ (normalized properly) would produce a quark charge that is isolated more toward the end of the quark (high z). Such a form produces a $zF(z)$ that tends to dip at small z, but this possibility cannot be ruled out since the data in fig. 3 cannot be used for $z \lesssim 0.3$. Hence one can view our jet model as a bit pessimistic. It results in a widely distributed quark charge and an order in rapidity and hierarchy that is quite mixed up. We must ultimately resort to experiment to see if $D_u^{h^+}(z) - D_u^{h^-}(z)$ is as widely distributed as our model suggests.

References

[1] D. Linglin, Observation of large-$p_\perp$ jets at the ISR and the problem of parton transverse momentum, Invited talk at 12th Rencontre de Moriond, March 1977.
[2] R. Moller, Jets and quantum numbers in the high-$p_\perp$ hadronic reactions at the CERN ISR, Invited talk at 12th Rencontre de Moriond, March 1977.
[3] C. Bromberg et al., Phys. Rev. Lett. 38 (1977) 1447;
E. Malamud, Comparison of hadron jets produced by π^- and p beams on hydrogen and aluminum targets, Invited talk presented at 8th Int. Symp. on multiparticle dynamics, Kaysersberg, France, June 1977.
[4] (FF1) R.D. Field and R.P. Feynman, Phys. Rev. D15 (1977) 2590.
[5] (FFF) R.P. Feynman, R.D. Field and G.C. Fox, Nucl. Phys. B128 (1977)1.
[6] A. Krzywicki and B. Petersson, Phys. Rev. D6 (1972) 924.
[7] J. Finkelstein and R.D. Peccei, Phys. Rev. D6 (1972) 2606.
[8] J.D. Bjorken, High-$p_\perp$ dynamics, Lecture given at SLAC Summer Institute on particle physics, November 1975.
[9] S.D. Ellis, M. Jacob and P.V. Landshoff, Nucl. Phys. B108 (1976) 93.
[10] A. Casher, J. Kogut and L. Susskind, Phys. Rev. Lett. 31 (1973) 792.
[11] G.R. Farrar and J.L. Rosner, Phys. Rev. D9 (1974) 2226.
[12] R.N. Cahn and E.W. Colglazier, Phys. Rev. D9 (1974) 2658.
[13] A. Seiden, Phys. Lett. 68B (1977) 157; Hadron distributions in a quark-parton model, SLAC-PUB-1962, Nucl. Phys. B, submitted.

[14] R.P. Feynman, Photon-hadron interactions (Benjamin, Reading, Mass., 1972).
[15] J.C. Vander Velde, Berkeley-Fermilab-Hawaii-Michigan Collaboration, Neutrino-proton interactions in the 15-Foot Bubble Chamber and properties of hadron jets, Invited talk presented at 4th Int. Winter Meeting on fundamental physics, Salardú, Spain, 1976.
[16] P. Pirilä, G.H. Thomas and C. Quigg, Phys. Rev. D12 (1975) 92;
C. Quigg, P. Pirilä and G.H. Thomas, Phys. Rev. Lett. 34 (1974) 290;
A. Krzywicki and D. Weingarten, Phys. Lett. 50B (1974) 265.
[17] C. Quigg, What have we learned from high-energy experiments?, Invited talk presented at the 1973 APS/DPF Meeting in Berkeley, AIP Proc. No. 14, subseries no. 6.
[18] E.L. Berger, Phys. Lett. 49B (1974) 369.
[19] T. Ferbel, Recent results from bubble-chamber experiments at Fermilab, invited talk at the SLAC Summer Institute on particle physics, SLAC report No. 179 (1974).
[20] S.J. Brodsky and J.F. Gunion, Hadronic fragmentation as a probe of the underlying dynamics of hadron collisions, UCD-77-9, SLAC-PUB-1939, Phys. Rev., to be published.
[21] M. Derrick et al., Properties of the hadronic system resulting from $\bar{\nu}$p interactions, ANL-HEP-PR-77-39, Phys. Rev. D, submitted.
[22] P. Stix et al., Transverse-momentum properties of multi-pion systems produced in high-energy collisions, Univ. of Rochester preprint UR-611.
[23] J.C. Vander Velde, Berkeley-Fermilab-Hawaii-Michigan Collaboration, Neutrino interactions in the 15-Foot Hydrogen Bubble Chamber and search for charmed particles, Invited talk presented at 12th Rencontre de Moriond, Flane, France 1977.
[24] C. Bromberg et al., Phys. Rev. D10 (1974) 3100.
[25] H. Fukuda and C. Iso, Prog. Theor. Phys. 57 (1977) 483.
[26] O. Sawada, Quark-cascade model; hadronization of quarks after a hard collision, Tohoku University preprint TU/77/166, 1977.
[27] U.P. Sukhatme, Quark jets: a quantitative description, University of Cambridge preprint, DAMTP 77/25.
[28] H. Frisch, A review of high transverse momentum processes, Invited talk at APS meeting, Brookhaven, October 1976;
D. Antreasyan et al., Chicago-Princeton preprint CP76-2.
[29] F.W. Busser et al., Phys. Lett. 55B (1975) 232.
[30] D. Darriulat et al., Nucl. Phys. B107 (1976) 429.
[31] J.A. Appel et al., Phys. Rev. Lett. 35 (1975) 9
[32] H. Bøggild, Event structures in large-$p_\perp$ hadron-hadron collisions, Invited talk given at 8th Int. Symp. on multiparticle dynamics, Kaysersberg, France, 1977.

VI. Quantum Gravity

Feynman worked seriously on the quantum theory of gravity for about a decade, beginning in the early 1950s.[1] In the introduction to paper [57] he wrote that his interest was "primarily in the relation of one part of nature to another," rather than in explaining phenomena or fitting data. In working out (absurdly small) one-loop quantum radiative corrections, he hoped that he would not be criticized "for the fact that there is no possible, practical reason for making these calculations."[2]

In fact, his work led to two sets of very useful results. The first, purely pedagogical, is embodied in the *Feynman Lectures on Gravitation* (publication [123]). In those lectures, Feynman develops the quantum field theory of a neutral massless spin 2 particle (the *graviton*), emphasizing the special features that arise, in comparison to theories of spin 0 and spin 1 particles, as well as the complications that result for a zero-mass particle in trying to create a self-consistent theory. As in the case of spin 1, masslessness results in redundant degrees of freedom, since Lorentz invariance requires that a *massless* particle can spin only along or opposite to its direction of momentum (positive or negative *chirality*), while a massive spin 2 particle may take up five different orientations relative to any arbitrary quantization direction. Eliminating the unwanted degrees of freedom is achieved by imposing certain "gauge conditions," which in the gravitational case brings about nonlinearity in the form of graviton–graviton interaction. Feynman shows that the classical limit of a properly gauged massless spin 2 theory is described by the Einstein gravitational field equations.[3]

The second result of Feynman's labors on gravitation is also pedagogical in part, with Feynman himself being one of the main students that he had in mind. For besides the elimination of unwanted degrees of freedom via the gauge conditions, a proof of self-consistency requires the quantum field theory to predict finite physical effects and satisfy certain conservation laws. For an interaction as weak as gravity, quantum predictions should be obtainable by a perturbation method, i.e. by expanding amplitudes for physical processes in a power series in the gravitational coupling constant (essentially the square root of Newton's constant). Then one must show that the successive terms in the series are finite, using the method of renormalization as required.[4]

The papers included here, namely [57], [85], and [86], address the question of renormalizing quantum gravity. Feynman begins in [57] by considering the quantum field theory of a massless spin 2 field in the "tree diagram" approximation, i.e. with neglect of graphs containing one or more closed loops. This is the theory that leads in the classical limit to the Einstein equations. Next he considers the class of diagrams with only one closed loop and develops a method for renormalizing such graphs (or, rather, the set of such graphs corresponding to a physical process, since individual graphs are not, in general, gauge-invariant.)

[1] Murray Gell-Mann recalled discussing the subject with Feynman in November 1954, and later suggesting to him that "he try the analogous problem in Yang–Mills theory." See Murray Gell-Mann, "Dick Feynman — the guy in the office down the hall," in *Most of the Good Stuff*, edited by L.M. Brown and J.S. Rigden (New York: American Institute of Physics, 1993), pp. 82–83.

[2] Paper [57], p. 697.

[3] This development is contained in publication [123]. A more complete description of Feynman's program, including citation of historical precedents, is contained in the preface to [123], written by John Preskill and Kip S. Thorne.

[4] See, e.g., *Renormalization*, edited by L.M. Brown (New York: Springer-Verlag, 1993).

His procedure involves two innovations. One is to show how loop diagrams can be written as the sum (integral) of tree diagrams. The second innovation is the use of a "ghost particle" (a fictitious vector particle) that must appear in a diagram added in the renormalization process in order to insure that the physical amplitude satisfies the principle of unitarity (i.e. conservation of probability). These two new procedures are described in paper [57], published in 1964, and are fully derived in Feynman's two contributions in the 1972 *Festschrift* to John Wheeler's sixtieth birthday.

Feynman's two innovations are probably even more important methodologically than as contributions to quantum gravity itself, insofar as they have been applied in treating the gauge theory of vector particles, the so-called Yang–Mills theory.[5] Feynman described how this came about:

> The algebraic complexity of the gravitational field equations is so great that it is not easy to do exploratory mathematical investigations and checks. Gell-Mann suggested to me that the Yang–Mills theory of vector particles with zero mass also is a non-linear theory with a gauge group and might show the same difficulties, and yet be easier to handle algebraically. This proved to be the case, and thereafter, all the work was done first with the Yang–Mills theory and then the corresponding expressions for gravitation were worked out. The connection is exceedingly close. Each difficulty and its resolution in one theory has its corresponding difficulty and resolution in the other. It becomes obvious that to find a completely satisfactory quantization of the zero-mass Yang–Mills field, is to find a completely satisfactory quantization of the general theory of relativity.[6]

The last sentence is probably an exaggeration, because while a satisfactory renormalizable theory of the Yang–Mills field does now exist, the same is not true for quantum gravity. It is also somewhat ironic that the Yang–Mills theory, regarded as a model theory for learning about how to quantize gravity, has achieved a status as a physical theory on the same level as general relativity itself. That is because the current Standard Model of strong, weak, and electromagnetic elementary particle interactions consists of leptons and quarks exchanging Yang–Mills fields, and full renormalizabilty has been proven. While Feynman did not succeed in going beyond the level of one-loop diagrams, his methods, including "ghost-particle" loops and path integrals, were the indispensable keys to the eventual success of this endeavor.

Selected Papers

[57] The quantum theory of gravitation. *Acta Physica Polonica* **24** (1963): 697–722.
[85] Closed loop and tree diagrams, in *Magic Without Magic: John Archibald Wheeler*, edited by John R. Klauder (New York: 1971).
[86] Problems in quantizing the gravitational field and the massless Yang–Mills field, in *Magic Without Magic: John Archibald Wheeler*, edited by John R. Klauder (New York: 1971).

[5] C.N. Yang and R. Mills, "Conservation of isotopic spin and a generalized gauge invariance," *Phys. Rev.* **96** (1954): 191–195.
[6] Paper [86], p. 378.

Vol. XXIV (1963) *ACTA PHYSICA POLONICA* Fasc. 6 (12)

QUANTUM THEORY OF GRAVITATION*

By R. P. Feynman

(Received July 3, 1963)

My subject is the quantum theory of gravitation. My interest in it is primarily in the relation of one part of nature to another. There's a certain irrationality to any work in gravitation, so it's hard to explain why you do any of it; for example, as far as quantum effects are concerned let us consider the effect of the gravitational attraction between an electron and a proton in a hydrogen atom; it changes the energy a little bit. Changing the energy of a quantum system means that the phase of the wave function is slowly shifted relative to what it would have been were no perturbation present. The effect of gravitation on the hydrogen atom is to shift the phase by 43 seconds of phase in every hundred times the lifetime of the universe! An atom made purely by gravitation, let us say two neutrons held together by gravitation, has a Bohr orbit of 10^8 light years. The energy of this system is 10^{-70} rydbergs. I wish to discuss here the possibility of calculating the Lamb correction to this thing, an energy, of the order 10^{-120}. This irrationality is shown also in the strange gadgets of Prof. Weber, in the absurd creations of Prof. Wheeler and other such things, because the dimensions are so peculiar. It is therefore clear that the problem we are working on is not the correct problem; the correct problem is what determines the size of gravitation? But since I am among equally irrational men I won't be criticized I hope for the fact that there is no possible, practical reason for making these calculations.

I am limiting myself to not discussing the questions of quantum geometry nor what happens when the fields are of very short wave length. I am not trying to discuss any problems which we don't already have in present quantum field theory of other fields, not that I believe that gravitation is incapable of solving the problems that we have in the present theory, but because I wish to limit my subject. I suppose that no wave lengths are shorter than one-millionth of the Compton wave length of a proton, and therefore it is legitimate to analyze everything in perturbation approximation; and I will carry out the perturbation approximation as far as I can in every direction, so that we can have as many terms as we want, which means that we can go to ten to the minus two-hundred and something rydbergs.

I am investigating this subject despite the real difficulty that there are no experiments. Therefore there is so real challenge to compute true, physical situations. And so I made

* Based on a tape-recording of Professor Feynman's lecture at the *Conference on Relativistic Theories of Gravitation*, Jabłonna, July, 1962. — Ed.

believe that there were experiments; I imagined that there were a lot of experiments and that the gravitational constant was more like the electrical constant and that they were coming up with data on the various gravitating atoms, and so forth; and that it was a challenge to calculate whether the theory agreed with the data. So that in each case I gave myself a specific physical problem; not a question, what happens in a quantized geometry, how do you define an energy tensor *etc.*, unless that question was necessary to the solution of the physical problem, so please appreciate that the plan of the attack is a succession of increasingly complex physical problems; if I could do one, then I was finished, and I went to a harder one imagining the experimenters were getting into more and more complicated situations. Also I decided not to investigate what I would call familiar difficulties. The quantum electrodynamics diverges; if this theory diverges, it's not something to be investigated unless it produces any specific difficulties associated with gravitation. In short, I was looking entirely for unfamiliar (that is, unfamiliar to meson physics) difficulties. For example, it's immediately remarked that the theory is non-linear. This is not at all an unfamiliar difficulty; the theory, for example, of the spin 1/2 particles interacting with the electromagnetic field has a coupling term $\bar{\psi}A\psi$ which involves three fields and is therefore non-linear; that's not a new thing at all. Now, I thought that this would be very easy and I'd just go ahead and do it, and here's what I planned. I started with the Lagrangian of Einstein for the interacting field of gravity and I had to make some definition for the matter since I'm dealing with real bodies and make up my mind what the matter was made of; and then later I would check whether the results that I have depend on the specific choice or they are more powerful. I can only do one example at a time; I took spin zero matter; then, since I'm going to make a perturbation theory, just as we do in quantum electrodynamics, where it is allowed (it is especially more allowed in gravity where the coupling constant is smaller), $g_{\mu\nu}$ is written as flat space as if there were no gravity plus $\varkappa$ times $h_{\mu\nu}$, where $\varkappa$ is the square root of the gravitational constant. Then, if this is substituted in the Lagrangian, one gets a big mess, which is outlined here.

$$\mathcal{L} = \frac{1}{\varkappa^2}\int R\sqrt{g}\,d\tau + \frac{1}{2}\int (\sqrt{g}\,g^{\mu\nu}\varphi_{,\mu}\varphi_{,\nu} - m^2\sqrt{g}\,\varphi^2)\,d\tau \tag{1}$$

$$g_{\mu\nu} = \delta_{\mu\nu} + \varkappa h_{\mu\nu}.$$

Substituting and expanding, and simplifying the results by a notation (a bar over a tensor means

$$\bar{x}_{\mu\nu} \equiv \frac{1}{2}(x_{\mu\nu} + x_{\nu\mu} - \delta_{\mu\nu}x_{\sigma\sigma});$$

notice that if $x_{\mu\nu}$ is symmetric, $\bar{\bar{x}}_{\mu\nu} = x_{\mu\nu}$) we get

$$\mathcal{L} = \int (h_{\mu\nu,\sigma}\bar{h}_{\mu\nu,\sigma} - 2\bar{h}_{\mu\sigma,\sigma}\bar{h}_{\mu\sigma,\sigma}) + \frac{1}{2}\int (\varphi_{,\mu}^2 - m^2\varphi^2)\,d\tau +$$

$$+ \varkappa\int\left(\bar{h}_{\mu\nu}\varphi_{,\mu}\varphi_{,\nu} - m^2\frac{1}{2}h_{\sigma\sigma}\varphi^2\right) + \varkappa\int \text{“}hhh\text{”} + \varkappa^2\int \text{“}hh\varphi\varphi\text{”} + \text{etc.} \tag{2}$$

First, there are terms which are quadratic in h; then there are terms which are quadratic

in φ, the spin zero meson field variable; then there are terms which are more complicated than quadratic; for example, here is a term with two φ's and one h, which I will write $h\varphi\varphi$ (I have written that one out, in particular); there are terms with three h's; then there are terms which involve two h's and two φ's; and so on and so on with more and more complicated terms. The first two terms are considered as the free Lagrangian of the gravitational field and of the matter.

Now we look first at what we would want to solve problem classically, we take the variation of this with respect to h, from the first term we produce a certain combination of second derivatives, and on the other side a mess involving higher orders than first. And the same with the φ, of course.

$$h_{\mu\nu,\,\sigma\sigma}-\bar{h}_{\sigma\nu,\,\sigma\mu}-\bar{h}_{\sigma\mu,\,\sigma\nu}=\bar{S}_{\mu\nu}\,(h,\,\varphi) \tag{3}$$

$$\varphi_{,\,\sigma\sigma}-m^2\varphi=\chi(\varphi,\,h). \tag{4}$$

We will speak in the following way: (3) is a wave equation, of which $S_{\mu\nu}$ is the source, just like (4) is the wave equation of which χ is the source. The problem is to solve those equations in succession, and to use the usual methods of calculation of the quantum theory. Inasmuch as I wanted to get into the minimum of difficulties, I just took a guess that I use the same plan as I do in electricity; and the plan in electricity leads to the following suggestion here: that if you have a source, you divide by the operator on the left side of (3) in momentum space to get the propagator field. So I have to solve this equation (3). But as you all know it is singular; the entire Lagrangian in the beginning was invariant under a complicated transformation of g, which in the form of h is the following; if you add to h a gradient plus more, the entire system is invariant:

$$h'_{\mu\nu}=h_{\mu\nu}+2\xi_{\mu,\,\nu}+2h_{\mu\sigma}\,\xi_{\sigma,\,\nu}+\xi_\sigma\,h_{\mu\nu,\,\sigma}, \tag{5}$$

where ξ_μ is arbitrary, and μ and ν should be made symmetric in all these equations. As a consequence of this same invariance in the complet Lagrangian one can show that the source $S_{\mu\nu}$ must have zero divergence $S_{\mu\nu,\,\nu}=0$. In fact equations (3) would not be consistent without this condition as can be seen by barring both sides and taking the divergence — the left side vanishes identically. Now, because of the invariance of the equations, in the same way that the Maxwell equations cannot be solved to get a unique vector potential — so these can't be solved and we can't get a unique propagator. But because of the invariance under the transformation some arbitrary choice of a condition on $h_{\mu\nu}$ can be made, analogous to the Lorentz condition $A_{\mu,\,\mu}=0$ in quantum electrodynamics. Making the simplest choice which I know, I make choice $\bar{h}_{\mu\sigma,\sigma}=0$. This is four conditions and I have free the four variables ξ_μ that I can adjust to make the condition satisfied by $h'_{\mu\nu}$. Then this equation (3) is very simple, because two terms in (3) fall away and all we have is that the d'Alemberian of h is equal to S. Therefore the generating field from a source $S_{\mu\nu}$ will equal the $\bar{S}_{\mu\nu}$ times $1/k^2$ in Fourier series, where k^2 is the square of the frequency, wave vector; the time part might be called the frequency ω, the space part $\boldsymbol{k}$. This is the analogue of the equation in electricity that says that the field is $1/k^2$ times the current. In the method of quantum field

theory, you have a source which generates something, and that may interact later with something else; the iteraction, of course, is $S_{\mu\nu}\ h_{\mu\nu}$; so that, I say, one source may create a potential which acts on another source. So, to take the very simplest example of two interacting systems, let's say S and S', the result would be the following: h would be generated by $S_{\mu\nu}$, and then it would interact with $S'_{\mu\nu}$, so we would get for the interaction of two systems, of two particles, the fundamental interaction that we investigate

$$\varkappa^2 \bar{S}^{\mu\nu} \frac{1}{k^2} S'_{\mu\nu}. \tag{6}$$

This represents the law of gravitational interaction expressed by means of an interchange of a virtual graviton. To understand the theory better and to see how far we already arrived we expand it out in components. Let index 4 represent the time, and 3 the direction of $\mathbf{k}$, so that 1 and 2 are transverse. The condition $k_\mu S_{\mu\nu} = 0$ becomes $\omega S_{4\nu} = kS_{3\nu}$ where k is the magnitude of $\mathbf{k}$. Using this, many of the terms involving number 3 component of S can be replaced by terms in number 4 components. After some rearranging there results

$$-2\bar{S}_{\mu\nu} \frac{1}{k^2} S'_{\mu\nu} = \frac{1}{k^2} [S_{44} S'_{44}] + \frac{1}{k^2} [S_{44}(S'_{11} + S'_{22}) + S'_{44}(S_{11} + S_{22}) +$$

$$+ S_{43} S'_{43} - 4S_{41} S'_{41} - 4S_{42} S'_{42}] + \frac{1}{k^2 - \omega + i\varepsilon} [(S_{11} - S_{22})\ (S'_{11} - S'_{22}) + 4S_{12} S'_{12}]. \tag{7}$$

There is a singular point in the last term when $\omega = k$, and to be precise we put in the $+i\varepsilon$ as is well-known from electrodynamics. You note that in the first two terms instead of one over a four-dimensional $\omega^2 - \mathbf{k}^2$ we have here just $1/\mathbf{k}^2$, the momentum itself. S_{44} is the energy density, so this first term represents the two energy densities interacting with no ω dependence which means, in the Fourier transform an interaction instantaneous in time; and $1/\mathbf{k}^2$ means $1/r$ in space, so there's an instantaneous $1/r$ interaction between masses, Newton's law. In the next term there's another instantaneous term which says that Newton's mass law should be corrected by some other components analogous to a kind of magnetic interaction (not quite analogous because the magnetic interaction in electricity already involves a $k^2 - \omega^2 + i\varepsilon$ propagator rather than just k^2. But the $k^2 - \omega^2 + it$ in gravitation comes even later and is a much smaller term which involves velocities to the fourth). So if we really wanted to do problems with atoms that were held together gravitationally it would be very easy; we would take the first term, and possibly even the second as the interaction. Being instantaneous, it can be put directly into a Schrödinger equation, analogous to the e^2/r term for electrical interaction. And that take care of gravitation to a very high accuracy, without a quantized field theory at all. However, for still higher accuracy we have to do the radiative corrections, which come from the last term.

Radiation of free gravitons corresponds to the situation that there is a pole in the propagator. There is a pole in the last term when $\omega = k$, of course, which means that the wave number and the frequency are related as for a mass zero particle. The residue of the pole, we see, is the product of two terms; which means that there are two kinds of waves, one generated by $S_{11} - S_{22}$ and the other generated by S_{12}, and so we have two kinds of trans-

verse polarized waves, that is there are two polarization states for the graviton. The linear combination $S_{11}-S_{22}\pm 2iS_{12}$ vary with angle Θ of rotation in the 1—2 plane as $e^{\pm 2i\theta}$ so the gravitaton has spin 2, component ± 2 along direction of polarization. Everything is clear directly from the expression (7); I just wanted to illustrate that the propagator (6) of quantum mechanics and all that we know about the classical situation are in evident coincidence.

In order to proceed to make specific calculations by means of diagrams, beside the propagator we need to know just what the junctions are, in other words just what the S's are for a particular problem; and I shall just illustrate how that's done in one example. It is done by looking at the non-quadratic terms in the Lagrangian I've written one out completely. This one has an h and two φ's in the Lagrangian (2). The rules of the quantum mechanics for writting this thing are to look at the h and two φ's: one φ each refers to the in and out particle, and the one h corresponds to the graviton; so we immediately see in that term a two particle interaction through a graviton (see Fig. 1). And we can immediately

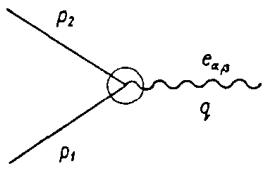


Fig. 1

read off the answer for the interaction this way: if the p_1 and p_2 are the momenta of the particles and q the momentum of the graviton; and $e_{\alpha\beta}$ is the polarization tensor of the plane wave representing the graviton, that is $h_{\alpha\beta}=e_{\alpha\beta}\,e^{iq\cdot x}$, the Fourier expansion of this term gives the amplitude for the coupling of two particles to a graviton

$$p_\mu^1 p_\nu^2 \bar{e}_{\mu\nu}-\frac{1}{2}m^2 e_{\sigma\sigma}. \tag{8}$$

So this is a coupling of matter to gravity; it is first order, and then there are higher terms; but the point I'm trying to make is that there is no mystery about what to write down — everything is perfectly clear, from the Lagrangian. We have the propagator, we have the couplings, we can write everything. A term like hhh implies a definite formula for the interaction of three gravitons; it is very complicated, and I won't write it down, but you can read it right off directly by substituting momenta for the gradients. That such a term exists is, of course, natural, because gravity interacts with any kind of energy, including its own, so if it interacts with an object-particles it will interact with gravitons; so this is the scattering of a graviton in a gravitational field, which must exist. So that everything is directly readable and all we have to do is proceed to find out if we get a sensible physics. I've already indicated that the physics of direct interactions is sensible; and I go ahead now to compute a number of other things.

To take just one example, we compute the Compton effect, or the analogue rather, of the Compton effect, in which a graviton comes in and out on a particle. The amplitude

for this is a sum of terms corresponding to the diagrams of Fig. 2. The amplidute for the first diagram of Fig. 2 is the coupling (8) times the propagator for the intermediate meson which reads $(p^2-m^2)^{-1}$, which is the Fourier transform of the equation (4) which is the propagation of the spin zero particle. Then there is another coupling of the same form as (8). We multiply these together, to get the amplitude for that diagram

$$\left(p_\mu^2 p_\nu \bar{e}^b_{\mu\nu} - \frac{1}{2} e_{\mu\mu} m^2\right) \frac{1}{p^2-m^2} \left(p_\sigma p^1_\tau e^a_{\sigma\tau} - \frac{1}{2} \bar{e}_{\sigma\sigma} m^2\right),$$

where we should substitute $p = p^2+q^b = p^1+q^a$. Then you must add similar contributions from the other diagrams.

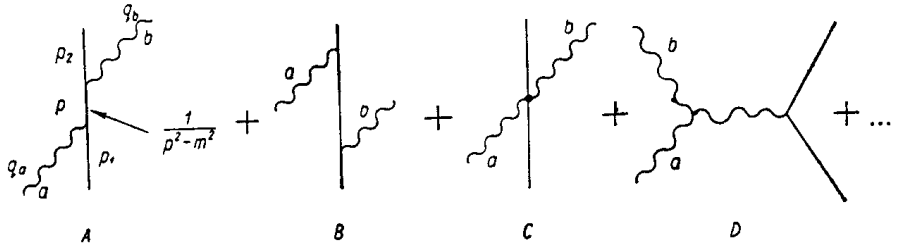


Fig. 2

The third one comes in because there are terms with two h's and two φ's in the Lagrangian. One adds the four diagrams together and gets an answer for the Compton effect. It is rather simple, and quite interesting; that it is simple is what is interesting, because the labour is fantastic in all these things.

But the thing I would like to emphasize is this; in this problem we used a certain wave $e^a_{\alpha\beta}$ for the incoming graviton number "a" say; the question is could we use a different one? According to the theory, it should really be invariant under coordinate transformations and so on, but what it corresponds to here is the analogue of gauge invariance, that you can add to the potential a gradient (see (5)). And therefore it should be that if I changed $e_{\alpha\beta}$ of a particular graviton to $e_{\alpha\beta}+q_\alpha\xi_\beta$ where ξ is arbitrary, and q_α is the momentum of the graviton, there should be no change in the physics. In short, the amplitude should be unchanged; and it is. The amplitude for this particular process is what I call gauge-invariant, or coordinate-transforming invariant. At first sight this is somewhat puzzling, because you would have expected that the invariance law of the whole thing is more complicated, including the last two terms in (5), which I seem to have omitted. But those terms have been included; you see asymptotically all you have to do is worry about the second term, the last two in h's times ξ's are in fact generated by the last diagram, Fig. 2D; when I put a gradient in here for this one, what this means is if I put for the incoming wave a pure gradient, I should get zero. If I put the gradient $q_\alpha\xi_\beta$ in for $e^a_{\alpha\beta}$ on this term D, I get a coupling between ξ and the other field $e^a_{\alpha\beta}$ because of the three graviton coupling. The result, as far as the matter line is concerned is that it is acted on in first order by a resultant field $e^b_{\mu\sigma}\xi_\sigma q^a_\nu +$ $+ \frac{1}{2} q^b_\sigma e_{\mu\nu}\xi_\sigma$ which is just the last two terms in (5). The rule is that the field which acts on the

matter itself must be invariant the way described by (5); but here in Fig. 2 I've already calculated all the corrections, the generator and all the necessary non-linear modifications if I take all the diagrams into account. In short, asymptotically far away if I include all kinds of diagrams such as D, the invariance need be checked only for a pure gradient added to an incoming wave. It takes care of the non-linearities by calculating them through the interaction.

I would like, now, to emphasize one more point that is very important for our later discussion. If I add a gradient, I said, the result was zero. Let's call a the one graviton coming in and b the other one in every diagram. The result is zero if I use a gradient for a, only if b is a free graviton with no source; that is if it is either really an honest graviton with $(q^b)^2 = 0$, or a pure potential, which is a solution of the free wave equation. That is unlike electrodynamics, where the field b could have been any potential at all and adding a gradient to a would have made no difference. But in gravity, it must be that b is a pure wave; the reason is very simple. There is no way to avoid this by changing any propagators; this is not a disease — there is a physical reason. The reason can be seen as follows: If this b had a source let me modify my diagrams to show the source of b, suppose some other matter particle made the b, so we add onto each b line a matter line at the end, like Fig. 3a. (*E.g.* Fig. 2a becomes Fig. 3b *etc.*)

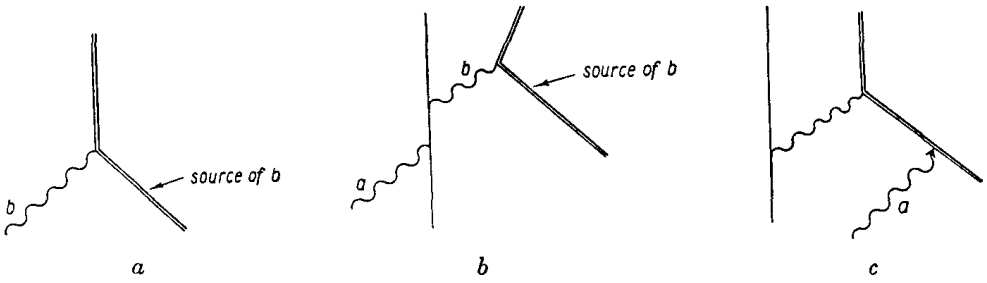


Fig. 3

Now, if b isn't a free wave, but it had a source, the situation is this. If this "a" field is taken as a gradient field which operates everywhere on everything in the diagram it should give zero. But we forgot something; there's another type of diagram, if the "a" is supposed to act on everything, one of which looks like Fig. 3c, in which the "a" itself acts on the source of b and then b comes over to interact with the original matter. In other words, among all the diagrams where there is a source, there's also these of type 3c. The sum of all diagrams is zero; but the sum of those like Fig. 2 without those of type 3c is not zero, and therefore if I were to just calculate the diagrams of Fig. 2 and forget about the source of b and then put a gradient in for "a" the result cannot be zero, but must be getting ready to cancel the terms from the likes of 3c when I do it right. That will turn out to be an important point to emphasize. I have done a lot of problems like this, without closed loops but I won't bore you with all the problems and answers; there's nothing new, I mean nothing interesting, in the sense that no apparent difficulties arise.

However, the next step is to take situations in which we have what we call closed loops, or rings, or circuits, in which not all momenta of the problem are defined. Let me just men-

tion something. I've analyzed this method both by doing a number of problems, and by a mathematical high-class elegant technique — I can do high class mathematics too, but I don't believe in it, that's the difference. I have to check it in a problem. I can prove that no matter how complicated the problem is, if you take it in the order in which there are no rings, in which every momentum is determined, the invariance is satisfied, the system is independent of what choice I made of gauge and of the propagator I made in the beginning; and everything is all right, there are no difficulties. I emphasize that this contains all the classical cases, and so I 'm really saying there are no difficulties in the classical gravitation theory. This is not meant as a grand discovery, because after all, you've been worrying about all these difficulties that I say don't exist, but only for you to get an idea of the calibration — what I mean by difficulties! If we take the next case, let's say the interaction of two particles in a higher order, then you get diagrams of which I'll only begin to write a few of them. One that looks like this in which two gravitons are exchanged,

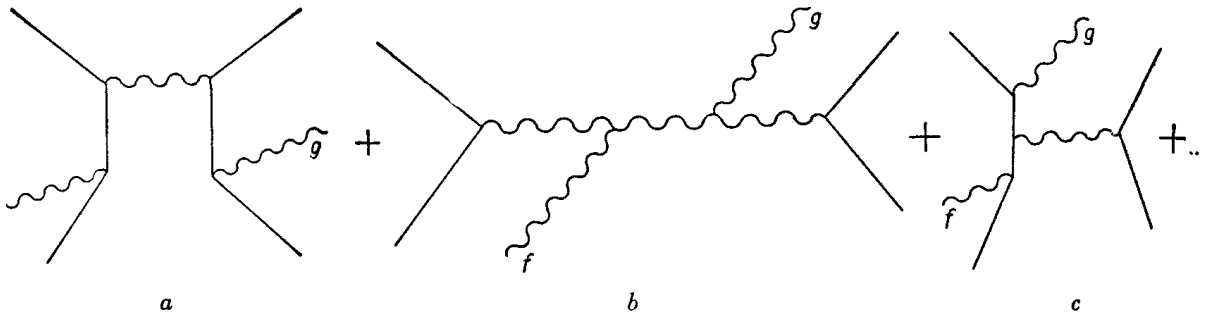


Fig. 4

or, for instance, a graviton gets split into two gravitons and then come back — these are only the beginning of a whole series of frightening-looking pictures, which correspond to the problem of calculating the Lamb shift, or the radiative corrections to the hydrogen atom. When I tried to do this, I did it in a straightforward way, following all the rules, putting in the propagator $1/k^2$, and so on. I had some difficulties, the thing didn't look gauge invariant but that had to do with the way I was making the cutoffs, because the stuff is infinite. Shortage of time does'nt permit me to explain the way I got around all those things, because in spite of getting around all those things the result is nevertheless definitely incorrect. It's gauge-invariant, it's perfectly O.K. looking, but it is definitely incorrect. The reason I knew it was incorrect is the following. In order to get it gauge-invariant, I had to do a lot of pushing and pulling, and I got the feeling that the thing might not be unique. I figured that maybe somebody else could do it another way or something, and I was rather suspicious, so I tried to get more tests for it; and a student of mine, by the name of Yura, tested to see if it was unitary; and what that means is the following: Let me take instead of this scattering problem, a problem of Fig. 4 in which time runs vertically, a problem which gives the same diagrams but in which time is running horizontally, which is the annihilation of a pair, to produce another pair, and we are calculating second order corrections to that problem. Let's suppose for simplicity that in the final state the pair is in the same state as before.

Then, adding all these diagrams gives the amplitude that if you have a pair, particle and antiparticle, they annihilate and recreate themselves; in other words it's the amplitude that the pair is still in the same state as a function of time. The amplitude to remain in the same state for a time T in general is of the form

$$e^{-i\left(E_0 - i\frac{\gamma}{2}\right)T}$$

you see that the imaginary part of the phase goes as $e^{-\frac{\gamma}{2}T}$; which means that the probability of being in a state must decrease with time. Why does the probability decrease in time? Because there's another possibility, namely, these two objects could come together, annihilate, and produce a real pair of gravitons. Therefore, it is necessary that this decay rate of the closed loop diagrams in Fig. 4 that I obtain by directly finding the imaginary part of the sum agrees with another thing I can calculate independently, without looking at the closed loop diagrams. Namely, what is the rate at which a particle and antiparticle annihilate into two gravitons? And this is very easy to calculate (same set of diagrams as Fig. 2, only turned on its side). I calculated this rate from Fig. 2, checked whether this rate agrees with the rate at which the probability of the two particles staying the same decreases (imaginary part of Fig. 4), and it does not check. Somethin'gs the matter.

This made me investigate the entire subject in great detail to find out what the trouble is. I discovered in the process two things. First, I discovered a number of theorems, which as far as I know are new, which relate closed loop diagrams and diagrams without closed loop diagrams (I shall call the latter diagrams "trees"). The unitarity relation which I have just been describing, is one connection between a closed loop diagram and a tree; but I found a whole lot of other ones, and this gives me more tests on my machinery. So let me just tell you a little bit about this theorem, which gives other rules. It is rather interesting. As a matter of fact, I proved that if you have a diagram with rings in it there are enough theorems altogether, so that you can express any diagram with circuits completely in terms of diagrams with trees and with all momenta for tree diagrams in physically attainable regions and on the mass shell. The demonstration is remarkably easy. There are several ways of demonstrating it; I'll only chose one. Things propagate from one place to another, as I said, with amplitude $1/k^2$. When translated into space, that's a certain propagation function which you might call $K_+(1, 2)$, a function of two positions, 1, 2, in space-time. It represents, in the past, incoming waves and in the future, it represents outgoing waves; so you have waves come in and out; and that's the conventional propagator, with the $i\varepsilon$ and so on, as usually represented. However, this is only a solution of the propagators's equation, the wave equation I mean; it is a special solution, as you all know. There are other solutions; for instance there is a solution which is purely retarded, which I'll call K_{ret} and which exists only inside the future light-cone. Now, if you have two Green's functions for the same equation they must differ by some solution of the homogeneous equation, say K_x. That means K_x is a solution of the free wave equation and $K_+ = K_{\text{ret}} + K_x$. In a ring like Fig. 4a we have a whole product of these K_+'s. For example, for four points 1, 2, 3, 4 in a ring we have a product like this: $K_+(1, 2)K_+(2, 3)K_+(3, 4)K_+(4, 1)$ (all K's are not the same, some of them belong to the gravitons and some are propagators for the particles and so on).

706

But now let us see what happens if we were to replace one (or more) of these K_+ by K_x, say $K_+(1, 2)$ is $K_x(1, 2)$? Then between 1, 2 we have just free particles, you've broken the ring; you've got an open diagram, because K_x is free wave solution, and this means it's an integral over all real momenta of free particles, on the mass shell and perfectly honest. Therefore if we replace one of K_+ by K_x then that particular line is opened; and the process is changed to one in which there is a forward scattering of an extra particle; there's a fake particle that belongs to this propagator that has to be integrated over, but it's a free diagram — it is now a tree, and therefore perfectly definite and unique to calculate. But I said that I could open every diagram; the reason is this. First I note that if I put K_{ret} for every K in a ring, I get zero

$$K_{\text{ret}}(1, 2)K_{\text{ret}}(2, 3)K_{\text{ret}}(3, 4)K_{\text{ret}}(4, 1) = 0 \tag{9}$$

for to be non zero t_1 must be greater than t_2, $t_2 > t_3$, $t_3 > t_4$ and $t_4 > t_1$ which is impossible. Now make the substitution $K_{\text{ret}} = K_+ - K_x$ in (9). You get either all K_+ in each factor, which is the closed loop we want; or at least one K_x, which are represented by tree diagrams. Since the sum is zero, closed loops can be represented as integrals over tree diagrams. I was surprised I had never noticed this thing before.

Well, then I checked whether these diagrams of Fig. 4 when opened into trees agreed with the theorem. I mean I hoped that the theorem proved for other meson theories would agree in principle for the gravity case, such that on opening a virtual graviton line the tree would correspond to forward scattering of free graviton waves. And it does not work in the gravity case. But, you say, how could it fail, after you just demonstrated that it ought to work? The reason it fails is the following: This argument has to do with the position of the poles in the propagators; a typical propagator is a factor $1/(k^2 - m^2 + i\varepsilon)$, the $+i\varepsilon$ due to the poles, and all I'm doing here is changing the rule about the poles and picking up an extra delta function $\delta(k^2 - m^2)$ as a consequence, which is the free wave coming in and out. What I want these free waves to represent in the gravity case are physical gravitons and not something wrong. They do represent waves of $q^2 = 0$ of course, but, as it turns out, not with the correct polarization to be free gravitons. I'd like to show it. It has to do with the numerator, not the denominator. You see the propagator that I wrote before, which was $S_{\mu\nu}$ times $1/(k^2 + i\varepsilon)$ times $\bar{S}'_{\mu\nu}$, is being replaced by $S_{\mu\nu}\delta(q^2)\bar{S}'_{\mu\nu}$. Now when I make $q^2 = 0$ I have a free wave instead of arbitrary momentum. This should be a real graviton or else there's going to be physical trouble. It ins't; although it is of zero momentum, it is not transverse. It does not make any difference in understanding the point so forget one index in $S_{\mu\nu}$ — it's a lot of extra work to carry the other index so just imagine there's one index: $S_\mu S_\mu \delta(q^2)$. This combination $S_\mu S'_\mu$, is $S_4S'_4 - S_3S'_3 - S_1S'_1 - S_2S'_2$, where 4 is the time and 3 is the direction, say, of momentum of the four-vector q. Then 1 and 2 are transverse, and those are the only two we want. (Please appreciate I removed one index — I can make it more elaborate, but it is the same idea.) That is we want only $-S_1S'_1 - S_2S'_2$ instead of the sum over four. Now what about this extra term $S_4S'_4 - S_3S'_3$? Well, it is $S_4 - S_3$ times $S'_4 + S'_3$ plus $S_4 + S_3$ times $S'_4 - S'_3$. But $S_4 - S_3$ is proprtional to $q_\mu S_\mu$ (suppressing one index) because q_4 in this notation is the frequency and equals q_3, if we assume the 3-direction is the direction of the momentum. So $S_4 - S_3$ is the response of the system to a gradient

potential, which we proved was zero in our invariance discussion. Therefore, we have shown $(S_4-S_3)/(S_4'+S_3') = 0$ and this should be accounted for by purely transverse wave contributions. But it ins't, and it isn't because *the proof that the response to a gradient potential is zero required that the other particle that was interacting was an honest free graviton.* And four plus three in $S_4'+S_3'$ is not honest — it's not transverse, it is not a correct kind of graviton. You see, the only way you can get a polarization 4+5 going in the 4—3 direction is to have what I call longitudinal response; it's not a transverse wave. Such a wave could only be generated by an artificial source here of some silly kind; it is not a free wave. When there's an artificial source for one graviton, even the another is a pure gradient, the sum of all the diagrams does not give zero. If the beam is not exactly that of a free wave, perfectly transverse and everything, the argument that the gradient has to be zero must fail, for the reason outlined previously.

Although this gradient for S_4-S_3 is what I want and I hoped it was going to be zero I forgot that the other end of it — $S_4'+S_3'$ is a funny wave which is not a gradient, and which is not a free wave — and therefore you do not get zero and should not get zero, and something is fundamentally wrong.

Incidentally I investigated further and discovered another very interesting point. There is another theory, more well-known to meson physicists, called the Yang-Mills theory, and I take the one with zero mass; it is a special theory that has never been investigated in great detail. It is very analogous to gravitation; instead of the coordinate transformation group being the source of everything, it's the isotopic spin rotation group that's the source of everything. It is a non-linear theory, that's like the gravitation theory, and so forth. At the suggestion of Gell-Mann I looked at the theory of Yang-Mills with zero mass, which has a kind of gauge group and everything the same; and found exactly the same difficulty. And therefore in meson theory it was not strictly unknown difficulty, because it should have been noticed by meson physicists who had been fooling around the Yang-Mills theory. They had not noticed it because they're practical, and the Yang-Mills theory with zero mass obviously does not exist, because a zero mass field would be obvious; it would come out of nuclei right away. So they didn't take the case of zero mass and investigate it carefully. But this disease which I discovered here is a disease which exist in other theories. So at least there is one good thing: gravity isn't alone in this difficulty. This observation that Yang-Mills was also in trouble was of very great advantage to me; it made everything much easier in trying to straighten out the troubles of the preceding paragraph, for several reasons. The main reason is if you have two examples of the same disease, then there are many things you don't worry about. You see, if there is something different in the two theories it is not caused by that. For example, for gravity, in front of the second derivatives of $g_{\mu\nu}$ in the Lagrangian there are other g's, the field itself. I kept worrying something was going to happen from that. In the Yang-Mills theory this is not so, that's not the cause of the trouble, and so on. That's one advantage — it limits the number of possibilities. And the second great advantage was that the Yang-Mills theory is enormously easier to compute with than the gravity theory, and therefore I continued most of my investigations on the Yang-Mills theory, with the idea, if I ever cure that one, I'll turn around and cure the other. Because I can demonstrate one thing; line for line it's a translation like music transcribed to a different

score; everything has its analogue precisely, so it is a very good example to work with. Incidentally, to give you some idea of the difference in order to calculate this diagram Fig. 4b the Yang-Mills case took me about a day; to calculate the diagram in the case of gravitation I tried again and again and was never able to do it; and it was finally put on a computing machine—I don't mean the arithmetic, I mean the algebra of all the terms coming in, just the algebra; I did the integrals myself later, but the algebra of the thing was done on a machine by John Matthews, so I couldn't have done it by hand. In fact, I think it's historically interesting that it's the first problem in algebra that I know of that was done on a machine that has not been done by hand.

Well, what then, now you have the difficulty; how do you cure it? Well I tried the following idea: I assumed the tree theorem to be true, and used it in reverse. If every closed ring diagram can be expressed as trees, and if trees produce no trouble and can be computed, then all you have to do is to say that the closed loop diagram is the sum of the corresponding tree diagrams, that it should be. Finally in each tree diagram for which a graviton line has been opened, take only real transverse graviton to represent that term. This then serves as the definition of how to calculate closed-loop diagrams; the old rules, involving a propagator $1/k^2+i\varepsilon$ *etc.* being superseded. The advantage of this is, first, that it will be gauge invariant, second, it will be unitary, because unitarity is a relation between a closed diagram and an open one, and is one of the class of relations I was talking about, so there's no difficulty. And third, it's completely unique as to what the answer is; there's no arbitrary fiddling around with different gauges and so forth, in the inside ring as there was before. So that's the plan.

Now, the plan requires, however, one more point. It's true that we proved here that every ring diagram can be broken up into a whole lot of trees; but, a given tree is not gauge invariant. For instance the tree diagram of Fig. 2A is not. Each one of the four diagrams of Fig. 2 is not gauge-invariant, nor is any combination of them except the sum of all four. So the thing is the following. Suppose I take all the processes, all of them that belong together in a given order; for example, all the diagrams of fourth order, of which Fig. 4 illustrates three; I break the whole mess into trees, lots of trees. Then I must gather

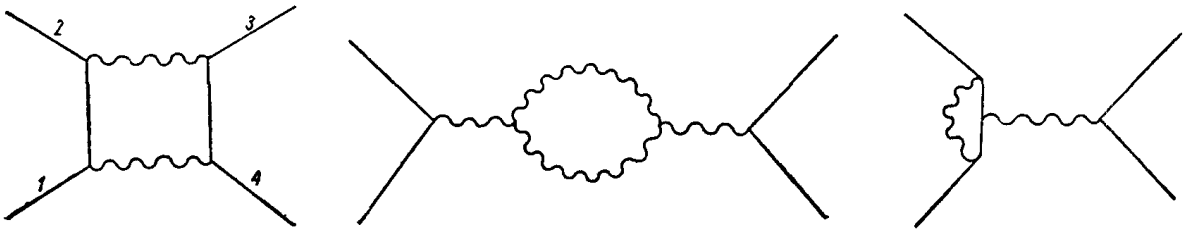


Fig. 5

the trees into baskets again, so that each basket contains the total of all of the diagrams of some specific process (for example the four diagrams of Fig. 2), you see, not just some particular tree diagram but the complet set for some process. The business of gathering the tree diagrams together in bunches representing all diagrams for complet processes is important, for only such a complet set is gauge invariant. The question is: Will any odd tree dia-

grams be left out or can they all be gathered into processes? The question is: Can we express the closed ring diagrams for some process into a sum over various other processes of tree diagrams for these processes?

Well, in the case with one ring only, I am sure it can be done, I proved it can be done and I have done it and it's all fine. And therefore the problem with one ring is fundamentally solved; because we say, you express it in terms of open parts, you find the processes that they correspond to, compute each process and add them together.

You might be interested in what the rule is for one ring; it's the sum of several pieces: first it is the sum of all the processes which you get in the lower order, in which you scatter one extra particle from the system. For instance, in Fig. 4 we have the rings for two particles scattering. There is no external graviton but there are two internal ones; now we compute in the same order a new problem in which there are two particles scattering, but while that's happening another particle, for example a graviton scatters forward. Some of the diagrams for this are illustrated in Fig. 5. State f the same state as g; so another graviton comes in and is scattered forward. In other words we do the forward scattering of an extra graviton. In addition, from breaking matter lines we have terms for the forward scattering of an extra positron, plus the forward scattering of an extra electron, and so on; one adds the forward scattering of every possible extra particle together. That is the first contribution. But when you break up the trees, you also sometimes break two lines, and then you get diagrams like Fig. 6 with two extra particles scattering (here a graviton and electron) so it turns out you must now subtract all the diagrams with two extra particles of all kinds scattering. Then add all diagrams with 3 extra particles scattering and so on. It's a nice rule, its's quite beautiful; it took me quite a while to find; I have other proofs for orther cases that are easy to understand.

Now, the next thing that anybody would ask which is a natural, interesting thing to ask, is this. Is it possible to go back and to find the rule by which you could have integrated the closed rings directly? In other words, change the rule for integrating the closed rings, so that when you integrate them in a more natural fashion, with the new method, it will

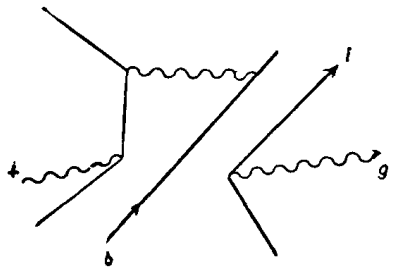


Fig. 6

give the same answer as this unique, absolute, definite thing of the trees. It's not necessary to do this, because, of course, I've defined everything; but it's of great interest to do this, because maybe I'll understand what I did wrong before. So I investigated that in detail. It turns out there are two changes that have to be made — it's a little hard to explain in

terms of the gravitation of which I'll only tell about one. Well, I'll try to explain the other, but it might cause some confusion. Because I have to explain in general what I'm doing when I do a ring. Most what it corresponds to is this: first you subtract from the Lagrangian this

$$\int \sqrt{g}\, \bar{H}^{\mu\nu}{}_{;\nu} H^{\sigma}_{\mu;\sigma}\, d\tau.$$

In that way the equation of motion that results is non-singular any more. Let me write what it really is so that there's no trouble. You say to me what is this, there's a g in it and an H in it? Yes. In doing a ring, there's a field variation over which you're integrating, which I call H; and there's a g — which is the representative of all the outside disturbances which can be summarized as being an effective external field g. And so you add to the complicated Lagrangian that you get in the ordinary way an extra term, which makes it no longer singular. That's the first thing; I found it out by trial and error before, when I made it gauge invariant. But then secondly, you must subtract from the answer, the result that you get by imagining that in the ring which involves only a graviton going around, instead you calculate with a different particle going around, an artificial, dopey particle is coupled to it. It's a vector particle, artificially coupled to the external field, so designed as to correct the error in this one. The forms are evidently invariant, as far as your g-space is concerned; these are like tensors in the g world; and therefore it's clear that my answers are gauge invariant or coordinate transformable, and all that's necessary. But are also quantum-mechanically satisfactory in the sense that they are unitary.

Now, the next question is, what happens when there are two or more loops? Since I only got this completely straightened out a week before I came here, I haven't had time to inwestigate the case of 2 or more loops to my own satisfaction. The preliminary investigations that I have made do not indicate that it's going to be possible so easily gather the things into the right barrels. It's surprising, I can't understand it; when you gather the trees into processes, there seems to be some loose trees, extra trees. I don't understand them at the moment, and I therefore do not claim that this method of quantization can be obviously and evidently carried on to the next order. In short, therefore, we are still not sure, of the radiative corrections to the radiative corrections to the Lamb shift, the uncertainty lies in energies of the order of magnitude of 10^{-255} rydbergs. I can therefore relax from the problem, and say: for all practical purposes everything is all right. In the meantime, unfortunately, although I could retire from the field and leave you experts who are used to working in gravitation to worry about this matter, I can't retire on the claim that the number is so small and that the thing is now really irrational, if it was not irrational before. Because, unfortunately, I also discovered in the process that the trouble is present in the Yang-Mills theory; and secondly I have incidentally discovered a tree-ring connection which is of very great interest and importance in the meson theories and so on. And so I'm stuck to have to continue this investigation, and of course you all appreciate that this is the secret reason for doing any work, no matter how absurd and irrational and academic it looks; we all realize that no matter how small a thing is, if it has physical interest and is thought about carefully enough, you're bound to think of something that's good for something else.

DISCUSSION

Møller: May I, as a non-expert, ask you a very simple and perhaps foolish question. Is this theory really Einstein's theory of gravitation in the sense that if you would have here many gravitons the equations would go over into the usual field equations of Einstein?

Feynman: Absolutely.

Møller: You are quite sure about it?

Feynman: Yes, in fact when I work out the fields and I don't say in what order I'm working, I have to do it in an abstract manner which includes any number of gravitons; and then the formulas are definitely related to the general theory's formulas; and the invariance is the same; things like this that you see labelled as loops are very typical quantum-mechanical things; but even here you see a tendency to write things with the right derivatives, gauge invariant and everything. No, there's no question that the thing is the Einsteinian theory. The classical limit of this theory that I'm working on now is a non-linear theory exactly the same as the Einsteinian equations. One thing is to prove it by equations; the other is to check it by calculations. I have mathematically proven to myself so many things that aren't true. I'm lousy at proving things — I always make a mistake. I don't notice when I'm doing a path integral over an infinite number of variables that the Lagrangian does not depend upon one of them, the integral is infinite and I've got a ratio of two infinities and I could get a different answer. And I don't notice in the morass of things that something, a little limit or sign, goes wrong. So I always have to check with calculations; and I'm very poor at calculations — I always get the wrong answer. So it's a lot of work in these things. But I've done two things. I checked it by the mathmatics, that the forms of the mathematical equations are the same; and then I checked it by doing a considerable number of problems in quantum mechanics, such as the rate of radiation from a double star held together by quantum-mechanical force, in several orders and so on, and it gives the same answer in the limit as the corresponding classical problem. Or the gravitational radiation when two stars — excuse me, two particles — go by each other, to any order you want (not for stars, then they have to be particles of specified properties; because obviously the rate of radiation of the gravity depends on the give of the starstides are produced). If you do a real problem with real physical things in in then I'm sure we have the right method that belongs to the gravity theory. There's no question about that. It can't take care of the cosmological problem, in which you have matter out to infinity, or that the space is curved at infinity. It could be done I'm sure, but I haven't investigated it. I used as a background a flat one way out at infinity.

Møller: But you say you are not sure it is renormalizable.

Feynman: I'm not sure, no.

Møller: In the limit of large number of gravitons this would not matter?

Feynman: Well, no; you see, there is still a classical electrodynamics; and it's not got to do with the renormalizability of quantum electrodynamics. The infinities come in different places. It's not a related problem.

Rosen: I'm not sure of this, not being one of the experts; but I have the impression that because of the non-linearity of the Einstein equations there exists a difficulty of the

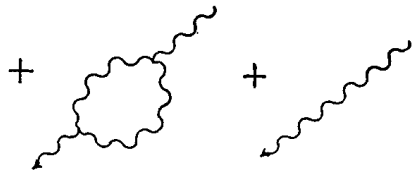
712

following kind. If the linear equations have a solution in the form of an infinite plane monochromatic wave, there does not seem to correspond to that a more exact solution; because you get piling up of energies in space and the solution then diverges at infinity. Could that have any bearing on the accuracy of this kind of calculation?

Feynman: No, I take that into account by a series of corrections. A single graviton is not the same thing as an infinite gravitational wave, because there's a limited energy in it. There's only one $\hbar\omega$.

Rosen: But you're using a momentum expansion which involves infinite waves.

Feynman: Yes, there are corrections. You see what happens if one calculates the corrections. If you have here a graviton coming in this way, then there are corrections for such a ring as this and so on. And these produce first, a divergence as usual; but second, a term in the logarithm of q^2; which means that if this thing is absolutely a free plane



wave, there's no meaning to the correction. So it must be understood in this way, that the thing was emitted some time far in the past, and is going to be absorbed some time in the future; and has not absolutely been going on forever. Then there's a very small coefficient in front of the logarithm and then for any reasonable q^2, like the diameter of the universe or something, I can still get a sensible answer; this is the shadow of the phenomenon you're talking about, that the corrections to the propagation of a graviton, dependent on the logarithm of the momentum squared carried by the graviton and which would be infinite if it were really a zero momentum graviton exactly. And so a free graviton just like that does not quite exist. And this is the correction for that. Strictly we would have to work with wave packets, but they can be of very large extent compared to the wave length of the gravitons.

Anderson: I'd like to ask if you get the same difficulty in the electromagnetic case that you did in the Yang-Mills and gravitational cases?

Feynman: No, sir, you do not. Gauge invariance of diagrams such as Fig. 2 (there is no 2*D*) is satisfied whether *b* is a free wave or not. That is because photons are not the source of photons; they are uncharged.

Anderson: The other thing I would like to suggest is that in putting of things into baskets, you might be able to get easily by always only starting out with vacuum diagrams and opening those successively.

Feynman: I tried that and it didn't go successfully.

Ivanenko: If I understood you correctly, you had used in the initial presentation the transmutation of two particles into gravitons. Yes?

Feynman: It was one of the examples.

Ivanenko: Yes. This process was considered, perhaps in a preliminary manner, by ourselves and by Prof. Weber and Brill. I ask you two questions. Do you possess the effective cross-section? Can you indicate the effects for which high-energy processes play an important role?

Feynman: I never went to energies more than one billion-billion BeV. And then the cross-sections of any of these processes are infinitesimal.

Ivanenko: They increase very, very sharply with energy. Yes, because the radiation is quadrupole, so it increases sharply in contrast to the electromagnetic transmutation of an electron-positron pair.

Feynman: It increases very sharply indeed. On the other hand, it starts out so low that one has to go pretty far to get anywhere. And the distance that you have to go is involved in this thing — the thing that's the analogue of $e^2/\hbar c$ in electricity, which is 1/137 is non-existent in gravitation; it depends on the problem; this is so because of the dimensions of G. So if E is the energy of some process, then if you take $GE^2/\hbar c$ you get an equivalent to this $e^2/\hbar c$. It may be less than that, but at least it can't be any bigger than this. So in order to make this thing to be of the order of 1%, in which case the rate is similar to the rate of photon annihilation, at ordinary energies, we need the GE^2 to be of the order of hc, and as has been pointed out many times, that's an energy of the order 10^{-5} grams, which is 10^{18} BeV. You can figure out the answer right away; just take the energy that you are interested in, square, multiply by G and divide by hc; if that becomes something, then you're getting somewhere. You still might not get somewhere, because the cross-section might not go up that fast, but at least it can't get up any worse than that. So I think that in order to get an appreciable effect, you've got to go to ridiculous energies. So you either have a ridiculously small effect or a ridiculous energy.

Weber: I have a cross-section which may be a partial answer to Ivanenko's question. Could I write it on the board? We have carried out a canonical quantization, which is not as fancy as the one you have just heard about; but considering the interaction of photons and gravitons; and it turns out that even in the linear approximation that one has the possibility of the graviton production by scattering of photons in a Coulomb field. And the scattering cross-section for this case turns out to be $8\pi^2$ times the constant of gravitation times the energy of the scatterer times the thickness of the scatterer in the direction of propagation of the photon through it divided by c^4. This assumes that all of the dimensions of the scatterer are large in comparison with the wave length of the photon. We obtained this result by quantization, and noticed that it didn't have Planck's constant in it, so we turned around and calculated it classically. Now, if one puts numbers in this, one finds that the scattering cross-section of a galaxy due to a uniform magnetic field through it is 10^{28} cm^2, a much larger number than the object that you talked about. This represents a conversion of photons into gravitons of about 1 part in 10^{16}. This is of course too small to measure. Also, we considered the possibility of using this cross-section for a laboratory experiment in which one had a scatterer consisting, say of a million gauss magnetic field over something like a cubic meter. This turned out to be entirely impossible, a result in total contradiction to what has appeared in the Russian literature. In fact, the theory of fluctuations shows that for a laboratory experiment involving the production of gravitons

by scattering of photons in a Coulomb field, the scattered power has to be greater than twice the square root of kT times the photon power divided by the averaging time of the experiment. I believe that the incorrect results that have appeared in the literature have been due to the statement that ΔP has to be greater than kT over τ; dimensionally these things are the same, but order of magnitude-wise this kind of experiment for the scatterer of which I spoke requires something like 10^{50} watts. Maybe I can say something about this afternoon; I don't want to take any more time.

De Witt: I should like to ask Prof. Feynman the following questions. First, to give us a careful statement of the tree theorem; and then outline, if he can to a brief extent, the nature of the proof of the theorem for the one-loop case, whish I understand does work. And then, to also show in a little bit more detail the structure and nature of the fictitious particle needed if you want to renormalize everything directly with the loops. And if you like, do it for the Yang-Mills, if things are prettier that way.

Feynman: I usually don't find that to go into the mathematical details of proofs in a large company is a very effective way to do anything; so, although that's the question that you asked me — I'd be glad to do it — I could instead of that give a more physical explanation of why there is such a theorem; how I thought of the theorem in the first place, and things of this nature; although I do have a proof — I'm not trying to cover up.

De Witt: May we have a statement of the theorem first?

Feynman: That I do not have. I only have it for one loop, and for one loop the careful statement of the theorem is... — look, let me do it my way. First — let me tell you how I thought of this crazy thing. I was invited to Brussels to give a talk on electrodynamics — the 50th anniversary of the 1911 Solvay Conference on radiation. And I said I'd make believe I'm coming back, and I'm telling an imaginary audience of Einstein, Lorentz and so on what the answer was. In other words, there are going to be intelligent guys, and I'll tell them the answer. So I tried to explain quantum electrodynamics in a very elementary way, and started out to explain the self-energy, like the hydrogen Lamb shift. How can you explain the hydrogen Lamb shift easily? It turns out you can't at all — they didn't even know there was an atomic nucleus. But, never mind. I thought of the following. I would explain to Lorentz that his idea that he mentioned in the conference, that classically the electromagnetic field could be represented by a lot of oscillators was correct. And that Planck's idea that the oscillators are quantized was correct, and that Lorentz's suggestion, which is also in that thing, that Planck should quantize the oscillators that the field is equivalent to, was right. And it was really amusing to discover that all that was in 1911. And that the paper in which Planck concludes that the energy of each oscillator was not $n\hbar\omega$ but $(n+1/2)\hbar\omega$ which was also in that, was also right; and that this produced a difficulty, because each of the harmonic oscillators of Lorentz in each of the modes had a frequency of $\hbar\omega/2$ which is an infinite amount of energy, because there are an infinite number of modes. And that that's a serious problem in quantum electrodynamics and the first one we have to remove. And the method we use to remove it is to simply redefine the energy so that we start from a different zero, because, of course, absolute energy doesn't mean anything. (In this gravitational context, absolute energy does mean something, but it's one of the technical points I can't discuss, which did require a certain skill to get rid of, in making

a gravity theory; but never mind.) Now look — I make a little hole in the box and I let in a little bit of hydrogen gas from a reservoir; such a small amount of hydrogen gas, that the density is low enough that the index of refraction in space differs from one by an amount proportional to A, the number of atoms. With the index being somewhat changed, the frequency of all the normal modes is altered. Each normal mode has the same wavelenght as before, because it must fit into the box; but the frequencies are all altered. And therefore the $\hbar\omega$'s should all be shifted a trifle, because of the shift of index, and therefore there's a slight shift of the energy. Although we subtract $\hbar\omega/2$ for the vacuum, there's a correction when we put the gas in; and this correction is proportional to the number of atoms, and can be associated with an energy for each atom. If you say, yes, but you had that energy already when you had the gas in back in the reservoir, I say, but let us only compare the difference in energy between the $2S$ and $2P$ state. When we change the excitation of the hydrogen gas from $2S$ to $2P$ then it changes its index without removing anything; and the energy difference that is needed to change the energy from $2S$ to the $2P$ for all these atoms is not only the energy that you calculate with disregard of the zero point energy; but the fact is that the zero point energy is changed very slightly. And this very slight difference should be the Lamb effect. So I thought, it's a nice argument; the only question is, is it true. In the first place it's interesting, because as you well know the index differs from one by an amount which is proportional to the forward scattering for γ rays of momentum k and therefore that shift in energy is essentially the sum over all momentum states of the forward scattering for γ rays of momentum k. So I looked at the forward scattering and compared it with the right formula for the Lamb shift, and it was not true, of course; it's too simple an argument. But then I said, wait, I forgot something. Dirac, explained to us that there are negative energy states for the electron but that the whole sea of negative energy states is filled. And, of course, if I put the hydrogen atoms in here all those electrons in negative energy states are also ascattering off the hydrogen atoms; and therefore their states are all shifted; and therefore the energy levels of all those are shifted a tiny bit. And therefore there's shift in the energy due to those. And so there must be an additional term which is the forward scattering of positrons, which is the same as scattering of negative energy electrons. Actually, for the symmetry of things it is better to take half the case where you make the positrons the holes and the other half where you make the electrons the holes; so it should be 1/2 forward scattering by electrons, 1/2 scattering by positrons and scattering by γ rays — the sum of all those forward scattering amplitudes ought to equal the self-energy of the hydrogen atom. And thats' right. And it's simple, and it's very peculiar. The reason it's peculiar is that these forward scatterings are real processes. At last I had discovered a formula I had always wanted, which is a formula for energy differences (which are defined in terms of virtual fields) in terms of actual measurable quantities, no matter how difficult the experiment may be — I mean I have to be able to scatter these things. Many times in studying the energy difference due to electricity (I suppose) between the proton and the neutron, I had hoped for a theorem which would go something like this — this energy difference between proton and neutron must be equal to the following sum of a bunch of cross-sections for a number of processes, but all real physical processes, I don't care how hard they are to measure. So this is the beginning of such a formula. It's rather surprising.

716

It's not the same as the usual formula — it's equal to it but it's not the same. I have no formulation of the laws of quantum gravidynamics; I have a proposal on how to make the calculations. When I make the proposal on how to do the closed loops, the obvious proposal does not work; it gives non-unitarity and stuff like that. So the obvious proposal is no good; it works O.K. for trees; so how am I going to define the answer for would correspond to a ring? The one I happen to have chosen is the following: I take the ring in general for any meson theory, one closed ring can be written as equivalent to a whole lot of processes each one of which is trees. I then define, as my belief as to what the ring ought to be in the grand theory, that it's going to be also equal to the corresponding physical set of trees. When I said this is equal to this. I didn't worry about gauge or anything else; what I means was, if these weren't gravitons but photons or any other neutral object — it doesn't make any difference what they are — this theorem is right. So I suppose it's right also for real gravitons, and I suppose also that what's being scattered is only transverse and is only a real free graviton with $q^2 = 0$. Therefore, I say let this ring equal this set of trees. Every one of these terms can be completely computed — it's a tree. And it's gauge invariant; that is, if I added an extra potential on the whole thing, another outside disturbance of a type which is nothing but a coordinate transformation — in short a pure gradient wave — to the whole diagram then it comes on to all of these processes; but it makes no effect on any of them, and therefore makes no effect on the sum; and therefore I know my definition of this ring is gauge-invariant. Second, unitarity is a property of the breaking of this diagram; the imaginary part of this equals something; if you take the imaginary part of this side, it's already broken up, in fact, and you can prove immediately that it's the correct unitarity rule. Therefore it's going to be unitarity and so on and so on. And so I therefore define gravity with one ring in this way. Now what prevents me from doing it with two rings? The lack of a complet statement of what two rings is equal to in terms of processes; that is I can open the ring all right; but I can't put the pieces — the broken diagrams — back together again into complete sets that each one is a complete physical process. In other words some of them correspond to the scattering of a graviton, but leaving out some diagrams. But the scattering of a graviton leaving out diagrams is no longer gauge invariant, I mean, not evidently gauge invariant, and so the power of the whole thing collapses. I don't know what to do with it. So that's the situation; that's why it is crucial to the particular plan. There's always, of course, another way out. And that's the following (and that's what I tried to describe at the end of the talk — maybe I talked too fast): After all now I've defined what this results is equal to — by definition not that you should do a loop some way and get this, but that a loop is equal to this by defition, and I'm not going to do a loop any other way. But, of course, from a practical point of view or from the point of view purely of interest, the question is, can you come back now and calculate the ring directly by some particular mathematical shenanigans, and get the same answer as you get by adding the trees. And I found the way to do that. I have another way, in other words, to do the ring integral directly. I have to subtract something from a vector particle going around the instead of a graviton to get the answer right. So I known the rule, and I know why the rule is, and I have a proof of the rule for one loop. I have two ways of extending. I can either break this two loop diagram open and get it back into the processes, like I did with the one ring — where so far I'm stuck. Or,

I can take the rule which I found here and try to guess the generalization for any number of rings. Also stuck. But I've only had a week, gentlemen; I've only been able to straighten out the difficulty of a single ring a week ago when I got everything cleaned up. It's more than a week — I had to take a lot of time checking and checking; but I was only finished checking to make sure of everything for this conference. And of course you're always asking me about the thing I haven't had time to make sure about yet, and I'm sorry; I worked hard to be sure of something, and now you ask me about those things I haven't had time. I hoped that I would be able to get it. I still have a few irons to try; I'm not completely stuck—maybe.

DeWitt: Because of the interest of the tricky extra particle that you mentioned at the end, and its possible connection, perhaps, with some work of Dr Białynicki-Birula, have you got far enough on that so that you could repeat it with just a little more detail? The structure of it and what sort of an equation it satisfies, and what is its propagator? These are technical points, but they have an interest.

Feynman: Give me ten minutes. And let me show how the analysis of these tree diagrams, loop diagrams and all this other stuff is done mathematical way. Now I will show you that I too can write equations that nobody can understand. Before I do that I should like to say that there are a few properties that this result has that are interesting. First of all in the Yang-Mills case there also exists a theory which violates the original idea of symmetry of the isotopic spin (from which was originally invented) by the simple assumption that the particle has a mass. That means to add to the Lagrangian a term $-\mu^2 a_\mu a^\mu$ where a_μ is an isotopic vector. You add this to the Lagrangian. This destroys the gauge invariance of the theory — it's just like electrodynamics with a mass, it's no longer gauge-invariant, it's just a dirty theory. Knowing that there is no such field with zero mass people say: „let's put the mass term on". Now when you put a mass term on it is no longer gauge invariant. But then it is also no longer singular. The Lagrangian is no longer singular for the same reason that it is not invariant. And therefore everything can be solved precisely. The propagator instead of being $\delta_{\mu\nu}$ between two currents is

$$\frac{\delta_{\mu\nu}-q_\mu q_\nu/\mu^2}{q^2-\mu^2}, \tag{10}$$

where q_μ is the momentum of propagating particle. The factor $1/(q^2-\mu^2)$ is typical for mass μ but the part $-q_\mu q_\nu/\mu^2$ is an important term which can be taken to be zero in electrodynamics but it is not obvious whether it can be taken to be zero in the case of Yang-Mills theory. In fact it has been proved it cannot be taken to be zero; this propagator is used between two currents. I am using the Yang-Mills example instead of the gravity example. I really want only the case $\mu^2 = 0$, and am asking whether I can get there by first calculating finite μ^2, then taking the limit $\mu^2 = 0$.

Now, with $\mu^2 \neq 0$ this is a definite propagator and there are no ambiguities at the closed rings, the closed loops. I have no freedom, I must compute this propagator. I mean there is no reason for trouble, and there is no trouble. There is no gauge invariance either.

And of course I checked. I broke the rings and I computed by the broken ring theorem method a closed loop problem of fair complexity (which in fact was the interaction of two

electrons). I computed it by the open ring method and by the closed ring method, and of course it agreed, there is no reason that it shouldn't. It turned out that for tree diagrams you don't have to worry about this $q_\mu q_\nu/\mu^2$ term, you can drop it — but not for the closed ring — only for tree. Therefore the tree diagrams have the definite limit as μ^2 goes to zero. And yet I have the closed ring diagram which is equal to the tree diagram when the mass is anything but zero, and therefore it ought to be true that the limit as μ^2 goes to zero of the ring is equal to the case when $\mu = 0$. It sounds like a great idea why don't you define the desired $\mu^2 = 0$ theory that way? Answer: You can't put μ^2 equal zero in the form (10). You can't do it because of the $q_\mu q_\nu/\mu^2$. So it was necessary next to see if there is a way to re-express the ring diagrams, for the case with $\mu^2 \neq 0$, in a new form with a propagator different from (10), that didn't have a μ^2 in it, in such a form that you can take the limits as μ^2 goes to zero. Then that would be a new way to do the μ equal zero case; and that's the way I found the formula. I'll try to explain how to find that theory.

We start with a definite theory, the Yang-Mills theory with a mass (the reason I do that is that there's no ambiguity about what I am trying to do) and later on I take the mass to zero, then the theory works something like this. You have the Lagrangian $\mathcal{L}(A, \varphi)$ which involves the vector potential of this field and the fields φ representing the matter with which this object is interacting for zero mass, to which, for finite mass we add the term $\mu^2 A_\mu A_\mu$. This is the Lagrangian that has to be integrated and the idea is that you integrate this over all fields A and φ; and that is the answer for the amplitude of the problem

$$X = \int e^{\int \mathcal{L}(A, \varphi) d\tau + \mu^2 A_\mu A_\mu d\tau} \, DAD\varphi. \tag{11}$$

But wait, what about the initial and final conditions? You have certain particles coming in and going out. To simplify things (this is not essential) I'll just study the case that corresponds only to gravitons in and out. I'll call them gravitons and mesons even though they are vector particles. The question is first, what is the right answer if you have gravitons represented by plane waves, $A_1, A_2, A_3 \ldots$ going in (positive frequency in A_1) or out (negative frequency). You make the following field up. Let A_{asym} be defined as α times the wave function A_1 that represents the first graviton coming in a plane wave, plus β times A_2 plus γ times A_3 and so on.

$$A_{\text{asym}} = \alpha A_1 + \beta A_2 + \gamma^{A_3} \ldots, \quad A \to A_{\text{asym}}. \tag{12}$$

Then you calculate this integral (11) subject to the condition that A approaches A_{asym} at infinity. The result of this is of course a function of $\alpha, \beta, \gamma \ldots$ and so on. Then what you want for X is just the term first order in $\alpha, \beta, \gamma \ldots$ That means just one of each these gravitons coming in and out. That's the right formula for a regular theory, for meson theory, You calculate the integral subject to the asymptotic condition, when you imagine all these waves, but you take the first order perturbation with respect to each one of the incoming waves. You never let the same photon operate twice; a photon operating twice is not a photon, it is a classical wave. So you take the derivative of this with respect to α, β, γ and so on, then setting them all equal to zero. That's problem. (In general there's φ asymptotic too.)

Now the way I happened to do this is the following: Let us call A_0 the A which satisfies the classical eqatiuons of motion, which in this particular case will be

$$\left.\frac{\partial \mathcal{L}}{\partial A}\right|_{A^\circ} + \mu^2 A^\circ = 0 \tag{13}$$

I solve this subject to the condition that A_0 equals A_{asym}. In other words, I find what is the maximum or minimum — whatever it is — of the action in (11), subject to the asymptotic condition. That's the beginning of analysing this.

The next thing is to make the simple substitution $A = A_0 + B$ and put it back in equation (11). Then if you take $\mathcal{L}$ od $A_0 + B$ (if B is negligible you get $\mathcal{L}$ of A_0 and so forth) so you get something like this

$$e^{i[\mathcal{L}(A_0)+\mu^2 A_0 A_0]} \int e^{[\mathcal{L}(A^0+B)-\mathcal{L}(A^0)]+\mu^2 BB+2\mu^2 AB}\, DB. \tag{14}$$

The integral is over all B, and B must go to zero asymptotically. This business can be expanded in powers of B.

$$\mathcal{L}(A+B)-\mathcal{L}(A)+\mu^2 BB+2\mu^2 AB = \text{Quad}\,(B)+\text{Cubic}\,(B)+\ldots+\mu^2 BB. \tag{15}$$

The zeroth power B is evidently zero. The first power of B is also zero because A_0 minimized the original thing. So this starts out quadratic in B plus cubic in B plus *etc.*, that's what this is here. These quadratic forms Quad (B) and so on of course depend on A_0, the cubic form involves A_0 in some complex, maybe very complicated, locked-up mess, but as far as B is concerned it is second power and higher powers.

Now I would like to point something out. First — it turns out if you analyze it, that the contribution of the first factor here alone (if you had forgotten the intergal and called it one) is exactly the contribution of all trees to the problem. So that's like the classical theories related to trees. Next, if you drop the term cubic in B in the exponent completely and just integrated the result over DB, that corresponds to the contribution from one ring, or from two isolated rings, or three isolated rings, but not interlocked rings. If you start to include the cubic term is has to come in a second power to do anything, because of the evenness and oddness of function. And as soon as it comes in second power, the cubic term, having three of these things come together twice, makes a terrible thing like ∞ which is a double ring. So you don't get to a double ring until you bring a cubic term down to the second order. So if I disregard that and just work with this second order term Quad (B) $+\mu^2 BB$, I'm studying the contribution from one ring. If I study this I am working from the trees. And now you see I have in my hands an expression for the contribution of a ring correct in all orders no matter how many lines come in. I also have expressions for the contributions from trees and so on. I can compare them in different mathematical circumstances, and it's on this basis that I have been able to prove everything I have been able to prove relating one ring to trees.

Now, let me explain how the theorem was obtained that takes the case for the mass and for a ring. Now we have to discuss a ring, which is a formula like this

$$X = \int e^{(\text{Quadr}(B)+\mu^2 B^4)}\, DB. \tag{15}$$

720

The quadratic form involves A_0 so the answer depends on A_0 — it's some complicated functional of A_0. Anyway I won't say that all the time, I'll just remember that. We have to integrate over all B. And the difficulty is — not difficulty, but the point is — that this quadratic form in B is singular, because it came from the piece of the action that has an invariance and this invariance keeps chasing us along. And there are certain transformations of B which leave this Quard B part unchanged in first order. That transformation in the Yang-Mills theory is

$$\vec{B}'_\mu = \vec{B}_\mu + \vec{\nabla}\alpha + (\vec{\alpha}\times\vec{A}) = \vec{B}_\mu + \alpha_{;\mu}\,. \tag{16}$$

where the vectors are in isotopic spin space and α is considered as first order. This transformation leaves the quadratic form invariant so the Quad (B) thing by itself is singular. But it doesn't make any difference, because of the addition of the $\mu^2 BB$. If $\mu^2 \neq 0$, there is no problem, but if $\mu^2 \to 0$, I'd be in trouble.

I discovered that if I make this change (16) in the actual Lagrangian and carry everything up to second order it is exact, in fact because it's only second order. If I do it with the exact change, the thing isn't invariant, it is only invariant to first order in α. But if I make the substitution exactly, then I get a certain addition to the Lagrangian, in other words the Lagrangian of B' (this includes the μ^2, the Lagrangian plus the μ^2 term in B) is the Lagrangian plus the μ^2 term in B plus something like this

$$\mu^2 B_\mu \cdot \alpha_{,\mu} + \frac{1}{2}\,\mu^2 \alpha_{,\mu}\,\vec{\alpha}^{;\mu}$$

I have to explain that the semicolon is analogous to the semicolon in gravity. The semicolon derivative $X_{;\mu}$ means the ordinary derivative of X minus A cross X and that's the analogue of the Christoffel symbols. Anyway, I find out what happens to L when I make this transformation. Now comes the idea, the trick, the nonsense: you start with the following thing; you, say, suppose instead of writing the original terms down, instead of writing the original Lagrangian I were to write the following:

$$\int e^{\mathcal{L}(B) + \frac{1}{2}(B_{\mu,\mu} - \alpha_{;\mu\mu} + \mu^2\alpha)^2}\,\mathcal{D}\alpha\mathcal{D}B.$$

Now I say that the integral over α is some constant or other. So all I have done is to multiply my original integral by $\mathcal{L}$ of B (by $\mathcal{L}$ of B I mean the whole thing, I mean this whole thing is going to be $\mathcal{L}$ of B). If I can claim that when I integrate α I get something which is independent of B, which is not self-evident. If I integrate over all α it does not look as if it is independent of B — but after a moment's consideration you see that it is. Because if I can solve a certain equation, which is $\alpha^\mu_{;\mu} - \mu^2\alpha = B^\mu_\mu$, I can shift the value of α by that amount, and then this term would disappear. In other words if I can solve this, and call this solution α_0 and change α to α_0, then the B would cancel and it would only be α' here. I did it a little abstractly which is a little easier to explain, therefore, this term that I've added can be thought of as an integral of the following nature: Integral of some B, plus an operator acting on α (this complicated operator is the second derivative and so on) squared $\mathcal{D}\alpha$. And then by that substitution I've just mentioned, this becomes equal to 1/2 the operator on A

times α' squared $\mathcal{D}\alpha$, which is equal to the integral e to the one half of α times A, the operator A, times the operator A times α integrated over primed α. Now when you integrate a quadratic form, which is a quadratic with an operator like this you get one over the square root of the determinant of the operator. So this thing is one over the square root of the determinant of the operator AA. The determinant of the operator A times A is square of the determinant of A. So this is one over the determinant of the operator A, or better it is one over square root of the determinant of the operator A squared, you'll see in a minute why I like to write it in this way. In other words, when I've written this thing down I've written the answer that I want. Let's call X the unknown answer that I want. Then this is equal to X divided by this determinant's square root squared. Now comes the trick — I now make the change from B to B'. We notice that B changed to B' is simply... oh!, this is wrong, that's what's wrong, it should be just this. Now I've got it. The change from B' to B is to add something to B. Therefore to the differential of B it adds nothing, it's just shifting the B to a new value. So I make the transformation from B to B' everywhere. So then I have $d\alpha$ and dB, and now I have a new thing up here where I make use of the formula for $\mathcal{L}$ of B':

$$\mathcal{L}(B') = \mathcal{L}(B) + \mu^2 B_\mu \alpha_\mu + \frac{1}{2}\mu^2 \alpha_{,\mu}\alpha_{;\mu}$$

You see there is a certain cross term generated here and another cross term coming from expanding this out and the net result, with a little algebra here, is that becomes $\mathcal{L}$ of B, but the quadratic term doesn't cancel out and is left; there's one half of $B_{\mu,\mu}$ squared; that's from this term; the cross term here cancels the cross term in there; and then we have only the quadratic — I mean the α terms

$$\int e^{\mathcal{L}(B)+\frac{1}{2}(B_{\mu,\mu})^2}\mathcal{D}B e^{\frac{\mu^2}{2}(\alpha_{,\mu}\alpha_{;\mu}+\mu^2\alpha^2)}\mathcal{D}\alpha.$$

And the problem is now to do this integral on α; well, another miraculous thing happens. I have the operator A, but that this down thing is $\alpha A\alpha$, and therefore its result is just determinant once; or the square of this integral is equal to this determinant, or something like that. Therefore, when you get all the factors right, X, the unknown, is equal to

$$X = \left[\int e^{\mathcal{L}(B)+\frac{1}{2}(B_{\mu,\mu})^2}\mathcal{D}B\right] : \left[\int e^{\frac{\mu^2}{2}(\alpha_{,\mu}\alpha_{;\mu}+\mu^2\alpha^2)}\mathcal{D}\alpha\right].$$

Sachs: I want to ask a question about long-range hopes. Perhaps for irrational reasons people are particularly interested in those parts of the theory where is a possibility of real qualitative differences: what do the coordinates or topology mean in a quantized theory, and this kind of junk. Now I wonder if you think that this perturbation theory can eventually be jazzed up to cover also this kind of questions?

Feynman: The present theory is not a theory as it is incomplete. I do not give a rule on how to do all problems. I expect of course that if I spend more time on figuring out how to untangle the pretzels I shall be able to make it into such a theory. So let's suppose I did. Now you can ask the question would the completed job, assuming it exists, be of any interest to esoteric question about the quantization of gravity. Of course it would be, because it

would be the expression of the quantum theory; there is today no expression of the quantum theory which is consistent. You say: but it's perturbation theory. But it isn't. I worked on the thing analyzing it in the series of increasing accuracy, but that's only, obviously, when I am doing problems and checking, or doing things like I just did. But even there I haven't said how many times the vector potential A_0 is attacking the diagram, there is no limit to what order of external lines are involved in the calculation of A_0, for example. And so if I get my general theorem for all orders, I'll have some kind of a formulation. The fact is, that in such things as electrodynamics and other theories, it has not been possible to figure out the consequences of the quantum field theory in the case of strong interactions, because of technical difficulties which are not technical difficulties just of the gravitation theory, but exist all over the quantum field theory. I do not expect that the gravitational problems will be any easier in that region than they are in any other field theory, so I can say very little there. But at least one should certainly formulate the theory that you're trying to calculate first, and then find out what the consequences are, before trying to do it the other way round. So I think that you'll be frustrated by the difficulties that do appear whenever any theory diverges. On other hand, if you ask about the physical significance of the quantization of geometry, in other words about the philosophy behind it; what happens to the metric, and all such questions, those I believe will be answerable, yes. I think you would be able to figure out the physics of it afterwards, but I won't to think about that until I have it completely formulated, I don't want to start to work out the anser to something unless I know what the equation is I am trying to analyze. But I dont' have the doubt that you will be able to do something, because after all you are describing the phenomena that you would expect, and if you deseribe the phenomena then you expect you can then find some kind of framework in which to talk to help to understand the phenomena.

RICHARD P. FEYNMAN

Closed Loop and Tree Diagrams

INTRODUCTION

It is the purpose of this paper to discuss some connections between the mathematical formulas for the amplitude of various processes in field theory when calculated in perturbation theory. The diagrams representing such amplitudes are generally of two forms; either the lines are interconnected so that a number of closed loops are formed, or there are no such loops. In the latter case we call the diagram a tree diagram. Examples of each type are given in Figure 1. For the moment we shall not be concerned with exactly what particles form the loops and the external lines; there is to be no implication that all parts of the loop represent propagation of the same particle.

In a tree diagram (Figure 1a), all the four-momenta of propagation of the internal lines are determined by the four-momenta of the external lines. Such a diagram can be directly evaluated (for incoming and outgoing states of given momenta) without any integration.

A diagram containing one closed loop (Figure 1b) represents a mathematical term involving an integration over a single four momentum which is not determined by the external momentum. It is, for example, the momentum of one of the propagators in the loop. If there are other loops, the term represented has further four momentum integrals; if the diagram has n independent loops, there is a $4n$-fold integral over the four components of each of n independent momentum variables.

We shall show that any diagram with closed loops can be expressed in terms of sums (actually integrals) of tree diagrams. In each of these tree diagrams there is, in addition to the external particles of the original closed loop diagram, certain particles in the initial and in the final state of the tree diagram. These

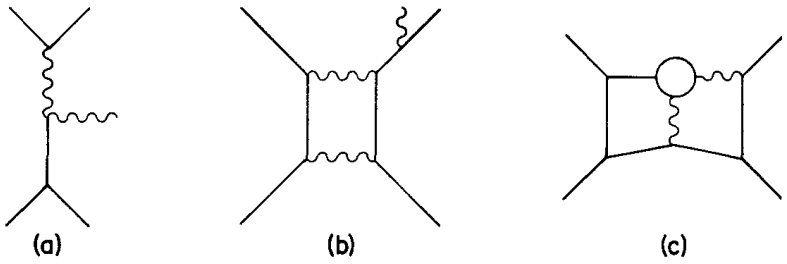


FIGURE 1.
(*a*) Tree diagram.
(*b*) Diagram involving a single closed loop.
(*c*) Diagram containing three independent closed loops.

additional particles are of the physical kind that appeared in the closed loop. For example, if an electron line appears, the tree diagrams into which the loop is converted may contain an extra external line entering and leaving, representing a positron of three-momentum $\mathbf{P}$, and of energy $\sqrt{m_e^2 + \mathbf{P}\cdot\mathbf{P}}$ on the electron mass shell m_e. The incoming and outgoing positron state is the same. That is, these diagrams are diagrams appearing in the forward scattering amplitude for the extra particles, scattering in the presence of the external particles. The sum on the tree diagrams is the sum for the two spin states of the extra positron and the integral over all momenta $\mathbf{P}$ of that positron.

If each individual closed loop diagram can be so expressed as tree diagrams, it is clear that the sum of all of the diagrams for a given process in any order can also be expressed in tree diagrams. But an interesting question arises as to whether the resulting multitude of tree diagrams can be represented as a sum of sets of tree diagrams, each set representing the complete set of tree diagrams expected for some given physical process. (In many cases it can be done and we will discuss these matters in the latter part of this paper.) In this way we obtain relations among the diagrams for various processes. As a special example of such a relation, we are all familiar with the unitarity relation which, for example, relates the imaginary part of a forward scattering to the integral over the square of certain particle production amplitudes. Our relations are more general.

These theorems are very close to those developed by Cutkosky. We are not sure whether they may be of any practical use in understanding field theory or strong interactions, but they have helped the author to resolve some ambiguities in the quantum theory of gravitation and of Yang-Mills vector mesons of zero mass. This application appears in the following paper.

REDUCTION OF A SINGLE LOOP TO TREES

A typical single closed loop diagram, like that of Figure 1b, will lead to an integral of the form

$$A = \int \frac{d^4k}{(2\pi)^4} N(k) \prod_i I_+^{(i)}(k) \tag{1}$$

where $N(k)$, the numerator, is some polynomial in the components of the four vector k_μ, and $I_+^{(i)}(k)$ is a factor for the i-th virtual line of the form

$$I_+^{(i)}(k) = \frac{i}{(k - p_i)^2 - m_i^2 + i\eta} \tag{2}$$

in the limit $\eta \to +0$, where m_i is the mass of the i-th particle and p_i is some momentum, being the sum of the momenta brought in by the external lines starting from some point on the loop and adding up to line i (that is, $p_{i+1} = p_i +$ the momentum brought into the loop by the external line at the junction of propagator i and $i + 1$). Any spin $\frac{1}{2}$ particles have had their usual propagator $(\not{p} - m)^{-1}$ rationalized to $(p + m)(p^2 - m^2)^{-1}$. The character of the poles of I_+ is defined in the usual way by the $i\eta$. It insures that positive frequencies are propagated forward in time, and negative frequencies backward. We shall suppose the integral converges; there are enough powers of k in the denominator D to offset those in $N(k)$ and d^4k. If they do not converge, the physical difficulties are offset by one or another convergence techniques (such as differentiating with respect to some of the m_i^2 and integrating back again) which relate the integral to others of the same type which do converge. These things can also be done with our theorems but we will not discuss them here.

If we write $k_\mu = (\omega, \mathbf{k})$, $p_i = (E_i, \mathbf{p}_i)$, our factor (2) is

$$\begin{aligned} I_+(k) &= \frac{i}{(\omega - E_i)^2 - ((\mathbf{k} - \mathbf{p}_i)^2 + m_i^2) + i\eta} \\ &= \frac{1}{2\epsilon_i}\left[\frac{1}{(\omega - E_i) - \epsilon_i + i\eta} - \frac{1}{(\omega - E_i) + \epsilon_i - i\eta}\right] \end{aligned} \tag{3}$$

where ϵ_i is written for $+\sqrt{(\mathbf{k} - \mathbf{p}_i)^2 + m_i^2}$ and we have written a new $i\eta$ to explain the position of the poles of ω, *viz.* at $E_i + \epsilon_i$ just below the real axis, and at $E_i - \epsilon_i$ just above. Thus, there are poles both above and below the real axis in ω, the integral (from $d^4k = d\omega\, d^3\mathbf{K}$) along the real axis. By changing the location of the pole in the last term, we can also write

$$I_+^{(i)}(k) = I_R^{(i)}(k) + \frac{\pi}{\epsilon_i}\delta(\omega - E_i + \epsilon_i) \tag{4}$$

where

$$I_R^{(i)}(k) = \frac{i}{2\epsilon_i}\left[\frac{1}{(\omega - E_i) - \epsilon_i + i\eta} - \frac{1}{(\omega - E_i) + \epsilon_i + i\eta}\right] \tag{5}$$

and we have used

$$\frac{i}{2\epsilon_i}\left[\frac{1}{(\omega - E_i) + \epsilon_i + i\eta} - \frac{1}{(\omega - E_i) + \epsilon_i - i\eta}\right] = \frac{\pi}{\epsilon_i}\delta(\omega - E_i + \epsilon_i). \tag{6}$$

If each I_i in the integral A of equation (1) is replaced by the sum (4), we shall have an integral A', like A, but all the I_+ replaced by I_R plus a residual in which one or more of the factors replacing $I_+^{(i)}$ has a δ function. This integral A' is

evidently zero, for all the poles are below the real axis of ω, so the contour can be closed in ω by a large semicircle above the real axis yielding zero for the integral (if the original integral on ω converges, the integral on the semi-circle approaches zero for large enough radius). Thus, the integral A is expressible by a sum of terms each of which contains at least one δ function. But such terms are related to tree diagrams.

If a term has just one δ function, as in (6), the integral over ω can of course be done immediately; contributions appear only for

$$\omega - E_i = -\sqrt{(\mathbf{K} - \mathbf{P}_i)^2 + m_i^2}, \tag{7}$$

that is, for such ω that particle i is on its mass shell (as an antiparticle). That is, this equation (7) can be used to substitute for ω. The integrand, if all the other factors are I_R's is just like the amplitude for a tree diagram for the same external lines but with one extra antiparticle of the type i of momentum $\mathbf{K}' = \mathbf{K} - \mathbf{P}_i$, energy $\sqrt{(\mathbf{K}')^2 + m_i^2}$ entering and leaving (for example, see Figure 2b). The remaining integral over $d^3\mathbf{K}$ (or equivalently over $d^3\mathbf{K}'$) is an integral over all momenta of the extra antiparticle. It is not precisely the amplitude for a tree diagram because the propagator in the remaining lines of the loop still present is I_R instead of I_+. But that we shall remedy in just a moment.

A term with δ functions arising from two of the factors, with the rest of the factors replaced by I_R, becomes a disjoint diagram (if there is just one closed loop) consisting of the product of two separate pieces (see Figure 2c). One kind of antiparticle comes in with momentum $\mathbf{K}'$, say, and another goes out with momentum $\mathbf{Q}$, which is entirely determined by $\mathbf{K}'$ and the external lines. In the other factor the antiparticle $\mathbf{Q}$ goes in and $\mathbf{K}'$ comes out. We sum all polarizations and integrate over all momenta $\mathbf{K}'$. Another way of expressing this, disregarding the fact that the diagram is disjoint, is to say we have a diagram with two antiparticles $\mathbf{K}'$, $\mathbf{Q}$ coming in, and two going out. All the couplings of the various particles must be exactly as prescribed by the connections in the original closed loop. More δ functions correspond to more extra antiparticles in and out.

The fact that the propagators in these tree diagrams is I_R instead of I_+ may be an annoyance, but it is easily remedied by substituting in these tree diagrams for I_R via equation (4) reversed:

$$I_R(k^2) = I_+(k^2) - \frac{\pi}{\sqrt{\mathbf{K}^2 + m^2}}\,\delta(\omega + \sqrt{\mathbf{K}^2 + m^2}) \tag{8}$$

where

$$I_+(k^2) = (k^2 - m^2 + i\eta)^{-1}.$$

The result is, algebraically, best seen by writing, first, in place of equation (1), the relation $A' = 0$, valid because all the poles are on one side of the axis:

$$A' = \int \frac{d^4k}{(2\pi)^4}\, N(k) \prod_i I_R^{(i)}(k) = 0 \tag{9}$$

and substituting (8) in this

$$\int \frac{d^4k}{(2\pi)^4} N(k) \prod_i \left[I_+^{(i)}(k) - \frac{\pi}{\sqrt{(\mathbf{K} - \mathbf{P}_i)^2 + m_i^2}} \times \delta(\omega - E_i + \sqrt{(\mathbf{K} - \mathbf{P}_i)^2 + m_i^2}) \right] = 0. \quad (10)$$

On expanding the product (10), the term with all I_+ is the closed loop (A, equation 1) desired, whereas all the other terms are the diagrams with various numbers of extra antiparticles scattering forward (each times -1 for each extra antiparticle), and this time all the remaining propagators are the standard I_+. Since the sum vanishes, we have expressed the closed loop in terms of tree diagrams.

THE PROCESS OF OPENING LOOPS

Let us call this process, using the principle of equation (10) to express the loop in terms of diagrams with extra particles (that is, integrals with at least one δ function), "opening the loop." Evidently we could have shifted the first pole in the expression (3) and obtained a different opening formula, in which particle scattering tree diagrams replace antiparticle scattering. It corresponds to considering the loop as going around in the opposite direction.

If there is more than one loop in the original diagram, the loops may be opened in succession. Choose any one loop; that is, integration over any one virtual momentum k, leaving the others to integrate later. Then this loop can be opened. What results is a diagram sum and integral over diagrams with extra particles, but which still has loops remaining in it. However, there is now one less loop, and in each remaining loop all the propagators are I_+ (if equation 10 is used). Therefore, a remaining loop may be treated in the same way, thus reducing the number of loops still further, until there are none left.

We can understand things very nicely if we look at what we have been doing in coordinate space. A single closed loop (without external propagators) can be written in the form (the symbol 1 stands for the space-time positions $x_{\mu 1}$ of the point 1, etc.):

$$L = \int\int\int \mathrm{Tr}\,[\chi_n(n) I_+(n, n-1) \cdots \chi_3(3) I_+(3, 2) \chi_2(2) I_+(2, 1) \chi_1(1) I_+(1, n)] \, d^4x_1 \, d^4x_2 \cdots d^4x_n \quad (11)$$

where $I_+(2, 1)$ as a function of the four vector $x_2 - x_1$ is the Fourier transform of the propagator $I_+(p)$ for the particle going from 1 to 2 (all these I's are not identical, because each should correspond to the mass of the appropriate particle, but we leave out the superscript i for the sake of simplicity). The various χ's are potentials or operators (derivatives, isopin operators, etc.); each is, however, local, depending on the one space-time point or a few of its derivatives.

Thus, $I_+(2, 1)$ is defined as the solution of

$$(\Box^2 - m^2)I_+(2, 1) = i\,\delta^4(2, 1) \tag{12}$$

where the d'Alembertian operates on variable 2 and $\delta^4(2, 1)$ is the four-dimensional δ function $\delta(t_2 - t_1)\,\delta(x_2 - x_1)\,\delta(y_2 - y_1)\,\delta(z_2 - z_1)$, which has only positive frequencies for $t_2 > t_1$ and negative frequencies for $t_2 < t_1$. That is,

$$\begin{aligned} I_+(2, 1) &= \int \frac{d^3P}{2E_P(2\pi)^3}\, e^{-iE_P(t_2 - t_1)} e^{iP\cdot(x_2 - x_1)} \qquad \text{for } t_2 > t_1 \\ &= \int \frac{d^3P}{2E_P(2\pi)^3}\, e^{-iE_P(t_1 - t_2)} e^{iP\cdot(x_2 - x_1)} \qquad \text{for } t_2 < t_1 \end{aligned} \tag{13}$$

where $E_P = +\sqrt{\mathbf{P}^2 + m^2}$.

We could also solve (12) with a different boundary condition, say that $I_R(2, 1)$ is the retarded solution, zero for $t_2 < t_1$, thus

$$(\Box^2 - m^2)I_R = i\,\delta^4(2, 1),$$

$$\begin{aligned} I_R &= \int \frac{d^3P}{2E_P(2\pi)^3}\,[e^{-iE_P(t_2 - t_1)} - e^{+iE_P(t_2 - t_1)}]e^{i\mathbf{P}\cdot(\mathbf{x}_2 - \mathbf{x}_1)} \qquad \text{for } t_2 > t_1 \\ &= 0 \qquad \text{for } t_2 < t_1. \end{aligned}$$

Then, since I_+ and I_R satisfy the same inhomogeneous linear equation the difference between them, $I_c(=I_+ - I_R)$, satisfies the homogeneous equation for a free particle

$$(\Box^2 - m^2)I_c(2, 1) = 0 \tag{14}$$

and

$$I_c = \int \frac{d^3P}{2E_P(2\pi)^3}\, e^{iE_P(t_2 - t_1)} e^{i\mathbf{P}\cdot(\mathbf{x}_2 - \mathbf{x}_1)}$$

for both $t_2 > t_1$ and $t_2 < t_1$.

Now it is evident that the expression in equation (11) vanishes if all the I_+ are replaced by I_R. That is because $I_R(2, 1)$ vanishes unless $t_2 > t_1$, so that if all the I_+ were I_R we would get zero unless $t_2 > t_1$, $t_3 > t_2$, $\cdots t_n > t_{n-1}$, $t_1 > t_n$, which is impossible. We cannot go around a closed ring with time always increasing and get back to our starting point.

If we substitute now $I_R = I_+ - I_c$ in this expression, equation (11) with all I_+ replaced by I_R, we get a relation between the term with all I_+ (which is the loop L we want) and a number of terms with I_c. Each term with I_c is evidently the scattering by an extra free antiparticle.

The sign is determined by what direction we choose to go around the loop in writing $I_R(2, 1)$. That is, draw an arrow around the loop in the direction of increasing time used in the original expressions with I_R. The labelling of a particle as electron or positron, say, is in accordance with this arrow and the proper quantum numbers in the diagram. Thus, for electron–electron bremsstrahlung scattering in Figure 2a, drawing the arrow in the direction 1, 2, 3, 4,

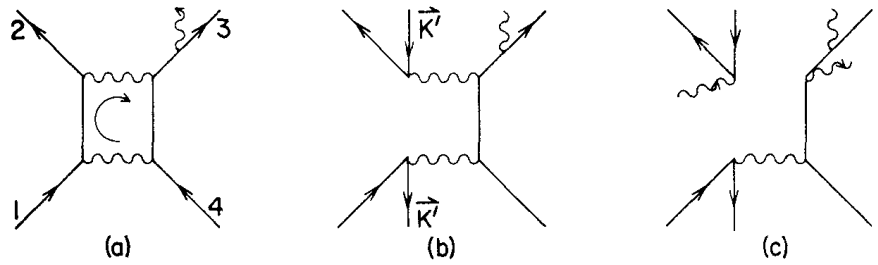


FIGURE 2.
A closed loop diagram (*a*) and some of the tree diagrams (*b*), (*c*) into which it can be broken up.

we see that the line 1, 2 corresponds to an electron; when opened, it means an extra positron coupled at 1, 2, as in Figure 2b. On the other hand, the arrow sees the particle from 3 to 4 as a positron and opening this section means an extra electron coupled there (Figure 3a). In the case of a particle identical to its anti-particle, we still have something to decide for the particular diagram and way of opening. In the case above, opening 2, 3 we have the tree 2c to add to 2b and 3a, but not the tree 3b (in this way of opening). That is, we want the photon coupled at 2 to bring in positive energy ω_Q, and at 3 to remove energy ω_Q, as in 2c and not the other way around 3b. This is for opening a particular diagram. What happens when we open all the diagrams for a given process we shall discuss later.

In order to put our equation in a little neater general mathematical form, consider first the case of a closed loop formed by a charged scalar meson mass m field ϕ coupled to a neutral meson field χ via a term $\int \phi^*(1)\chi(1)\phi(1)\, d\tau_1$ in the Lagrangian. Let it be a loop with n external lines corresponding to neutral mesons of momentum q_a, q_b, etc., and we wish the sum as all the lines change their order around the loop (see Figure 4). By choosing this simple example, we will clear our minds of complications.

From what we have said, the loop with I_R inside, which is zero, can be written as a sum of terms. The first term is the loop L. The next term with one I_c has one line or another open and a positive meson coming in. This is a scattering amplitude in the n potentials for a positive meson of momentum $\mathbf{P}$. Let us call $\langle \mathbf{P}'|T|\mathbf{P}\rangle$ the scattering amplitude of such a meson from $\mathbf{P}$ to $\mathbf{P}'$ in the potential χ (in this case, to n-th order); $\langle \mathbf{P}'|T|\mathbf{P}\rangle$ does *not* contain the $\delta_{\mathbf{PP}'}$ corresponding to the positive meson going right through without interaction. Then if L was the sum of the loops with the various χ_a, χ_i in any order, it is seen that the first term

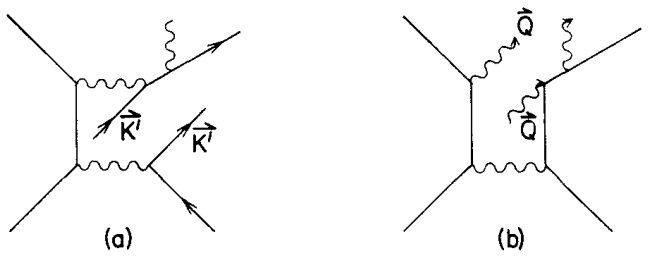


FIGURE 3.
See text.

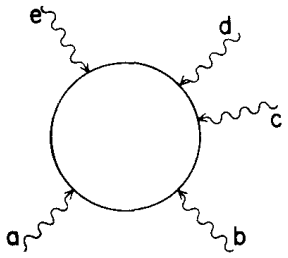


FIGURE 4.
See text.

has them in every order also. We are to integrate over all momenta $\mathbf{P}$, of the incident positive meson, symbolized by $\Sigma_{\mathbf{P}}$. We shall end up by proving that

$$0 = L - \Sigma_{\mathbf{P}}\langle\mathbf{P}|T|\mathbf{P}\rangle + \tfrac{1}{2}\Sigma_{\mathbf{P},\mathbf{Q}}\langle\mathbf{P}|T|\mathbf{Q}\rangle\langle\mathbf{Q}|T|\mathbf{P}\rangle - \tfrac{1}{3}\Sigma_{\mathbf{P},\mathbf{Q},\mathbf{R}}\langle\mathbf{P}|T|\mathbf{Q}\rangle\langle\mathbf{Q}|T|\mathbf{R}\rangle\langle\mathbf{R}|T|\mathbf{P}\rangle. \quad (15)$$

The first two terms are now clear. The third term is that coming from two I_c factors. Between these two factors, the various χ_a are distributed in various ways (for example, χ_a, χ_c, χ_d to the first, and χ_b, say, to the second). That is, in $\langle\mathbf{P}|T|\mathbf{Q}\rangle\langle\mathbf{Q}|T|\mathbf{P}\rangle$ we imagine the first scattering to be in a field such as χ_a, χ_c, χ_d, and the second in χ_b. That is, the product of two scatterings in a potential calculated so that the product is of the requisite order in the potentials. The factor $\frac{1}{2}$ is because putting, say χ_a, χ_c, χ_d in the first factor and χ_b in the second factor, is equivalent to putting χ_a, χ_c, χ_d in the second and χ_b in the first.

This relation (15) is the general relation for any loop. If there are quantum numbers like isotope spin on the virtual particle, of course these are to be summed over in the $\Sigma_{\mathbf{P}}$. If a loop consists of several different particles in different propagator sections, such as electron $1 \to 2$, photon $2 \to 3$, this is abstractly equivalent to some single particle which can change its type from electron to photon, etc., under the influence of the external lines. Therefore, if the $\Sigma_{\mathbf{P}}$ includes a sum also over the types of particle in the ring and the external potentials χ carry "quantum numbers," which determine how they change the virtual particles "quantum numbers," the formula can be used directly.

RULES FOR COMPLETE PROCESSES

We should like to develop some rules for complete processes. That is, we wish to add all diagrams for a given process and, if it contains closed loops, try to express it in terms of the sum of tree diagrams for other processes.

We start, for example, with our charged scalar mesons in an external scalar potential χ. Suppose now that χ is a fixed external potential. We know how to express the amplitude for scattering with external lines χ_a, χ_b, etc., in terms of the

scattering amplitude $S[\chi(1)]$ in an arbitrary potential by functional derivatives. *Viz.*, the amplitude with three external scalar mesons in and out, χ_a, χ_b, χ_c, is

$$\int\int\int \chi_a(1)\chi_b(2)\chi_c(3)\,\frac{\delta^3 S[\chi]}{\delta\chi(1)\,\delta\chi(2)\,\delta\chi(3)}\,d^4x_1\,d^4x_2\,d^4x_3. \tag{16}$$

Therefore, we begin by calculating in a general external potential.

Evidently (15) is then valid where L is the loop in the external potential, and all the amplitudes $\langle \mathbf{P}|T|\mathbf{Q}\rangle$ are scattering in this potential. But L is not the diagram for a complete process. On the other hand, e^L is. It is the amplitude that an initial vacuum (that is, no charged scalar meson) will still be a vacuum in the final state in spite of the action of the potentials (it is C_v in reference [1]). To study it, we take the exponential of both sides of (15) and expand it out to obtain:

$$\begin{aligned}
1 = {} & e^L \\
& - \Sigma_{\mathbf{P}}\langle \mathbf{P}|T|\mathbf{P}\rangle e^L \\
& + \tfrac{1}{2}\Sigma_{\mathbf{P},\mathbf{Q}}[\langle \mathbf{P}|T|\mathbf{P}\rangle\langle \mathbf{Q}|T|\mathbf{Q}\rangle + \langle \mathbf{Q}|T|\mathbf{P}\rangle\langle \mathbf{P}|T|\mathbf{Q}\rangle]e^L \\
& - \tfrac{1}{6}\Sigma_{\mathbf{P},\mathbf{Q},\mathbf{R}}[\langle \mathbf{P}|T|\mathbf{P}\rangle\langle \mathbf{Q}|T|\mathbf{Q}\rangle\langle \mathbf{R}|T|\mathbf{R}\rangle \\
& \qquad + 3\langle \mathbf{P}|T|\mathbf{Q}\rangle\langle \mathbf{Q}|T|\mathbf{P}\rangle\langle \mathbf{R}|T|\mathbf{R}\rangle \\
& \qquad + 2\langle \mathbf{P}|T|\mathbf{Q}\rangle\langle \mathbf{Q}|T|\mathbf{R}\rangle\langle \mathbf{R}|T|\mathbf{P}\rangle]e^L + , \text{ etc.}
\end{aligned} \tag{17}$$

Each term on the right side is the complete amplitude for some definite process

$$e^L = \text{amp vac to vac.}$$

$\langle \mathbf{P}|T|\mathbf{P}\rangle e^L$ = amplitude that an initial state with one positive meson ends in the same state (forward scattering of one meson in the potential), assuming it interacts. (Remember, $\langle \mathbf{P}'|T|\mathbf{P}\rangle$ does not contain the direct $\delta_{\mathbf{PP}'}$ term). We sum over all values of $\mathbf{P}$.

$[\langle \mathbf{P}|T|\mathbf{P}\rangle\langle \mathbf{Q}|T|\mathbf{Q}\rangle + \langle \mathbf{Q}|T|\mathbf{P}\rangle\langle \mathbf{P}|T|\mathbf{Q}\rangle]e^L$ is the amplitude that two positive mesons in the initial state, $\mathbf{P}$, $\mathbf{Q}$, are both scattered forward correctly, including exchange for the Bose particles (so the final state is $\langle \mathbf{P}|\langle \mathbf{Q}| + \langle \mathbf{Q}|\langle \mathbf{P}|$). The factor $\frac{1}{2}$ normalizes the states or, put another way, the $\frac{1}{2}$ connects the sum on both $\mathbf{P}$ and $\mathbf{Q}$ to the sums over each pair $\mathbf{P}$, $\mathbf{Q}$ counted once.

The next term is an analogous forward scattering from the state of three positive mesons with momenta $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ to the same state, again including exchange.

Thus we have for *charged bosons* in an external potential:

$$\begin{aligned}
1 = {} & \langle \text{no meson}|\text{no meson}\rangle - \Sigma\,\langle \text{forward scattering of one positive meson}\rangle \\
& + \Sigma\,\langle \text{forward simultaneous scattering two positive mesons}\rangle \\
& - \Sigma\,\langle \text{forward scattering three mesons}\rangle + , \text{ etc.}
\end{aligned} \tag{18}$$

This is an interesting relationship between amplitudes for various real physical processes. How general is it? Obviously it applies as well to scalar external

potentials, like χ, as to vector external potentials, as in the electrodynamics of charged scalar mesons. And the charged mesons could be other spin than zero also, if you like. It will *not* work directly if the mesons of the loop are neutral, a case to be discussed below. It will work if the mesons of the loop are not all the same kind and the propagators are successively of different masses and spins. But what is essential is that in every state the particle is distinct from the antiparticle—otherwise the equation is meaningless unless the phrase "positive extra meson scattering forward" has content. For if we have two such mesons in the third term, say, we do not wish them to be able to annihilate each other.

There is a corresponding theorem for Fermi particles. Consider electrons in an external potential (for example, electrodynamics). Equation (15), which opens a loop, is general and does not depend in any way on statistics. It works for spin $\frac{1}{2}$ as well as any other; it even works when parts of the loop are spin $\frac{1}{2}$ and other parts have other spin (as we have seen). It is therefore valid for electrons in an external potential.

The amplitude to go from a state with no electrons or positrons to the same state is, however, e^{-L} for Fermi statistics, rather than e^{+L} for Bose. Let us expand the exponential of the negative of each side of (15). We get

$$\begin{aligned} 1 = e^{-L} &+ \Sigma_{\mathbf{P}}\langle\mathbf{P}|T|\mathbf{P}\rangle e^{-L} + \tfrac{1}{2}\Sigma_{\mathbf{P},\mathbf{Q}}(\langle\mathbf{P}|T|\mathbf{P}\rangle\langle\mathbf{Q}|T|\mathbf{Q}\rangle \\ &- \langle\mathbf{P}|T|\mathbf{Q}\rangle\langle\mathbf{Q}|T|\mathbf{P}\rangle)e^{-L} + \tfrac{1}{6}\Sigma_{\mathbf{P},\mathbf{Q},\mathbf{R}}(\langle\mathbf{P}|T|\mathbf{P}\rangle\langle\mathbf{Q}|T|\mathbf{Q}\rangle\langle\mathbf{R}|T|\mathbf{R}\rangle \\ &- 3\langle\mathbf{P}|T|\mathbf{Q}\rangle\langle\mathbf{Q}|T|\mathbf{P}\rangle\langle\mathbf{R}|T|\mathbf{R}\rangle + 2\langle\mathbf{P}|T|\mathbf{Q}\rangle\langle\mathbf{Q}|T|\mathbf{R}\rangle\langle\mathbf{R}|T|\mathbf{P}\rangle)e^{-L}. \end{aligned} \tag{19}$$

It is evident that the terms represent scatterings, with exchange, into antisymmetric states, just as expected for the Fermi statistics. We have then, in the Fermi case,

$$\begin{aligned} 1 = &\langle\text{no positron}\,|\,\text{no positron}\rangle \\ &+ \Sigma_{\mathbf{P}}\,\langle\text{forward scattering of one positron}\rangle \\ &+ \Sigma\,\langle\text{forward scattering of two positrons}\rangle \\ &+ \Sigma\,\langle\text{forward scattering of three positrons}\rangle, \text{ etc.} \end{aligned} \tag{20}$$

Since with Fermi particles we can always distinguish particle and antiparticle, (19) is valid for any Fermi particles, charged or neutral, or having other indices such as the octet of nucleons (since the antiparticles, called positrons here, would be the distinct octet of antinucleons) in scalar or vector meson external fields, etc.

Equation (19) is easy to demonstrate directly. We take the case of electrons and positrons. Consider a state with no electrons in either the positron energy states, and in the old view of Dirac, *none* in the negative energy sea either. There being no electrons with which to interact, the external potentials would have no effect—the amplitude to remain in this same state is unity. Put in more modern language, consider an initial state which I will call "Dirac vacuum," which consists of no electrons present, but every possible positron state occupied. Then a little consideration will show that an external potential can do nothing. It

cannot scatter an electron—there are none. It cannot scatter a positron to another state—they are all full. It cannot annihilate a positron, for no electrons are present to annihilate—and it cannot create a pair, for all the positron states are already occupied. Thus, the Dirac vacuum remains undisturbed, the amplitude for it to do so being 1, the left side of equation (19).

But we can calculate this same amplitude another way, for it is well known that we can disregard the exclusion principle in intermediate states, as long as we are careful to make either the initial or final state antisymmetric.

Now, this amplitude, that the Dirac vacuum remains so in a potential, can be analyzed in the following way (following the conventions of reference [1]). There is some amplitude that none of the positrons interact with the potential and all go right through; this is just e^{-L}. Or one positron, say of momentum $\mathbf{P}$, may be scattered by the potential, the rest not. If the others are not, the final state of this positron must be the same in spin and momentum. The amplitude for this is $\langle \mathbf{P}|T|\mathbf{P}\rangle e^{-L}$, but we must sum over all the positrons which can do this (sum over two spins and all three-momentum). Again, perhaps only two of the positrons interact, the others not. If they are of momentum $\mathbf{P}$, $\mathbf{Q}$, they must either have these same momenta or else they interchange states. This leads to the third term, etc.

We could have started with this simple derivation of (20) and then taken logarithms to obtain (15), and finally argue that (15) is valid without respect to statistics in view of its nature involving only one loop. But this close connection no longer surprises us when we realize that the propagator appropriate to the old Dirac view of electrons in negative energy states is just I_R. That is, in each case we compare an amplitude calculated with I_R to one calculated with I_+.

I have not found a derivation for the Bose equation (18) as simple as the one just presented (using the Dirac vacuum) in the Fermi case.

It is evident, of course, that although we have used extra positrons in writing equation (20), we could just as well have used extra electrons and obtained another formula with electron scattering. It is not easy to find a simple expression in which the $\Sigma_{\mathbf{P}}$ can mean sum over electron and over positron states both. Just adding $\frac{1}{2}$ of equation (20) with positrons to $\frac{1}{2}$ of equation (20) with electrons will not do. The term in the scattering of two particles will not include the scattering of one electron and one positron.

Thus, I have not found a simple expression for processes in the case of neutral Bose particles (I mean a particle equivalent to its own antiparticle, like photons). It is to be remarked that the vacuum to vacuum amplitude is $e^{(1/2)L}$ because in the expression L in equation (15) I mean the loop calculated considering an order of potential interactions $a, b, c, \ldots, d$ to be distinct from the order $d, \ldots, c, b, a$ (as for charged bosons), while for neutral bosons there is no distinction of going around the loop one way or the other, and the contribution is $\frac{1}{2}$ of the expression used in (15). In any perturbation expansion it is possible, if artificial quantum numbers are put on the particles, to keep track. An expression for processes with neutral particles would be very useful. (We emphasize equation (15) is valid in any case—we are discussing the extensions to processes, analogous to equation (19) or (17).) Such an expression could (like (15)) be immediately generalized to any situation at all, with loops containing particles of different kinds.

How much more general can we make (19)? First of all, we can differentiate both sides with respect to the external field to any order to obtain relations valid in perturbation theory—but all such results are probably obtained more easily and in greater generality directly from (15).

But (18) and (20) apply also to interacting particles. For example, the amplitude $T_{e^2}[A]$ for a process with particles mutually interacting via electromagnetic fields can be obtained from the amplitude $T_0[A]$ for the same process in an external potential by suitably summing over various functions A. Explicitly,

$$T_{e^2}[A] = \int T_0[A + B]e^{iS(B)}\,\mathcal{D}B \tag{21}$$

where $S(B)$ is the action of the electromagnetic field [2]. Hence, performing this operation on both sides of equation (20), or (18), we find the relation valid for interacting particles. An expression such as (19) is no longer valid, because the two-particle scattering can no longer be written in the form of independent products $\langle \mathbf{P}|T|\mathbf{P}\rangle\langle \mathbf{Q}|T|\mathbf{Q}\rangle - \langle \mathbf{P}|T|\mathbf{Q}\rangle\langle \mathbf{Q}|T|\mathbf{P}\rangle$ when there is interaction.

In the expressions found in this way, closed loops of interacting fermions (or charged bosons) can be re-expressed in terms of scatterings without such loops (that is, the diagrams associated with charge renormalization, photon-photon scattering, etc.). However, all closed loops are not thereby eliminated, as diagrams such as Figure 2a, or self-energy diagrams, etc., are not opened by this process.

It is also clear from the argument for (20), starting from the Dirac vacuum, that the equation is equally valid if the particles mutually interact.

We might ask if we can generalize (17) by aiming not for the vac vac amplitude (e^L) but for a specific amplitude for antiparticle (or particle) scattering, say for a particle to go from state x to state y. The true amplitude for this is $\langle y|x\rangle e^L$. But if we calculate it by multiplying $\langle y|x\rangle$ by e^L from (17), we get

$$\langle y|x\rangle = \langle y|x\rangle e^L - \Sigma_{\mathbf{P}}\langle y|x\rangle\langle \mathbf{P}|\mathbf{P}\rangle e^L + \text{, etc.}$$

The left side is a simple tree diagram, but the right is not a sum of processes. The first term is satisfactory, being the amplitude for antiparticle scattering from x to y. But the second term is not the scattering of an antiparticle at x and at $\mathbf{P}$ to one at y and $\mathbf{P}$ for the exchange (or annihilation, if x is a particle) term is missing. The expression for this scattering should be

$$\Sigma_{\mathbf{P}}(\langle \mathbf{P}|\mathbf{P}\rangle\langle y|x\rangle + \langle y|\mathbf{P}\rangle\langle \mathbf{P}|x\rangle)e^L.$$

It is possible to sum the correct series of processes, but the sum is not $\langle y|x\rangle$—it is $\langle y|x\rangle_R$. We mean by $\langle y|x\rangle_R$ the scattering amplitude for antiparticles at x to go to y *calculated* using the retarded propagator I_R. It is therefore a tree diagram, for I_R permits no closed loops. That is,

$$\begin{aligned}\langle y|x\rangle_R = \langle y|x\rangle e^L &- \Sigma_{\mathbf{P}}[\langle \mathbf{P}|\mathbf{P}\rangle\langle y|x\rangle + \langle \mathbf{P}|x\rangle\langle y|\mathbf{P}\rangle]e^L \\ &+ \tfrac{1}{2}\Sigma_{\mathbf{P},\mathbf{Q}}[\langle \mathbf{P}|\mathbf{P}\rangle\langle \mathbf{Q}|\mathbf{Q}\rangle\langle y|x\rangle + \langle \mathbf{P}|\mathbf{Q}\rangle\langle \mathbf{Q}|\mathbf{P}\rangle\langle y|x\rangle \\ &+ 2\langle \mathbf{P}|\mathbf{P}\rangle\langle \mathbf{Q}|x\rangle\langle y|\mathbf{Q}\rangle + 2\langle \mathbf{P}|x\rangle\langle \mathbf{Q}|\mathbf{P}\rangle\langle y|\mathbf{Q}\rangle]e^L - \cdots\end{aligned}$$

(If x and y are particles, written $\bar{x}$, $\bar{y}$, the new exchange terms are appropriately annihilations, etc.) Evidently the terms on the left are scatterings of one antiparticle from x to y; the scattering of two antiparticles at x, $\mathbf{P}$ to y, $\mathbf{P}$ including exchange; the scattering of three antiparticles at x, $\mathbf{P}$, $\mathbf{Q}$ to y, $\mathbf{P}$, $\mathbf{Q}$ appropriately symmetrized, etc.

Evidently the general law for charged bosons making a transition from any initial state I, including any number of negative and positive particles, to a final state F under the influence of any potentials, or of mutual interaction of any kind, is

$$\{F|I\}_R = \{F|I\}_+ - \Sigma\{F, \mathbf{P}|I, \mathbf{P}\}_+ + \Sigma\{F, \mathbf{P}, \mathbf{Q}|I, \mathbf{P}, \mathbf{Q}\}_+ - \text{etc.} \quad (22)$$

where $\{F|I\}_R$ is the amplitude calculated using the retarded operator for any propagation of a particle or antiparticle (hence there are no particle loops in $\{F|I\}_R$); $\{F|I\}_+$ is the amplitude calculated in the usual way and therefore the one expected in the physical world (there are closed loops in $\{F|I\}_+$). The quantity $\{F, \mathbf{P}|I, \mathbf{P}\}_+$ is the usual physical amplitude for scattering from a state $|I, \mathbf{P}\}$ with an extra antiparticle in a state of momentum $\mathbf{P}$ and some spin, and others as in I, to a corresponding state $\{F, \mathbf{P}|$ with an antiparticle in the *same* spin and momentum state $\mathbf{P}$; but there are other particles in F, except that part of the amplitude in which the particle $\mathbf{P}$ does not interact with anything is to be omitted. The states I, $\mathbf{P}$ and F, $\mathbf{P}$ are to be suitably symmetrized as required by Bose statistics. The sum $\Sigma_{\mathbf{P}}$ is taken over all spin and momenta $\mathbf{P}$ of the extra antiparticle.

The state $|I, \mathbf{P}, \mathbf{Q}\}$ is the state I with two extra antiparticles and the sum here is on all the distinct possibilities—that is, $\frac{1}{2}$ the sum over all states $\mathbf{P}$ and all states $\mathbf{Q}$.

For particles obeying Fermi statistics the states are antisymmetrized, of course, and the minus signs on the right side of (22) should all be plus.

The scattering calculated via I_R involves only trees in the case of external potentials. The scattering of an electron, say in third order, appears as Figure 5a, and is nonvanishing only if $t_3 > t_2 > t_1$ (in fact, only if 2 is not outside the forward light cone of 1, 3 not outside that of 2, and therefore of 1 also, etc.). For positron scattering, the diagram appears as in Figure 5b; again, 1, 2, 3 are in time order, but the external positron connects at the time of the latest perturbation, the outgoing one at the time of the earliest perturbation.

Since all these terms represent forward scatterings, dispersion theory may be used in re-expressing them further. I do not know to what extent these formulas may be practical or useful in the study of strong interactions or in the simplification of calculations in electrodynamics.

If we desire not expressions for trees in terms of real processes, but real processes in terms of trees, we may turn these equations around and write

$$\{I|F\}_+ = \{I|F\}_R + \Sigma\{I, \mathbf{P}|F, \mathbf{P}\}_R + \Sigma\{I, \mathbf{P}, \mathbf{Q}|F, \mathbf{P}, \mathbf{Q}\}_R +, \text{etc.} \quad (22)'$$

for the Bose case, and alternating signs $[(-1)^n$ for n extra particles] for the Fermi case. I have found these useful in developing the quantum theory of gravitation (and the Yang-Mills theory of vector meson of zero mass). There I had con-

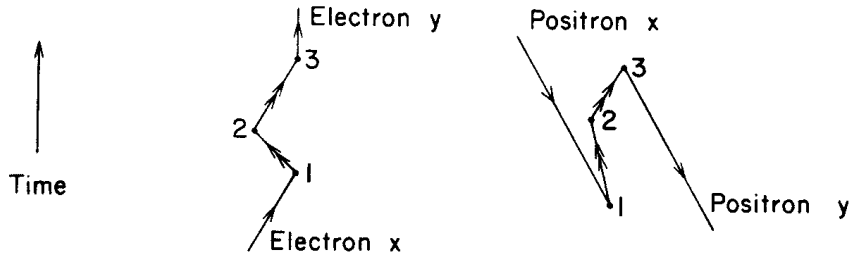


FIGURE 5.
Time relations of the retarded propagator in electron and in positron scattering.

siderable difficulty in finding a prescription for keeping closed loop diagrams gauge invariant, but no difficulty whatever with tree diagrams. The relations such as (15) or (22)′ helped me to resolve this difficulty, at least for the case of diagrams having one closed loop. The fact that I have not found the general expression analogous to (22)′ for neutral mesons (particle, antiparticle identity) has made this approach harder to handle in application to larger numbers of closed loops. This work will be published later.

MATHEMATICAL EXPRESSION FOR TREE DIAGRAMS

In this section we shall express in a more formal mathematical way just what the tree diagram for a given process is. We shall not be concerned with showing the relation of closed loop diagrams to a sum of tree diagrams. We shall limit ourselves to Bose fields for ease of representation; for tree diagrams this is no limitation.

Let there be a number of fields in interaction and call them collectively A, so A is a vector (function of position) having components for electrons, scalar mesons, photons, or what have you. The Lagrangian for the system is written $L(A)$, and is nonquadratic:

$$L(A) = L_2(A) - I(A) \tag{23}$$

where $L_2(A)$ is quadratic in A and describes all the propagators of the fields, and $I(A)$ is of higher than second order in the fields A and represents all the interactions.

The full quantum field amplitude for a given process including trees and closed loops can be written, as usual, as an integral over all fields B:

$$l = \text{amp with loops} = \int_{B_{\text{asym}} = A_0} \left(\exp i \int L(B)\, d\tau\right) \mathcal{D}B. \tag{24}$$

How in (24) do we represent the external particles? They are represented by a number of free particle waves $A_1, A_2, \ldots$ coming in and going out. We want each of them to act once. Thus, we form, with arbitrary coefficients $\alpha_1, \alpha_2, \ldots$, the expression

$$A_0 = \sum \alpha_i A_i \tag{25}$$

and perform the integral in (24) subject to the condition that asymptotically B (at $+\infty$ and $-\infty$) equals A_o (representing asymptotically waves coming in for $t \to -\infty$ and going out at $t = +\infty$). Then the amp in (24) depends on $\alpha_1, \alpha_2, \ldots$ The amplitude for the process with one particle coming in with wave A_1, one with wave A_2, etc., is just the first order action of each potential and is the coefficient of $\alpha_1\alpha_2\alpha_3 \ldots$ in this expression (24). If external potentials are acting, they are put into A_o and not differentiated (p. 337).

To be a little more explicit, let us define $R(A)$ as the first functional derivative of $L(A)$, and $R_2(A)$ likewise of $L_2(A)$,

$$R(A) = \delta\left(\int L(A)\, d\tau\right)\Big/\delta A = R_2(A) - K(A), \tag{26}$$

$$R_2(A) = \delta\left(\int L_2(A)\, d\tau\right)\Big/\delta A, \tag{27}$$

$$K(A) = \delta\left(\int I(A)\, d\tau\right)\Big/\delta A. \tag{28}$$

R_2 is then simply a linear differential operator, the "free particle propagator" whose reciprocal we shall write as R_2^{-1} and define precisely so that at $t = +\infty$ only positive frequencies are propagated, and at $t = -\infty$ only negative frequencies. Thus, if $R_2 = \Box^2 - m^2$, or $q^2 - m^2$, in momentum space, $R_2^{-1} = (q^2 - m^2 + i\eta)^{-1}$ in the limit as $\eta \to +0$. What happens when R_2 has no inverse due to a gauge group, such as in gravitation, will not concern us here. It presents no serious problem for trees, but it does in (24). For details in this case, see my other contribution to this volume (p. 377).

The asymptotic free particles each satisfy $R_2(A_i) = 0$ or in total $R_2(A_o) = 0$.

Another useful mode of expression which resolves certain ambiguities of surface integrals at infinity is to suppose that A_1, etc., were not strictly free but came from certain classical sources s_1, etc., at infinity, or else to suppose A_1 to be arbitrarily cut off near infinity and define $R_2(A_i) = s_i$. Thus, we write

$$A_o = R_2^{-1}s \tag{29}$$

where $s = \sum \alpha_i s_i$ is the effective source at infinity. The amplitude with all loops can also be stated as

$$l = \int_{B_{\text{asym}}=0} \exp i \int (L(B) - sB)\, d\tau\, \mathcal{D}B. \tag{30}$$

Now, s could be considered finite if there were external classical potentials acting, but if we have only quanta coming in and out we shall have to take only first order in all the α_i in s, thus s is infinitesimal.

All this is standard field theory.

We desire now to express, for the same problem, defined by the same A_o, equation (25), or better yet, by the same sources s, the amplitude for trees (no loops). It is

$$T = \text{amp with no loops} = \exp i \int (L(A) - As)\, d\tau \tag{31}$$

where A satisfies the equation

$$R(A) = s. \tag{32}$$

This equation (32) has the explicit solution

$$A = R_2^{-1}(K(A)) + A_o. \tag{33}$$

If we wish to avoid speaking of sources s, we should say the amplitude with no loops is $\exp i\int (L(A) - AR(A))\, d\tau$ where A makes $\int L(A)\, d\tau$ extremum subject to $A_{\text{asym}} = A_o$, and therefore given by (33). We should have to be very careful with ambiguous integrals like $\int AR_2(A)\, d\tau$, which must be interpreted as $\int (AK(A) + A_oR_2(A))\, d\tau$.

The result (31) may not be entirely obvious but we outline how it may be derived. If one considers how a connected tree diagram is formed with a set of lines $A_1, A_2, \ldots$ going in, and one extra F, coming out, it is clear that we are constructing an A'' being the first order in $\alpha_1, \alpha_2, \ldots$ of A' where A' is a solution of

$$A' = R_2^{-1}(K(A')) + \sum \alpha_i A_i \tag{34}$$

as can be seen by solving (34) by iteration and noticing term by term how the correct diagrams are formed by the equation, the final amplitude being

$$t = \int FK(A'')\, d\tau.$$

If we imply that the correct order (first in all α_i) is to be taken, we may write this

$$t = \int FR_2(A')\, d\tau.$$

Using the method of sources, so we can integrate by parts, this is

$$t = \int s_F A'\, d\tau$$

where s_F is the source of F or $s_F = R_2(F)$.

To get a more convenient and symmetrical form, we would like an expression not in terms of A' but in terms of A, which comes from (34) with the extra term ϕF added to the $\sum \alpha_i A_i$, and later take the derivative with respect to ϕ at $\phi = 0$. Since we now go to first order in ϕ, we can write

$$t = \phi \int s_F A\, d\tau \tag{35}$$

(the change from A' to A brings only in a ϕ^2 term). That is, upon varying a source by a perturbation δs, the change in t must be

$$\delta t = \int (\delta s) A\, d\tau. \tag{36}$$

This is solved[1] by

$$t = -\int (L(A) - sA)\, d\tau \tag{37}$$

because its first variation is

$$\delta t = -\int (\delta L/\delta A) - s)\, \delta A\, d\tau + \int \delta s A\, d\tau$$

where δA is the change in A induced via equation (32) by the change δs in s, and the first term vanishes by the equation (32) determining A. When it is considered that for the S matrix we want disjoint as well as connected diagrams the complete sum becomes the exponential (31).

The analysis of tree diagrams is very close to classical field theory, for it involves only the solution of the classical equations of motion (32). This is done, however, via (33), with complex boundary conditions. As closed loops are expressible in terms of trees, this suggests another way to get from classical to quantum field theory without ambiguity.

CONTRIBUTIONS FROM SINGLE LOOPS

For formal analysis, the trees may now be separated from the diagrams with loops by substituting $B = A + D$ in (30) to get (note, in the system with sources, $A_{\text{asym}} = 0$)

$$l = \int_{D_{\text{asym}}=0} \left(\exp i \int (L(A + D) - s(A + D))\, d\tau\right) \mathscr{D}D \tag{38}$$

factoring out T, the amplitude for trees from equation (31) we find

$$l = T \int \left(\exp i \int (L(A + D) - L(A) - R(A)D)\, d\tau\right) \mathscr{D}D$$

where the extra factor f represents the contributions from loop diagrams only is

$$f = \int \left(\exp i \int L_A(D)\, d\tau\right) \mathscr{D}D \tag{39}$$

where the field D goes vacuum to vacuum, all reference to sources or asymptotic fields is contained in the effective $L_A(D)$ Lagrangian for motion of D in an external field A,

$$L_A(D) = L(A + D) - D(\delta L(A)/\delta A) - L(A). \tag{40}$$

The Lagrangian $L_A(D)$ is purely of second order and higher in D. We define

$$R_A(D) = \delta L_A(D)/\delta D = R(A + D) - R(A),$$

[1] The connected diagram with a total of n lines in and out can also be written as: the first order in α_1 to α_n of $(n - 1)^{-1} \int L(A)\, d\tau$. For if, in (35), we consider F as the line A_1, we get $t = \alpha_1 \int s_1 A\, d\tau$, likewise for A_2, etc., so summing, $nt = \int sA\, d\tau$.

an operator which has a linear part, coming from the quadratic part $L_{A2}(D)$ of $L_A(D)$, which we call $R_{A2}(D)$. Thus, R_{A2} is a linear operator (depending upon the form of A) having an inverse R_{A2}^{-1} defined with the usual quantum boundary conditions. That is, we can formally study all problems of opening loops with external lines, by considering what happens when opening a loop without external lines, but with the mathematically more complex propagator R_A^{-1}.

Expanded only to second order, we get

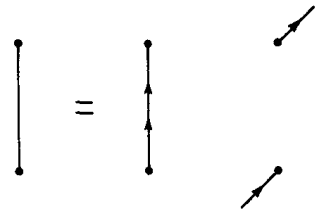
$$f_2 = \int \left(\exp i \int L_{A2}(D)\, d\tau\right) \mathscr{D}D,$$

which gives the usual closed loop expression, simply single loops with no external lines, propagator R_{A2}^{-1}; that is, $f_2 = e^L$, $L = \text{trace } (\ln R_{A2})$.

This propagator, say, $(R_{A2}^{-1})_+$, to be more explicit, can, by redefining its poles, be represented as in equation (4) as a retarded propagator $(R_{A2}^{-1})_r$ plus a free particle part. That is, $x = (R_{A2}^{-1})_r s$ represents a solution of $R_{A2}x = s$ with no field asymptotically in the past, while $(R_{A2}^{-1})_+ s = y$ represents a solution of $R_{A2}y = s$ with quantum boundary conditions. Their difference $x - y$, satisfying $R_{A2}(x - y) = 0$, is a solution of the homogeneous equation for antiparticle scattering (in the external potential, of course). We thus symbolize equation

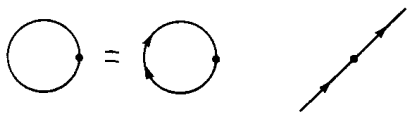
$$(R_{A2}^{-1})_+ = (R_{A2}^{-1})_r + \text{free scatt.} \tag{41}$$

by the diagrammatic symbolism



(42)

The retarded propagator is represented by a line with a double arrow in the direction of the latter terminal. The single arrows on the external lines indicate that they carry positive frequency in (at the lower terminal) or out (at the upper terminal). Therefore, for the first order perturbation of one closed loop, we have (the dot represents the perturbation)



(43)

The loop with the arrow representing the retarded propagator vanishes, for final termination time must exceed the initial time, and cannot be equal as the loop requires. Thus, one loop is equivalent to one extra particle scattering.

Opening such a loop gives simply a straight line for a simple particle propagated from past to future under the influence of the potential A. That is, the extra particle is in interaction with all the physically real external particles. In this case, the diagram when one loop is opened corresponds to a definite order of a complete physical problem, namely the forward scattering of one extra particle under the influence of all the others. When two connected loops are opened, the result is not always exactly the same as a definite physical problem as we shall see in the next section.

OPENING TWO CONNECTED LOOPS

As we have seen, we now need only to study loops with no external lines, but with propagator R_{A2}^{-1}. If we expand the operator $L_A(D)$ in (39) to still higher order,

$$L_A(D) = L_{A2}(D) + L_{A3}(D) + L_{A4}(D) + \cdots \tag{44}$$

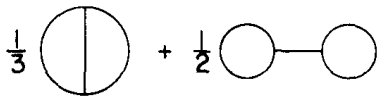


FIGURE 6.
The contributions of second order in $L_{A3}(D)$.

we shall obtain in the next order the diagrams of Figure 6. We also obtain from $L_{A4}(D)$ a diagram in the form of a figure eight but when we deal with those in Figure 6 it is immediately obvious how to include the figure eight. Each junction in Figure 6 represents a coupling of three lines via $L_{A3}(D)$ and each line represents an action of a propagator R_{A2}^{-1}.

Using (42) we can transform the first diagram of Figure 6 as indicated in Figure 7. The closed ring of propagators in the second line of Figure 7 vanishes for it contains a closed ring of retarded commutators. In Figure 7 xtra particles of different asymptotic momenta are indicated by different kinds of lines: dashed

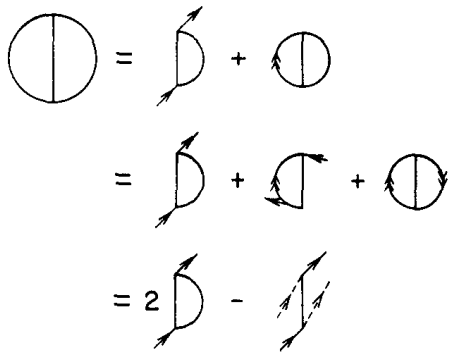


FIGURE 7.
Opening a double ring.

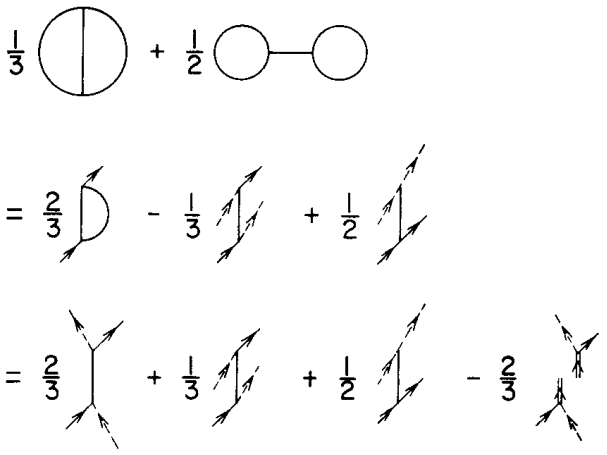
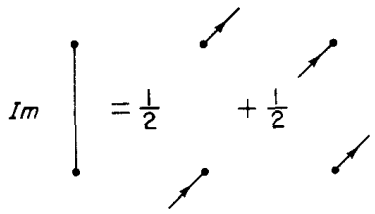


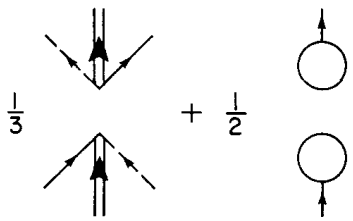
FIGURE 8.
Reduction of the Figure 6 diagrams to trees.

or straight. Opening the second diagram of Figure 6 is direct via (43), so we have Figure 8. The last diagram in Figure 8 is an entirely disjoint diagram, consisting of two factors, one being two extra particles scattering into a third, and the other its inverse.

Incidentally, we can check unitarity by noting that the imaginary part of the propagator $(R^{-1})_+$ is $\frac{1}{2}$ with retarded poles, $\frac{1}{2}$ advanced, or



We can then easily verify from the last line of Figure 8 that the imaginary part of the diagrams of Figure 8 is just



as required by unitarity.

The combination of the diagrams A, B, C, in Figure 8, namely $\frac{2}{3}A + \frac{1}{3}B + \frac{1}{2}C$ is *not* what an ordinary process would give. If we simply draw all the diagrams to represent the scattering of two particles on each other in this order, we obtain diagrams A, B, C with equal weight, $A + B + C$. It is true that if we imagined

the extra particles to carry certain quantum numbers or charges, then by making rules about what quantum numbers the intermediate propagator can carry the diagrams can be selected (for example, if the extra particles each carried a different charge and the intermediate must be neutral, only C survives). It is also true that in many cases such rules exist which lets the combination in Figure 8 remain as a true process. However, for completely neutral objects like gravitons the combination of Figure 8, found on opening the diagrams of Figure 6, is not a true process. Therefore, an attempt to formulate quantum gravitational rules by first saying how trees work and then using them to calculate closed loops is not completely straightforward. That is because a property, like gauge invariance, which is not valid for each diagram alone, but is valid for each set of diagrams representing a complete process, may not be valid if the tree diagrams are assembled into closed loops carelessly. The application of the methods in this paper to the problem of the quantum theory of gravitation appears in the following paper.

REFERENCES

1. R. P. Feynman, *Phys. Rev.*, **76**, 749 (1949).
2. R. P. Feynman, *Phys. Rev.*, **80**, 440 (1950), Eq. (46).

RICHARD P. FEYNMAN

Problems in Quantizing the Gravitational Field, and the Massless Yang-Mills Field

INTRODUCTION

Some years ago I started extensive work on the quantum theory of gravitation, studying difficulties in formulating it, as well as detailed analysis of its divergence and renormalization characteristics. However, I became involved in other interests and the work was never completed. Many people, including Professor Wheeler, have asked me what I found. I shall not discuss the renormalization work here, but I will take this opportunity to publish the material I have, incomplete though it may be. Therefore, in this paper we shall discuss some problems which arise in attempting to formulate a quantum theory of gravitation. I have heard that these problems have been solved by others,[1] but I will report how the problem looked to me some years ago, correcting, however, some errors in the equations for the gravitational case which I discovered while preparing this manuscript.

The questions about making a "quantum theory of geometry" or other conceptual questions are all evaded by considering the gravitational field as just a spin-2 field nonlinearly coupled to matter and itself (one way, for example, is by expanding $g_{\mu\nu} = \delta_{\mu\nu} + h_{\mu\nu}$ and considering $h_{\mu\nu}$ as the field variable) and attempting to quantize this by following the prescriptions of quantum field theory, as one expects to do with any other field. The central difficulty springs from the fact that the Lagrangian is invariant under a gauge group, and therefore the propagator is singular and generally undefined. In the classical theory and in quantum electrodynamics this problem may be evaded by starting with a modified Lagrangian which is no longer completely gauge invariant (for example,

[1] For related references, see *note added in proof* at the end of the article.

in electrodynamics add a term like $(\nabla_\mu A_\mu)^2$ to the Lagrangian). This leads to a nonsingular propagator (for example, to $\delta_{\mu\nu}q^2$ in electrodynamics) but because of conservation laws that the sources satisfy (a consequence of the original gauge group invariance) it can be shown that the extra term added to the Lagrangian has no final physical effect. In the quantum gravitational theory, adding to the Lagrangian a term of this kind permits a definite prescription for calculating all diagrams. It is found that all tree diagrams (diagrams which involve no closed loops) are, like the classical theory, completely satisfactory. However, diagrams with closed loops are affected. In particular, if a diagram with one closed loop is compared to its related tree diagrams by unitarity (or by the loop opening processes of reference I[2]) a discrepancy is found. This discrepancy can easily be removed by a further modifying rule—that for each closed ring of gravitons (calculated with the modified nonsingular Lagrangian) a special term must be subtracted. This special term is describable as the contribution of a corresponding closed loop around which, in place of the graviton, goes a special auxiliary vector particle (propagated in a definite manner and coupled to gravitation but not directly to matter).

When a diagram contains two contiguous rings (like the letter θ) I have not been able to find the corresponding rule. Part of the difficulty here is to find an unambiguous definition of what the double ring should equal in a correct theory. The main purpose of this paper is to discuss various attempts to resolve these difficulties and to describe some incidental relations that have been found in attempting to do so. It is not even clear that the problems are serious, or are not easily solved, as I have fallen into calculational confusion.

The algebraic complexity of the gravitational field equations is so great that it is not easy to do exploratory mathematical investigations and checks. Gell-Mann suggested to me that the Yang-Mills theory of vector particles with zero mass also is a nonlinear theory with a gauge group and might show the same difficulties, and yet be easier to handle algebraically. This proved to be the case, and thereafter, all the work was done first with the Yang-Mills theory and then the corresponding expressions for gravitation were worked out. The correspondence is exceedingly close. Each difficulty and its resolution in one theory has its corresponding difficulty and resolution in the other. It becomes obvious that to find a completely satisfactory quantization of the zero mass Yang-Mills field, is to find a completely satisfactory quantization of the general theory of relativity.

For this reason, all the work in this paper will deal primarily with the Yang-Mills theory. The correspondence of the equations in the gravitation theory will be developed in an appendix.

YANG-MILLS FIELD THEORY

The source of gravitation is energy and momentum, quantities which are locally conserved. The gravitational field carries energy and momentum and hence must

[2] See the previous paper "Closed Loops and Tree Diagrams" by R. P. Feynman which hereafter we call I.

be coupled to itself. The source of the Yang-Mills field is isotopic spin current, a quantity locally conserved. The field carries isotopic spin and hence must be coupled to itself, to make a nonlinear field theory.

We shall formulate the Yang-Mills theory in the special case that it is interacting with Dirac matter of spin $\frac{1}{2}$ which carries half integral isospin (that is, is a representation of SU_2). It is readily generalizable to matter of any spin, or of several kinds—but the problems we wish to deal with are not at all affected by that. In fact, we find that our problems arise entirely with the field acting upon itself alone, the addition of other matter as sources only making a small additional algebraic complication, and no new physical problems. Again, extension into SU_n can be made immediately, merely by a reinterpretation of symbols, so to get to our real problems as quickly as possible we suppose matter represented by an SU_2 spinor, Dirac spinor ψ, in interaction with an isospin vector and four-space vector field $\mathbf{A}_\mu$ with Lagrangian

$$L_{\text{matter}} = \bar{\psi}\gamma_\mu(i\nabla_\mu - \boldsymbol{\tau}\cdot\mathbf{A}_\mu)\psi + m\bar{\psi}\psi \tag{1}$$

where $\boldsymbol{\tau}$ are the Pauli matrices for isospin. A rotation in the SU_2 space induces on ψ the transformation $\psi \to e^{i\boldsymbol{\tau}\cdot\boldsymbol{\alpha}}\psi = (1 + i\boldsymbol{\tau}\cdot\boldsymbol{\alpha})\psi$ for infinitesimal rotation and, on $\boldsymbol{\tau}\cdot\mathbf{A}_\mu$ the transformation $e^{i\boldsymbol{\tau}\cdot\boldsymbol{\alpha}}\boldsymbol{\tau}\cdot\mathbf{A}_\mu e^{-i\boldsymbol{\tau}\cdot\boldsymbol{\alpha}} = \boldsymbol{\tau}\cdot(\mathbf{A}_\mu - \boldsymbol{\alpha} \times \mathbf{A}_\mu)$ for infinitesimal $\boldsymbol{\alpha}$, where we define the cross product $\times$ between two isovectors $\mathbf{A}$, $\mathbf{B}$, with components $\mathbf{A}_i$, $\mathbf{B}_i$,

$$(\mathbf{A} \times \mathbf{B})_k = \lambda_{kij}A_iB_j,$$

where the coefficients of the Lie group λ_{kij} are defined by

$$\tau_i\tau_j - \tau_j\tau_i = i\lambda_{ijk}\tau_k.$$

(For SU_2 in the three-dimensional isovector space, this is just the ordinary cross product of vector analysis in three dimensions.) Then $[\boldsymbol{\tau}\cdot\mathbf{A}, \boldsymbol{\tau}\cdot\mathbf{B}] = i\boldsymbol{\tau}\cdot[\mathbf{A} \times \mathbf{B}]$.

If we now ask that our expressions be invariant for local SU_2 transformations, the quantity $\boldsymbol{\alpha}$ becomes a function of four-space position, and does not commute with $i\nabla_\mu$, but this changes to $i\nabla_\mu - \boldsymbol{\tau}\cdot(\nabla_\mu\boldsymbol{\alpha})$. The expression (1) will be invariant if $\mathbf{A}_\mu$ changes by

$$\mathbf{A}_\mu \to \mathbf{A}_\mu - \nabla_\mu\boldsymbol{\alpha} - \boldsymbol{\alpha} \times \mathbf{A}_\mu. \tag{2}$$

Henceforth, our notation will be $\mathbf{C}_{,\mu}$ for $\nabla_\mu\mathbf{C}$ and

$$\mathbf{C}_{;\mu} = \nabla_\mu\mathbf{C} - \mathbf{A}_\mu \times \mathbf{C} \tag{3}$$

is a "covariant differentiation" of any isovector $\mathbf{C}$ (which may carry other space indices, of course). This has the property that $\mathbf{C}_{;\mu}$ transforms as an isovector if $\mathbf{C}$ does. We call (2) a gauge transformation.

Yang and Mills found an expression for a Lagrangian for the field $\mathbf{A}_\mu$ that would be invariant under (2). The part that contains $\mathbf{A}_\mu$ only is

$$L_{YM}(\mathbf{A}) = \tfrac{1}{4}\mathbf{E}_{\mu\nu}\cdot\mathbf{E}_{\mu\nu} \tag{4}$$

where

$$\mathbf{E}_{\mu\nu} = \mathbf{A}_{\mu,\nu} - \mathbf{A}_{\nu,\mu} + \mathbf{A}_\mu \times \mathbf{A}_\nu. \tag{5}$$

From (3) we find this antisymmetric $\mathbf{E}_{\mu\nu}$ might also be defined in terms of covariant differentiation via

$$\mathbf{C}_{;\mu;\nu} - \mathbf{C}_{;\nu;\mu} = -\mathbf{E}_{\mu\nu} \times \mathbf{C}. \tag{6}$$

The variation of $\int L_{YM}(A)\, d\tau$ with respect to $\mathbf{A}_\mu$ we call $R_\mu(\mathbf{A})$. It is

$$R_\mu(\mathbf{A}) = -\mathbf{E}_{\mu\nu;\nu} \tag{7}$$

so that from (6) we can deduce directly that

$$R_{\mu;\mu}(\mathbf{A}) = 0. \tag{8}$$

This R_μ is a nonlinear operator, its linear part $R_{2\mu}(\mathbf{A})$ coming from the quadratic part of the YM Lagrangian

$$L_{2YM}(\mathbf{A}) = \tfrac{1}{2}(\mathbf{A}_{\mu,\nu}\cdot\mathbf{A}_{\mu,\nu}) - \tfrac{1}{2}(\mathbf{A}_{\mu,\mu})^2 \tag{9}$$

is

$$R_{2\mu}(\mathbf{A}) = \mathbf{A}_{\mu,\nu\nu} - (\mathbf{A}_{\nu,\nu})_\mu \tag{10}$$

or in momentum space

$$R_{2\mu}(\mathbf{A}) = (q^2\,\delta_{\mu\nu} - q_\mu q_\nu)\mathbf{A}_\nu. \tag{11}$$

The operator $q^2\,\delta_{\mu\nu} - q_\mu q_\nu$ is singular and does not have an inverse to serve as a natural propagator.

We can write the full Lagrangian in interaction as a piece representing free matter $\bar\psi i\gamma_\mu\nabla_\mu\psi + m^2\bar\psi\psi$ and a part involving the fields

$$L(\mathbf{A}) = L_{YM}(\mathbf{A}) - \mathbf{J}_\mu\cdot\mathbf{A}_\mu \tag{12}$$

where J_μ, the current or sources for the field, comes from matter and is $\bar\psi\boldsymbol{\tau}\gamma_\mu\psi$. It is then a consequence of the equations of motion of the matter (from varying $\bar\psi$ in [1])

$$\gamma_\mu(i\nabla_\mu - \boldsymbol{\tau}\cdot\mathbf{A})\psi = m^2\psi$$

(since they are SU_2 invariant) that this current is conserved:

$$\mathbf{J}_{\mu;\mu} = 0. \tag{13}$$

The field equations of motion upon varying A in (12) are

$$R_\mu(\mathbf{A}) = \mathbf{J}_\mu. \tag{14}$$

The operator R_μ is singular and the equation cannot hold for a general $\mathbf{J}_\mu$. In view of the identity (8), this equation (14) can only hold if J_μ is conserved (13).

To study finite transformations, write $\sigma = e^{i\boldsymbol{\tau}\cdot\boldsymbol{\alpha}}$ and $a_\mu = \boldsymbol{\tau}\cdot\mathbf{A}_\mu$. Then the transformation of ψ is to $\psi' = \sigma\psi$ and of an ordinary isovector $C(=\boldsymbol{\tau}\cdot\mathbf{C})$ to C' is

$$C' = \sigma C\sigma^{-1} \tag{15}$$

but the field a_μ transforms in a special way, with a gradient term, $SU(2)$

$$a_\mu' = \sigma(a_\mu + t_\mu)\sigma^{-1} \tag{16}$$

where

$$t_\mu = i\sigma^{-1}(\nabla_\mu\sigma) \tag{17}$$

with

$$\epsilon_{\mu\nu} = a_{\mu,\nu}a_{\nu,\mu} - a_\mu a_\nu + a_\nu a_\mu, \tag{18}$$

$$L = \tfrac{1}{4}\,\mathrm{Tr}\,(\epsilon_{\mu\nu}\epsilon_{\mu\nu})$$

is an invariant.

These expressions can be transformed to ordinary notation via the equation for commutators of τ. For example, equation (15) would become

$$\mathbf{C}' = \mathbf{C} - \boldsymbol{\alpha} \times \mathbf{C} + \tfrac{1}{2}\boldsymbol{\alpha} \times (\boldsymbol{\alpha} \times \mathbf{C}) - \frac{1}{3!}\,\boldsymbol{\alpha} \times (\boldsymbol{\alpha} \times (\boldsymbol{\alpha} \times \mathbf{C})), \tag{15'}$$

or, in an obvious notation, $\mathbf{C}' = e^{-(\alpha\times)}\mathbf{C}$. The quantity t_μ is $\boldsymbol{\tau}\cdot\mathbf{T}_\mu$, where

$$\mathbf{T}_\mu = -\boldsymbol{\alpha}_{,\mu} - \tfrac{1}{2}\boldsymbol{\alpha} \times \boldsymbol{\alpha}_{,\mu} - \tfrac{1}{6}\boldsymbol{\alpha} \times (\boldsymbol{\alpha} \times \boldsymbol{\alpha}_{,\mu}) + \cdots \tag{17'}$$

or

$$T_\mu = (\boldsymbol{\alpha} \times)^{-1}(1 - e^{(\alpha\times)})\boldsymbol{\alpha}_{,\mu}.$$

We shall have to consider two other related theories with slightly modified Lagrangians which avoid the singularities of the propagator coming from (4), (9). The first we shall call *YMM*, Yang-Mills with mass, and adds a term $m^2\mathbf{A}_\mu\cdot\mathbf{A}_\mu$ to the Lagrangian:

$$L_{YMM}(\mathbf{A}) = L_{YM}(\mathbf{A}) - \tfrac{1}{2}m^2\mathbf{A}_\mu\cdot\mathbf{A}_\mu. \tag{19}$$

The equations (14) now become

$$R_\mu(\mathbf{A}) - m^2\mathbf{A}_\mu = \mathbf{J}_\mu \tag{20}$$

which has solutions for *any* J_μ, for

$$-m^2\mathbf{A}_{\mu;\mu} = \mathbf{J}_{\mu;\mu}, \tag{21}$$

but if the sources are conserved we have (note $\mathbf{A}_{\mu;\mu} = \mathbf{A}_{\mu,\mu}$)

$$\mathbf{A}_{\mu,\mu} = 0. \tag{22}$$

As $m^2 \to 0$ the classical solution of (20) approaches solutions of the zero mass equation (14), solved with the subsidiary condition, equation (22). The second order operator R_2 is now, in momentum space (see equation (11)) $q^2\,\delta_{\mu\nu} - q_\mu q_\nu - m^2\,\delta_{\mu\nu}$, an operator which has the inverse "propagator,"

$$\mathrm{Prop} = -(\delta_{\mu\nu} - q_\mu q_\nu m^{-2})/(q^2 - m^2). \tag{23}$$

Thus, it leads to a definite quantum field theory without difficulties (we do not discuss divergence problems) if we proceed, in the usual way, by making dia-

grams with junctions determined by the nonquadratic part of the full Lagrangian and lines with the propagator (23) with m^2 replaced by $m^2 - i\epsilon$ for $\epsilon \to +0$. One possible way we shall try to define the theory with zero mass is as a limit of this theory with mass, as $m^2 \to 0$. The term $q_\mu q_\nu/m^2$ in the propagator makes it, at first sight, doubtful that such a limit exists, but for trees and single closed loops it does exist.

The subsidiary condition (22) suggests, for $m^2 = 0$, another modification which we shall call *YMD* (Yang-Mills plus divergence terms). It is obtained from the Yang-Mills Lagrangian with the divergence terms $(A_{\mu,\mu})^2$ left out, or

$$L_{YMD}(\mathbf{A}) = L_{YM}(\mathbf{A}) + \tfrac{1}{2}(\mathbf{A}_{\mu,\mu})^2 \tag{24}$$

so that

$$L_{2YMD}(\mathbf{A}) = \tfrac{1}{2}(\mathbf{A}_{\mu,\nu})\cdot(\mathbf{A}_{\mu,\nu}). \tag{25}$$

The equations of motion now become

$$R_\mu(\mathbf{A}) + \mathbf{A}_{\nu,\nu\mu} = \mathbf{J}_\mu \tag{26}$$

so that, for conserved sources (13), we have

$$\mathbf{A}_{\nu,\nu\mu;\mu} = 0. \tag{27}$$

This, of course, has a solution

$$\mathbf{A}_{\nu,\nu} = 0 \tag{22}$$

and therefore (26) is equivalent to the unmodified equation (14) solved with the subsidiary condition (22).

In quantum theory, *YMD* leads to a nonsingular propagator

$$-\delta_{\mu\nu}/(q^2 - i\epsilon). \tag{28}$$

The couplings are given by the nonquadratic parts of L_{YMD} (or L_{YM}, for they are the same). This is a definite quantum theory, and agrees with the limit of *YMM* as $m^2 \to 0$ for trees, but not for diagrams with closed loops. Further modifications are necessary. We discuss these later.

Thus, in classical theory we can define things either by L_{YM} or by L_{YMD}, or by L_{YMM} as $m^2 \to 0$. All three give the same physical result. In quantum theory, L_{YM} has difficulty in definition because of the singular propagator, but L_{YMD} and L_{YMM} as $m^2 \to 0$ agree for trees because the theory of trees is so close to classical theory. For diagrams with one closed loop, L_{YMM} as $m^2 \to 0$ is satisfactory and the loop when opened by the methods of *I* gives the corresponding trees. However, as we shall see, for a closed loop L_{YMD} is not satisfactory; upon opening the loop the correct tree is not obtained, and unitarity suffers. It must be modified by subtracting a contribution equal to a corresponding closed loop in which a space-time scalar (spin 0) isovector particle propagates coupled to A_μ (but not to other matter).

The difficulties will be discussed in detail below but some hint may be given here. It appears that we must arrange, for quantum mechanics, that the equation (27) must not be left as a consequence of a classical equation of motion, but

must dynamically come from a Lagrangian also. If P is a spin 0 isovector field, the Lagrangian

$$L_P = -\tfrac{1}{2}\mathbf{P}_{;\mu}\cdot\mathbf{P}_{,\mu} \tag{29}$$

leads to the equation of motion

$$\tfrac{1}{2}(\mathbf{P}_{;\mu,\mu} + \mathbf{P}_{,\mu;\mu}) = \mathbf{P}_{,\mu;\mu} - \tfrac{1}{2}\mathbf{A}_{\mu,\mu} \times \mathbf{P} = 0. \tag{30}$$

If we set

$$\mathbf{P} = \mathbf{A}_{\mu,\mu}, \tag{31}$$

equation (30) reduces to (27), so if $\mathbf{P} = \mathbf{A}_{\mu,\mu}$ is true initially, it will be always, and it appears that the addition of $\frac{1}{2}(\mathbf{A}_{\mu,\mu})^2$ to L_{YM} must be compensated by the subtraction of L_P.

QUANTUM THEORY

There are many ways suggested to pass from classical to quantum theory. The one we shall use, to study the difficulties, is the path integral method which says the quantum mechanical amplitude for an event arises from the path integral

$$\text{amp} = \int \exp\left[i \int L_T(A, \psi)\, d\tau\right] \mathscr{D}A\mathscr{D}\psi \tag{32}$$

over all fields subject to special boundary conditions, where L_T is the total Lagrangian and the integral is over all fields, symbolized by $\mathscr{D}A\mathscr{D}\psi$. If the matter field is a Dirac field suitable modifications are made either by using operator calculus, or by changing certain signs in final diagram expressions or, if you prefer, using scalar matter for an example or no matter field at all. In this paper we are concentrating on the real difficulty, which lies in the properties of the Yang-Mills, or gravity Lagrangians, and a detailed study has shown that addition of matter makes no new problems, provided its Lagrangian is invariant under the correct transformation (2). Therefore, let us leave the matter field out, for later integration perhaps, and simply study (32) with (12) substituted for $L_T(A, \psi)$:

$$\text{amp} = \int \exp\left[i \int \left[L_{YM}(\mathbf{A}) - \mathbf{J}_\mu\cdot\mathbf{A}_\mu\right] d\tau\right] \mathscr{D}A. \tag{33}$$

Expression (33) is meaningless, however. This is because there is a direction in which $\mathbf{A}$ can be changed which does not alter the integrand, namely the change (16) [in first order, (2)]. Thus, when moving A in this way, the integral is infinite. It is analogous to a multiple integral over many variables over an infinite range but the integrand does not depend on one of the variables. What the author did at first was try to ignore this difficulty and to proceed to compute a number of processes in detail (analogues of Compton effect, bremsstrahlung, radiationless scattering, vacuum polarization, and second-order scattering with scalar matter interacting with gravitational fields or Yang-Mills field quanta

instead of photons). The aim was to see if any difficulties arose, and to study the convergence problem, which will not concern us here. The difficulties which did arise having to do with the gauge invariance were then studied carefully by many tedious examples and finally by more formal and powerful methods. It is these latter results which will be described here; although they are more difficult to understand directly, they do permit a larger overview of the problem. The various theorems have all been checked by examples (in fact, in nearly all cases the examples suggested the theorems and guided the "more powerful methods" on what to prove).

The expression (32) was immediately converted to rules for diagrams. To be explicit, a *YM* field quantum is represented by a polarization four-vector isovector $\mathbf{e}_\mu$. Its coupling with matter in the Dirac case is $\boldsymbol{\tau}\cdot\mathbf{e}_\mu\gamma_\mu$, while matter propagates with its usual propagator; for example, $(\not{p} - m)^{-1}$. The term of third order in A in L_{YM}, namely $(\mathbf{A}_\mu \times \mathbf{A}_\nu)\cdot\mathbf{A}_{\mu,\nu}$, leads to an interaction where three field quanta come together. For such a junction with quanta of momenta q^a, q^b, q^c, (so $q^a + q^b + q^c = 0$) of polarizations **a**, **b**, **c**, the amplitude contribution is

$$(q_\nu^a - q_\nu^c)\mathbf{b}_\nu\cdot(\mathbf{a}_\mu \times \mathbf{c}_\mu) + (q_\nu^b - a_\nu^a)\mathbf{c}_\nu\cdot(\mathbf{b}_\mu \times \mathbf{a}_\mu) + (q_\nu^c - q_\nu^b)\mathbf{a}_\nu\cdot(\mathbf{c}_\mu \times \mathbf{b}_\mu).$$

The fourth-order term $\frac{1}{4}(A_\mu \times A_\nu)\cdot(A_\mu \times A_\nu)$ results in junctions where four field quanta come together.

A free quantum entering or leaving has a polarization $\mathbf{a}_\mu$ satisfying (from equation 10)

$$(q^a)^2\mathbf{a}_\mu - (q_\nu^a\mathbf{a}_\nu)q_\mu^a = 0. \tag{34}$$

If it is physical $(q^a)^2 = 0$ and, further, it is transverse,

$$q_\mu^a\mathbf{a}_\mu = 0. \tag{35}$$

On the other hand, (34) has solutions even if $(q^a)^2 \neq 0$, namely a quantum polarized in the direction of its momentum, a pure divergence

$$\mathbf{a}_\mu = q_\mu^a\boldsymbol{\alpha}. \tag{36}$$

Such an incoming field must have no physical effect. The result of substituting q_μ^a for $\mathbf{a}_\mu$ in the sums of all diagrams for a process always turns out to give zero. This is an expression of the gauge invariance condition. It also means that for a free photon, the polarization $\mathbf{a}_\mu$ is not unique, a gauge change $\mathbf{a}_\mu \to \mathbf{a}'_\mu + q_\mu^a\boldsymbol{\alpha}$ has no effect. Although equation (2) is nonlinear, in the quantum perturbation theory it becomes the condition that an extra perturbation of the form (36) acting to first order in all the diagrams in which it can act produces zero.

For the Yang-Mills propagator of the field quanta in the case $m^2 = 0$ (*YM*), we replace e_μ, e_ν at the ends of a virtual quantum by equation (28), $-\delta_{\mu\nu}/q^2$. This is not strictly the reciprocal of the singular R_μ, for R_μ has no reciprocal, but in analogy with electrodynamics it was hoped that this propagator would be satisfactory since the sources were conserved, $q_\mu J_\mu = 0$, and an extra photon

coupled via a gradient, (36), had no effect. For example, it was expected that the $-q_\mu q_\nu/m^2$ term of the propagator with mass, equation (23), would vanish for this reason, and the limit $m^2 \to 0$ could be taken to lead *YMM* to this *YMD*.

All this was confirmed with tree diagrams, and with trees no difficulties arose.

A difficulty arose with a closed loop. Its origin can be understood in the following way when considering the relation of the loop to the unitarity condition or to trees if the loop is opened. Let the q of one propagator of the loop be in the z direction and write the $-\delta_{\mu\nu}/q^2$ as $-e_\mu e_\mu/q^2$ or as

$$\frac{1}{q^2}(e_x e_x + e_y e_y + (e_z - e_t)(e_z + e_t)). \tag{37}$$

Thus, at the pole of q^2 we have as necessary, and expected, free quanta entering and leaving with transverse polarizations x, y. But we also have an extra quantum that we do not want, coupled as $e_z - e_t$ entering, and $e_z + e_t$ leaving.

At first sight we might expect this extra quantum entering as it does via $e_z - e_t$ (which in the direction of q, for $q_t = q_z$) to give zero. Does not gauge invariance, which we have maintained carefully, insure that the action of such a potential (as equation 36) gives zero, when diagrams are summed? But pure divergences of the type equation (36) give zero only for real physical circumstances when all the other quanta acting either have their sources explicit in the diagrams (so $\mathbf{a}_\mu = q_\mu \alpha$ can act directly on these sources, too), or are genuine external free quanta satisfying equation (34), so that they are purely transverse equation (35) or else they are pure divergences themselves, as in equation (36). The trouble is that the quantum *leaving* upon opening (37) is polarized in the direction $e_t + e_z$ or $(x, y, z, t) = (0, 0, -1, 1)$, which is neither transverse to $q = (0,0, Q, Q)$ nor in the q direction, and has no apparent source. Thus, the last term in (37) does not give zero upon opening the loop but contributes a residual. This residual is easily calculated and is a rather simple expression, so it was easy to find a rule for subtracting a diagram to eliminate it. The rule is formulated in connection with equation (29). In perturbation theory we subtract a term corresponding to a space-scalar isovector meson going around the loop. It propagates via $1/q^2$ and is coupled via $(A_\mu \times P) \cdot P_{,\mu}$. That is, the scalar meson is neither created nor destroyed but two meet at a junction with a vector quantum. With in and out momenta q^1 and q^2, isospin quality $\mathbf{p}^1$, $\mathbf{p}^2$, they couple to an isovector $\mathbf{a}_\mu(q^2 = q^1 + q^a)$ via

$$\mathbf{a}_\mu \cdot (\mathbf{p}_1 \times \mathbf{p}_2^*)(q_\mu^1 + q_\mu^2). \tag{38}$$

They appear only in closed loops, never as free mesons. Where a loop is opened, their contribution cancels the unwanted last terms of equation (37).

For tree diagrams, the scalar mesons do not appear and there is no difficulty with the propagator of equation (37). This may be seen, for on opening a line for a propagator in a tree in (37) the $e_t - e_z$ couples to a piece entirely disjoint from the $e_t + e_z$ piece and, therefore, being a gradient coupling it gives zero on this piece.

The questions can also be studied using *YMM* and the propagator (23), expecting to take the limit as $m^2 \to 0$. The reader must be warned, however, in doing that the free wave external quanta must first *not* be taken to satisfy (34),

but rather that equation appropriate to the $m^2 \neq 0$ case (namely, equation 34 with 0 replaced by $m^2\mathbf{a}_\mu$ on the right-hand side). If he uses (34) unchanged, the limit for trees or one loop will exist, but it will not agree with the aforementioned $m^2 = 0$ theory. With the proper modified equation (34), the limit again exists but is different, and agrees with the aforementioned theory.

We cannot here go further into these explanatory details via perturbation theory, which serves so admirably as a laboratory to learn about these matters, but will now go directly to a more general and abstract mathematical way of dealing with them. First we discuss the theory of trees, and can follow I directly, except to correct small details, for R_μ has no inverse. Next we discuss one loop, and prove a relation between *YMD* and *YMM* plus a scalar meson which verifies our statement about one such loop. Finally we say what we know about double loops.

TREES

In this section we discuss tree diagrams in the *YM* theory and show first that the obvious prescription for writing them down using the propagator $-\delta_{\mu\nu}/q^2$ gives results which are completely satisfactory and gauge invariant. Next we show that the theory with mass, for which no ambiguities arise, yields as $m^2 \to 0$, the same result as the previous $-\delta_{\mu\nu}/q^2$ propagator. In the next section we will discuss problems connected with the trees expected from opening closed loop diagrams. We shall use the notations and ideas of I in our discussions.

Let us formulate the theory of tree diagrams if we use the propagator $-\delta_{\mu\nu}/q^2$ which we symbolize as R_{2D}^{-1} to analyze the Lagrangian (12), (1). We separate $R_\mu(\mathbf{A})$ into a second order (in A) part $R_{2\mu}(\mathbf{A})$ of equation (11) and a part of higher order in A,

$$R_\mu(\mathbf{A}) = R_{2\mu}(\mathbf{A}) - K_\mu(\mathbf{A}) \tag{39}$$

and

$$R_{2\mu}(\mathbf{A}) = R_{2D\mu}(\mathbf{A}) - (\mathbf{A}_{\nu,\nu})_\mu. \tag{40}$$

This $R_{2D\mu}(\mathbf{A})$ is just $\mathbf{A}_{\mu,\nu\nu}$, of course. Then the construction of trees with R_{2D}^{-1} as the propagator means solving by iteration the equation

$$\mathbf{A}_\mu = R_{2D}^{-1}\{K_\mu(\mathbf{A}) + \mathbf{J}_\mu\} + \mathbf{A}_\mu^0 \tag{41}$$

where

$$\mathbf{J}_\mu = \bar{\psi}\boldsymbol{\tau}\gamma_\mu\psi \tag{42}$$

and an accompanying equation for matter

$$\psi = (i\gamma_\mu\nabla_\mu - m)^{-1}\{\gamma_\mu(\boldsymbol{\tau}\cdot\mathbf{A}_\mu)\psi\} + \psi^0 \tag{43}$$

where A^0 and ψ^0 are appropriate (see I) asymptotic functions representing the incoming and outgoing external lines, satisfying

$$(i\gamma_\mu\nabla_\mu - m)\psi^0 = 0$$

and

$$R_{2\mu}(\mathbf{A}^0) = R_{2D\mu}(\mathbf{A}^0) - (\mathbf{A}^0_{\nu,\nu})_{,\mu} = 0. \tag{44}$$

The resultant tree t, can then be expressed in terms of A, ψ. It is easiest to express the effect on it of one extra disturbing external line, say a field $\mathbf{F}_\mu$ acts on it in first order as

$$\delta t = \int \mathbf{F}_\mu \cdot (K_\mu(\mathbf{A}) + \mathbf{J}_\mu)\, d\tau. \tag{45}$$

Equation (43) means that ψ satisfies the equation of motion for its field (just below equation 12) and hence, its current equation (42) satisfies (13). Multiplying the equation (41) by R_{2D} and using (39), (40), and (44) shows that A satisfies

$$R_\mu(\mathbf{A}) = \mathbf{J}_\mu + \mathbf{A}_{\nu,\nu\mu} - \mathbf{A}^0_{\nu,\nu\mu}. \tag{46}$$

This appears, at first sight, not to be the expected equation (14), but to carry a residual $\mathbf{D}_{,\mu}$ where we write

$$\mathbf{D} = \mathbf{A}_{\nu,\nu} - \mathbf{A}^0_{\nu,\nu}. \tag{47}$$

However, taking the covariant derivative of each side of (46), we find

$$\mathbf{D}_{,\mu;\mu} = 0 = \mathbf{D}_{,\mu,\mu} - \mathbf{A}_\mu \times \mathbf{D}_{,\mu} \tag{48}$$

which we are solving (as one readily finds by taking the divergence of equation (41)) by the iteration

$$\mathbf{D} = R_{2D}^{-1}\{\mathbf{A}_\mu \times \mathbf{D}_{,\mu}\}. \tag{49}$$

Now asymptotically $\mathbf{A}_\mu = \mathbf{A}^0_\mu$, so that asymptotically $\mathbf{D} = 0$, so that (48) says $\mathbf{D} = 0$ everywhere (it has no source), so equation (46) implies that $\mathbf{A}$ does satisfy the expected equation (14). It does so with the special condition $\mathbf{A}_{\mu,\mu} = \mathbf{A}^0_{\mu,\mu}$. If, for example, $\mathbf{A}^0_\mu$ were always chosen to satisfy $\mathbf{A}^0_{\mu,\mu} = 0$, then $\mathbf{A}_{\mu,\mu}$ would be that solution of (14) satisfying the supplementary condition (22). (Of course, no restriction need be made on $\mathbf{A}^0_{\mu,\mu}$, so none exists on $\mathbf{A}_{\mu,\mu}$.)

This is satisfactory but to be complete we must now show that the tree value does not depend on $\mathbf{A}^0_\mu$ if it is changed by a gauge transformation. We need only prove this for arbitrary first order changes $\mathbf{A}^0_\mu \to \mathbf{A}^0_\mu + \nabla_\mu \boldsymbol{\alpha}$. That means only that we must show that a pure gradient external line (36) has zero coupling. This is easiest done by letting this line be the last line $\mathbf{F}_\mu$ in (45) to get

$$\delta t = \int \nabla_\mu \boldsymbol{\alpha} \cdot (K_\mu(\mathbf{A}) + \mathbf{J}_\mu)\, d\tau. \tag{50}$$

Integration by parts and noting that equations (8) and (13) imply (since $R_{2\mu,\mu} = 0$)

$$K_{\mu,\mu}(\mathbf{A}) + \mathbf{J}_{\mu,\mu} = \mathbf{A}_\mu \times (-R_\mu(\mathbf{A}) + \mathbf{J}_\mu)$$

so that in view of (46) and (47) we get

$$\delta t = \int \boldsymbol{\alpha} \cdot (\mathbf{A}_\mu \times \mathbf{D}_{,\mu})\, d\tau = 0 \tag{51}$$

since $\mathbf{D} = 0$. Therefore, the theory is invariant under gauge transformation, and pure divergence external lines have no influence.

The reader may be curious to know, when $\mathbf{A}^0_\mu$ is changed to $\mathbf{A}^0_\mu + \nabla_\mu \boldsymbol{\alpha}$, what happens to $\mathbf{A}_\mu$. We do not require this for our further studies, but it is interesting. It is simply changed by a gauge transformation (ψ is also changed in the same transformation) and becomes $\mathbf{A}_\mu + \nabla_\mu \boldsymbol{\beta} - \mathbf{A}_\mu \times \boldsymbol{\beta} = \mathbf{A}_\mu + \boldsymbol{\beta}_{;\mu}$. The condition that $\mathbf{D} = 0$ or $\mathbf{A}_{\mu,\mu} = \mathbf{A}^0_{\mu,\mu}$ becomes $\boldsymbol{\beta}_{,\mu,\mu} - \boldsymbol{\alpha}_{,\mu\mu} - (\mathbf{A}_\mu \times \boldsymbol{\beta})_{,\mu} = 0$ which is solved by iteration as

$$\boldsymbol{\beta} = R_{2D}^{-1}\{(\mathbf{A}_\mu \times \boldsymbol{\beta})_{,\mu}\} + \boldsymbol{\alpha} \tag{52}$$

since asymptotically $\boldsymbol{\beta} = \boldsymbol{\alpha}$. This equation has a kernel adjoint to (60).

Alternatively, we could define the trees as those calculated for *YMM* in the limit $m^2 \to 0$. For *YMM* we add $-m^2\mathbf{A}_\mu \cdot \mathbf{A}_\mu$ to the Lagrangian so the propagator becomes unique. Thus, the expression for trees from (1) is unambiguous, we solve (43) and, instead of (41),

$$\mathbf{A}_\mu = (R_{2D} - m^2)^{-1}(\delta_{\mu\nu} + m^{-2}\nabla_\mu\nabla_\nu)\{K_\nu(\mathbf{A}) + \mathbf{J}_\nu\} + \mathbf{A}^0_\mu \tag{41'}$$

where $\mathbf{A}^0_\mu$ now satisfies

$$R_{2\mu}(\mathbf{A}^0) - m^2\mathbf{A}^0_\mu = 0, \tag{44'}$$

which implies that

$$\mathbf{A}^0_{\mu,\mu} = 0. \tag{53}$$

Now equation (41') implies that A satisfies

$$R_\mu(\mathbf{A}) = \mathbf{J}_\mu + m^2\mathbf{A}_\mu, \tag{46'}$$

which implies $\mathbf{A}_{\mu,\mu} = 0$.

In view now of the relation preceding equation (51), whose left side vanishes from (46'), since $\mathbf{A}_\mu \times \mathbf{A}_\mu = 0$, the gradient terms in (41') vanish and we could have used instead

$$\mathbf{A}_\mu = (R_{2D} - m^2)^{-1}\{K_\nu(\mathbf{A}) + \mathbf{J}_\nu\} + \mathbf{A}^0_\mu. \tag{41''}$$

This now has an evident limit as $m^2 \to 0$ (in equation 41' the limit was not evident because of the m^{-2}). This limit is (41) with the supplementary condition (53) and therefore (14). We have already seen this supplementary requirement does not destroy the gauge invariance.

SINGLE CLOSED LOOP

We have seen that the single closed loop calculated (for $m^2 = 0$) with propagator R_{2D}^{-1} does not agree with expectations if the loop is opened. This is because the propagator generates some new unwanted terms according to (37). There $(e_z - e_t)(e_z + e_t)$ can be represented as a scattering in which a particle of momentum q_μ (now $q^2 = 0$) comes in polarized in the direction q_μ and goes out with the same momentum but polarized in the direction N_μ where N_μ is a four vector—(0 0 −1, 1) in our case—of zero length such that $q \cdot N \neq 0$. We shall assume that the correct theory is what would come if you would just close the

tree diagram with transverse quanta. If we call this loop formed on closing trees with transverse quanta "closed loop *YM*," we have

$$\text{Closed Loop } YM = \text{Closed Loop } YMD - X \tag{54}$$

where X is the result of closing a tree with two extra quanta of momentum q, polarization $q_\mu\boldsymbol{\alpha}$ and $N_\mu\boldsymbol{\alpha}/(q\cdot N)$. Closing the tree would be equivalent to calculating this extra tree and in place of $\boldsymbol{\alpha}\cdots\boldsymbol{\alpha}$ putting the propagator $1/q^2$ (and taking the trace of the SU_2 indices). During all this, the other external lines are represented by $\mathbf{A}_0$, etc., in the usual way. We calculate X now. We shall leave out the inessential complication of the matter field ψ, and take $J_\mu = 0$.

Therefore, we are able to calculate the tree X with an external gradient $F_\mu = \nabla_\mu\boldsymbol{\alpha}$ and another external photon, say $\mathbf{b}_\mu$, where

$$\mathbf{b}_{\mu,\mu} = \boldsymbol{\alpha}. \tag{55}$$

To do this, we can use the formula equation (45) where the $\mathbf{A}_\mu$ is to be calculated from (41) where $\mathbf{A}^0_\mu$ contains properly all the physically external lines and also (in first order, of course) the field $\mathbf{b}^*_\mu$. For that reason, equation (44) no longer holds but we have instead (defining $\mathbf{S}_\mu$)

$$R_{2\mu}(\mathbf{A}^0) = R_{2\mu}(\mathbf{b}) = \mathbf{S}_\mu.$$

Since $\mathbf{b}_\mu$ is a free wave, $q^2 = 0$, but it is not transverse so $R_{2\mu}(\mathbf{b}) = q^2\mathbf{b}_\mu - q_\mu q_\nu \mathbf{b}_\nu = +\mathbf{b}_{\nu,\nu\mu}$, so that we have from (55)

$$\mathbf{S}_\mu = \boldsymbol{\alpha}_{,\mu}. \tag{56}$$

The equation (41) now leads to a modification of equation (46), namely,

$$R_\mu(\mathbf{A}) = \mathbf{J}_\mu + \mathbf{A}_{\nu,\nu\mu} - \mathbf{A}^0_{\nu,\nu\mu} + \mathbf{S}_\mu, \tag{57}$$

but we recover our original equations if we now put

$$\mathbf{D} = \mathbf{A}_{\nu,\nu} - \mathbf{A}^0_{\nu,\nu} + \boldsymbol{\alpha} \tag{58}$$

so that $\mathbf{D}$ asymptotically is $\boldsymbol{\alpha}$ this time. Now our tree (50) becomes (following the steps to equation (51))

$$X = \int \boldsymbol{\alpha}\cdot(\mathbf{A}_\mu \times \mathbf{D}_{,\mu})\, d\tau. \tag{59}$$

We see from taking the covariant derivative of (57) that $\mathbf{D}$ satisfies (48) but that it is being solved by the iteration

$$\mathbf{D} = R_{2D}^{-1}\{\mathbf{A}_\mu \times \mathbf{D}_{,\mu}\} + \boldsymbol{\alpha}. \tag{60}$$

It is evident from inspection of (59) and (60) that X is simply the tree that would result from a closed loop diagram of a field $\mathbf{P}$ coming from the Lagrangian (29) or $\mathbf{P}_{,\mu}\mathbf{P}_{,\mu} + \mathbf{P}\cdot(\mathbf{A}_\mu \times \mathbf{P}_{,\mu})$, thus leading to the propagation equation (60) and the closed loop (59). This is most obvious if we choose the gauge $\mathbf{A}^0_{\mu,\mu} = 0$.

It is also readily seen that *YMM* leads to the same thing, as $m^2 \to 0$. The easiest way is first to calculate the closed loop via the propagator of *YMMD*, namely $(q^2 - m^2)^{-1}\,\delta_{\mu\nu}$ and show that it exceeds that calculated by

$$(q^2 - m^2)^{-1}(\delta_{\mu\nu} - q_\mu q_\nu/m^2)$$

by a term which is just what you would get if a scalar particle went around a loop with the Lagrangian

$$L = -\tfrac{1}{2}(\mathbf{P}_{;\mu}\cdot\mathbf{P}_{,\mu} - m^2\mathbf{P}\cdot\mathbf{P}); \tag{61}$$

therefore, with propagator $(q^2 - m^2)^{-1}$ and coupling $\mathbf{P}_{,\mu}\cdot(\mathbf{A}_\mu \times \mathbf{P})$. In all these new propagators, the limit as $m^2 \to 0$ may be taken. You can do all this arguing from tree closures just as before or by more formal methods. Because these latter methods may be interesting in further investigations, we describe them in the next section.

FORMAL THEORY OF SIMPLE LOOPS

We will prove in a more formal way that in diagrams containing up to one loop, if the propagator $\delta_{\mu\nu}$ is used, we must subtract the action of a particle acting under the action from the Lagrangian (29). The easiest way to proceed is to prove the corresponding theorem for *YMM*, the Yang-Mills theory with mass, and then to take the limit as m^2 approaches zero (which limit exists for one loop diagrams). We shall write *CLYMM* to represent the expression for one closed loop, for *YMM*, that is with propagator (23). We have shown, in the previous paper that the contribution for all closed loops is given by (see I, equation 39 and 40)

$$\int \exp\left\{i\int [L(A + D) - D\,\delta L(A)/\delta A - L(A)]\,d\tau\right\}\mathscr{D}D \tag{62}$$

where $L(A)$ is the appropriate Lagrangian and we integrate over all D such that asymptotically $D_{asym} \to 0$. To get the term corresponding to only simple closed loops, the expression in the exponent must be expanded to second order in D and higher terms dropped. This is evidently just the terms of second order in D in the expansion of $L(A + D)$ in powers of D. Call it $L''(D)$, a quadratic function of D containing A implicitly. For example, for the Lagrangian (19), a direct substitution of $\mathbf{A}_\mu + \mathbf{D}_\mu$ for $\mathbf{A}_\mu$ and selection of the second order terms shows that in this case

$$L''_{YMM}(D) = \tfrac{1}{4}\mathbf{F}_{\mu\nu}\mathbf{F}_{\mu\nu} + \tfrac{1}{2}\mathbf{E}_{\mu\nu}\cdot(\mathbf{D}_\mu \times \mathbf{D}_\nu) - \tfrac{1}{2}m^2\mathbf{D}_\mu\cdot\mathbf{D}_\mu. \tag{63}$$

Here $\mathbf{E}_{\mu\nu}$ is given by (5), a function just of $\mathbf{A}$, while

$$\mathbf{F}_{\mu\nu} = \mathbf{D}_{\mu;\nu} - \mathbf{D}_{\nu;\mu} \tag{64}$$

with the definition of the covariant derivative being that appropriate to just the external field A given by (3). We here explicitly suppose that there is no matter

present in addition to the field in order to avoid inessential mathematical complications. The reader can reinstate the terms representing matter and its coupling to the field and verify that they have no effect on the sense of the relations we shall prove. Therefore,

$$CLYMM = \int \exp\left[i \int L''_{YMM}(D)\, d\tau\right] \mathscr{D}D. \tag{65}$$

In the same way for the Yang-Mills theory with mass with divergence terms left out, so the propagator is $\delta_{\mu\nu}(q^2 - m^2)^{-2}$, the closed loop expression is

$$CLYMMD = \int \exp\left\{i \int [L''_{YMM}(D) + \tfrac{1}{2}(D_{\mu,\mu})^2]\, d\tau\right\} \mathscr{D}D. \tag{66}$$

We shall show that leaving out the divergence terms does make a difference, (66) is not equal to (65). However, if we subtract from each loop of (66) the action of a scalar particle, they do become equal. That is, we shall prove

$$CLYMM = CLYMMD/CLSM \tag{67}$$

where $CLSM$ is the closed loop from a scalar particle of mass m^2;

$$CLSM = \int \exp\left\{-\frac{i}{2} \int (\mathbf{P}_{;\mu}\mathbf{P}_{,\mu} - m^2\mathbf{P}\cdot\mathbf{P})\, d\tau\right\} \mathscr{D}\mathbf{P}. \tag{68}$$

Since the right-hand side of (67) evidently has a limit as $m^2 \to 0$ (because all propagators are nonsingular), we have proven that the left side does have a limit even though the singular kernel makes that less than obvious. We may therefore take the right side substituting $m^2 = 0$ (that is, $CLYMD/CSL$) as the single closed loop action for the YM theory with zero mass. This is what we wished to prove.

We do not really prove (67) but only that both sides are equal within a constant (independent of A) normalizing factor (infinite as $m^2 \to 0$), but such a constant factor in path integrals can be included in their definitions and has no effect on the physics that results from them. We now proceed to a proof of (67). First, in (65), replace $\mathbf{D}_\mu$ by $\mathbf{D}'_\mu$ and then substitute

$$\mathbf{D}'_\mu = \mathbf{D}_\mu + \boldsymbol{\alpha}_{;\mu} \tag{69}$$

for arbitrary $\boldsymbol{\alpha}$. Evidently, $\mathscr{D}\mathbf{D}' = \mathscr{D}\mathbf{D}$ and $\mathbf{F}'_{\mu\nu} = \mathbf{F}_{\mu\nu} - \mathbf{E}_{\mu\nu} \times \boldsymbol{\alpha}$ in virtue of (6). Hence,

$$\begin{aligned} L''_{YMM}(\mathbf{D}') = L''_{YMM}(\mathbf{D}) &- \tfrac{1}{2}(\mathbf{E}_{\mu\nu} \times \boldsymbol{\alpha})\cdot\mathbf{F}_{\mu\nu} + (\mathbf{E}_{\mu\nu}\cdot\mathbf{D}_\mu \times \boldsymbol{\alpha}_{;\nu}) \\ &+ \tfrac{1}{4}(\mathbf{E}_{\mu\nu} \times \boldsymbol{\alpha})\cdot(\mathbf{E}_{\mu\nu} \times \boldsymbol{\alpha}) + \tfrac{1}{2}\mathbf{E}_{\mu\nu}\cdot(\boldsymbol{\alpha}_{;\mu} \times \boldsymbol{\alpha}_{;\nu}) \\ &- m^2\boldsymbol{\alpha}_{;\mu}\mathbf{D}_\mu - \frac{m^2}{2}\boldsymbol{\alpha}_{;\mu}\boldsymbol{\alpha}_{;\mu}. \end{aligned} \tag{70}$$

When calculating $\int L''_{YMM}(\mathbf{D}')\, d\tau$, however, the third and fifth terms containing $\boldsymbol{\alpha}_{;\mu}$ can be integrated by parts, using (7) to produce

$$\int L''_{YMM}(\mathbf{D}')\, d\tau = \int \left\{L''_{YMM}(\mathbf{D}) - (\mathbf{D}_\mu + \tfrac{1}{2}\boldsymbol{\alpha}_{;\mu})\cdot(m^2\boldsymbol{\alpha}_{;\mu} + \mathbf{R}_\mu \times \boldsymbol{\alpha})\right\} d\tau.$$

However, the external field **A** is such that $\mathbf{R}_\mu(\mathbf{A}) = m^2\mathbf{A}_\mu$ (equation (20), with $\mathbf{J}_\mu = 0$, for we are omitting matter terms) and this substitution tells us

$$CLYMM = \int \exp\left\{i \int [L''_{YMM}(\mathbf{D}) - m^2\mathbf{D}_\mu\boldsymbol{\alpha}_{,\mu} - \tfrac{1}{2}m^2\boldsymbol{\alpha}_{;\mu} \times \boldsymbol{\alpha}_{,\mu}]\, d\tau\right\}\mathscr{D}\mathbf{D} \quad (71)$$

valid for any α.

We use this transformation in the following way. First, multiply the original expression (65) with **D** called **D′** by the path integral (a functional of **A**)

$$X = \int \exp\left\{\frac{i}{2}\int (m^2\boldsymbol{\alpha}' + \boldsymbol{\alpha}'_{;\mu,\mu})^2\, d\tau\right\}\mathscr{D}\boldsymbol{\alpha}' \quad (72)$$

to get

$$X\cdot CLYMM = \int \exp\left\{\frac{i}{2}\int [L''_{YMM}(\mathbf{D}') + \tfrac{1}{2}(m^2\boldsymbol{\alpha}' + \boldsymbol{\alpha}'_{;\mu,\mu})^2]\, d\tau\right\}\mathscr{D}\mathbf{D}'\mathscr{D}\boldsymbol{\alpha}'. \quad (73)$$

Now define the linear operator G:

$$G\boldsymbol{\alpha}' = \boldsymbol{\alpha}'_{;\mu,\mu} + m^2\boldsymbol{\alpha}' \quad (74)$$

to substitute $\boldsymbol{\alpha}'$ by

$$\boldsymbol{\alpha}' = \boldsymbol{\alpha} - G^{-1}\mathbf{D}'_{\mu,\mu} \quad (75)$$

so $G\boldsymbol{\alpha}' = G\boldsymbol{\alpha} - \mathbf{D}'_{\mu,\mu}$ and $\mathscr{D}\boldsymbol{\alpha}' = \mathscr{D}\boldsymbol{\alpha}$, so that we get

$$X\cdot CLYMM = \int \exp\left\{\frac{i}{2}\int [L''_{YMM}(\mathbf{D}') + \tfrac{1}{2}(\mathbf{D}'_{\mu,\mu} - m^2\boldsymbol{\alpha} - \boldsymbol{\alpha}_{;\mu,\mu})^2]\, d\tau\right\}\mathscr{D}\mathbf{D}'\mathscr{D}\boldsymbol{\alpha}.$$

Now substitute **D′** via (69) and use the algebraic steps leading from (65) to (71) to get

$$\begin{aligned} X\cdot CLYMM &= \int \exp\left\{\frac{i}{2}\int [L''_{YMM}(\mathbf{D}) - m^2\mathbf{D}_{,\mu}\boldsymbol{\alpha}_{,\mu} - \tfrac{1}{2}m^2\boldsymbol{\alpha}_{,\mu}\boldsymbol{\alpha}_{;\mu}\right. \\ &\qquad \left. + \tfrac{1}{2}(\mathbf{D}_{\mu,\mu} - m^2\boldsymbol{\alpha})^2]\, d\tau\right\}\mathscr{D}\mathbf{D}\mathscr{D}\boldsymbol{\alpha} \\ &= \int \exp\left\{\frac{i}{2}\int [L''_{YMM}(\mathbf{D}) + \tfrac{1}{2}(\mathbf{D}_{\mu,\mu})^2 - \tfrac{1}{2}m^2\boldsymbol{\alpha}_{,\mu}\boldsymbol{\alpha}_{;\mu}\right. \\ &\qquad \left. + \tfrac{1}{2}m^4\boldsymbol{\alpha}^2]\, d\tau\right\}\mathscr{D}\mathbf{D}\mathscr{D}\boldsymbol{\alpha}. \end{aligned} \quad (76)$$

The integral on **D**, **α** has again separated, that on **D** being just *CLYMMD* and that on **α** (putting $\boldsymbol{\alpha} = \mathbf{P}/m$, and changing normalization constants, so $\mathscr{D}\boldsymbol{\alpha} \to \mathscr{D}\mathbf{P}$) is just *CLSM* of (68) so that in (76) we have shown that

$$X\cdot CLYMM = CLYMMD\cdot CLSM. \quad (77)$$

To finish the proof of (67), we must merely show that X defined in (72) is

$$X = (CLSM)^2. \quad (78)$$

This can be seen by noting, with the definition (74),

$$X = \int \exp\left\{\frac{i}{2}\int (\boldsymbol{\alpha} G G \boldsymbol{\alpha})\, d\tau\right\} \mathscr{D}\boldsymbol{\alpha} \tag{79}$$

while

$$CLSM = \int \exp\left\{\frac{i}{2}\int (\mathbf{P} G \mathbf{P})\, d\tau\right\} \mathscr{D}\mathbf{P}. \tag{80}$$

Gaussian integrals are proportional to the reciprocal square root of the determinant of the coefficient matrix in the quadratic exponent. Therefore, (80) is proportional to the $-1/2$ power of the functional determinant of G, while for (79) it is G^2. But the determinant of the square of a matrix is the square of the determinant of the matrix so that, except for an unimportant constant factor, as far as dependence on A is concerned, equation (78) holds, and equation (67) is established.

We may now see what is going on if we attempt to calculate $CLYM$ directly for $m^2 = 0$. When we do, the integrals on $\mathbf{D}_\mu$ usually have a convergent integral but now there is one "direction" into which we may move $\mathbf{D}_\mu$, namely, $\mathbf{D}_\mu \to \mathbf{D}_\mu + \boldsymbol{\alpha}_{\mu;\mu}$, which does not affect the integrand (see equation 71 for $m^2 = 0$) so the integral in this "direction" diverges. We attempt to put a convergence factor on it of the form $m^2\mathbf{D}_\mu\mathbf{D}_\mu$ and everything turns out to be all right as $m^2 \to 0$. But if instead we try $\frac{1}{2}(\mathbf{D}_{\mu,\mu})^2$ as a convergence factor ($CLYMD$), we are controlling the integral in a way that depends on A, and we shall have to divide by CLS to take out this dependence.

As a final note, we give another way of dealing with things that appear to be simpler and to make the gauge invariance of the final result more obvious. On the other hand, however, it seems not to give the same result. Suppose, first, we form a different version of YMD, say YMD', by modifying L''_{YMD} of (63) by adding the more obviously covariant $\frac{1}{2}\mathbf{D}_{;\mu}\mathbf{D}_{;\mu}$, instead of $\frac{1}{2}\mathbf{D}_{,\mu}\mathbf{D}_{,\mu}$ as in (66). Next, suppose that in calculating $CLYMM$ from (65), the external field $\mathbf{A}_\mu$ is calculated for the case $m^2 = 0$, and is not altered when we consider adding the mass term $m^2\mathbf{D}_\mu\mathbf{D}_\mu$ and later varying m^2 toward its limit zero. Call this $CLYMM'$. This means that we use $R_\mu(\mathbf{A}) = 0$ (rather than $R_\mu(\mathbf{A}) = m^2\mathbf{A}_\mu$) in the steps after (70). Then in (72) and (74) replace $\alpha_{;\mu,\mu}$ by $\alpha_{;\mu;\mu}$, and in (75) use $\mathbf{D}'_{\mu;\mu}$ for $\mathbf{D}'_{\mu,\mu}$. What one finally finds is that

$$CLYMM' = CLYMMD'/CLSM' \tag{67'}$$

where

$$CLYMMD' = \int \exp\left\{i\int [L''_{YMM}(D) + \tfrac{1}{2}(\mathbf{D}_{\mu;\mu})^2]\, d\tau\right\} \mathscr{D}\mathbf{D} \tag{66'}$$

and

$$CLMS' = \int \exp\left\{-\frac{i}{2}\int [\mathbf{P}_{;\mu}\mathbf{P}_{;\mu} - m^2\mathbf{P}\cdot\mathbf{P}]\, d\tau\right\} \mathscr{D}\mathbf{P}. \tag{68'}$$

The quantity $CLYMM'$ is calculated from an expression that looks exactly like (65) except $\mathbf{A}_\mu$ is different. The $\mathbf{A}_\mu$ is calculated with $m^2 = 0$; it satisfies $\mathbf{R}_\mu(\mathbf{A}_\mu) = 0$. Now the result on the right-hand side of (67') has a limit for $m^2 \to 0$, namely

just put $m^2 = 0$ in the Lagrangians. One is tempted to define this expression for $m^2 = 0$ as the value for one closed loop for the YM theory; it is more evidently gauge invariant than our previous expression, the right side of (67) with $m^2 = 0$. However, it is not evident that this limit of (67′) is the same as the limit (67). Direct calculations of examples indicate that they are not the same. Expression (67) seems more satisfactory because it comes from the trees argument, and it is the limit of a complete theory (YMM) using the same rules for diagrams for all lines whether they be external or are on a loop. In the gravitational theory, in which the method of including a mass in the original equations is not completely obvious (although it is in the second order Lagrangian L''), it is much easier to find and define the analogue of (67′) (see the Appendix).

PROCESSES WITH MORE THAN ONE CLOSED LOOP

Processes with more than one closed loop can probably be defined and calculated in terms of lower order processes by using unitarity to obtain the imaginary part and a dispersion relation to obtain the real part. Since the lower order processes (even up to one loop) have been successfully defined, probably a complete theory now exists, at least in principle. However, I have tried to find an explicit formula for these higher processes in the massless case in terms of rules for diagrams so that an explicit expression for a diagram of any complexity can be written down. I have not yet succeeded in doing this, but should like to report in this section on a few observations and relations noticed in attempting to do this.

The problem is to find a definition of closed loop diagrams for $m^2 = 0$ (say for definiteness, one with two closed loops), so that the result is consistent with unitarity and also with gauge invariance. The latter condition takes the form of the condition that an additional perturbing potential that is a pure gradient, $\delta A_\mu = \nabla_\mu \alpha$, will have no effect in first order.

We cannot define the answer directly as a path integral, for that integral is undefined because of the singular kernel ($m^2 = 0$). We expect, by analogy to the one loop case, that we could give the answer in terms of the YMD theory, which is not singular, with some modification (like subtracting the contribution from a scalar particle) but I do not know what the modification must be in the general case.

Alternatively, we could try to define the answer (for two loops, say) by putting together lower diagrams with one loop, according to the principles of [I]. However, here the parts which must be assembled (see [I], last section) are not complete sets of diagrams for a complete process in lower order. For a set of diagrams which is not a complete process, the result of the gradient potential is not zero but is somewhat more complicated, so it becomes difficult to insure gauge invariance. I have not yet carried it through carefully to see where this leads.

The attempt that I did try at first, and the only one for which the study is virtually complete, was to try to define the general result as the limit of YMM as $m^2 \to 0$. First I explicitly calculated the result for two connected loops (as in

the Greek letter θ) in an external potential for the Yang-Mills field with $m^2 \neq 0$ using the propagator (23) or rather its analogue for propagation in an external field $\mathbf{A}_\mu$. I then compared the result to what one calculates with propagator analogous to $\delta_{\mu\nu}(q^2 - m^2)^{-1}$ (that is, of *YMMD* in an external potential) and calculated the difference. The idea was to find an expression for the difference so I would know just how *YMMD* must be modified to get *YMM*. If this modification had a limit as $m^2 \to 0$, I would find the natural modification of *YMD* needed to define and calculate *YM*.

For two coupled loops there are three propagators, so the $q_\mu q_\mu / m^2$ of equation (23) leads at first sight to terms of order m^{-6}, and having no limit, as $m^2 \to 0$. But we learned in the one loop case that the m^{-2} term can be simplified, and in fact does have a limit. I have succeeded in simplifying the m^{-6} terms and have gotten rid of terms of order m^{-6} and m^{-4} but am still left with a term of order m^{-2} which I am unable to simplify further. I conclude that for two coupled loops there is no limit as $m^2 \to 0$, and therefore, in general, the Yang-Mills theory as usually formulated and interpreted in diagrams in the usual way is a theory that does not have a limit as $m^2 \to 0$ and that such a limit cannot serve as a definition of *YM*. (By Yang-Mills with mass, we mean to add simply $-\frac{1}{2}m^2\mathbf{A}_\mu \cdot \mathbf{A}_\mu$ to the Lagrangian. This is arbitrary; perhaps other terms depending on m might be added to make a better definition, and one that does yield a limiting theory as $m^2 \to 0$.)

I must therefore return to other methods of defining *YM* for $m^2 = 0$; I have not yet done this.

During the course of investigating these matters, I did find a way to express the m^{-2} terms which I was able to generalize to all orders of closed loops in the complete theory. What I did first was to express the m^{-2} terms in terms of the *YMMD* propagator so that limits could be easily studied, but I could not guess at the general form of this expression for loops of arbitrary complexity. However, when I expressed these m^{-2} terms in terms of the *YMM* propagator I was able to see the generalization. This relation, however, was no longer in a form that permits one to see that the limit surely does not exist, but it is nevertheless interesting, and we discuss it in the next section.

RELATION OF *YMMD* TO *YMM*

Ideally we should like to express the complete *YMM* in terms of *YMMD* with some modification, and study the modification as $m^2 \to 0$. The expression may be difficult because in first order it says that to find *YMM* one should *subtract* from *YMMD* the result of a scalar particle loop. This subtraction process may be hard to define and find in higher order. I thought, therefore, that since this also says that in first order *YMMD* is equivalent to the action of a pure vector particle *YMM* plus a scalar particle, such a relation is physical and might be generalized. Therefore, although it is not what I really desired, I was led to see whether I could express *YMMD* as the action of a vector and a scalar particle (both acting nonlinearly, and in interaction). It is possible and the result is given here.

We shall no longer separate out tree diagrams and closed loop diagrams, but consider the theory as a whole in all orders as a path integral.

To do this, we first express the amplitude for the Yang-Mills theory with $m^2 \neq 0$ and with the divergence term added (omitting other matter for simplicity of exposition) as a path integral, as in equation (33),

$$\text{amp}\,(YMMD) = \int \exp\left\{i \int [L_{YM}(\mathbf{A}) + \tfrac{1}{2}(\mathbf{A}_{\mu,\mu})^2 - \tfrac{1}{2}m^2\mathbf{A}_\mu\cdot\mathbf{A}_\mu]\, d\tau\right\}\mathscr{D}\mathbf{A}. \quad (81)$$

We may multiply by the inessential constant

$$\text{const} = \int e^{-i/2m^2\mathbf{P}^2}\mathscr{D}\mathbf{P} \quad (82)$$

where $\mathbf{P}$ is a space-scalar isospin-vector field, to get

$$\text{amp}\,(YMMD) = \int \exp\left\{i \int [L_{YM}(\mathbf{A}) + \tfrac{1}{2}(\mathbf{A}_{\mu,\mu})^2 - \tfrac{1}{2}m^2\mathbf{A}_\mu\cdot\mathbf{A}_\mu - \tfrac{1}{2}m^2\mathbf{P}^2]\, d\tau\right\}\mathscr{D}\mathbf{A}\mathscr{D}\mathbf{P}. \quad (83)$$

We shall now make some transformations of the integrand in the exponent. First we substitute $\mathbf{P} \to \mathbf{P} - m^{-1}\mathbf{A}_{\mu,\mu}$ (this does not change $\mathscr{D}\mathbf{P}$) and perform an integration by parts to convert this integrand to

$$L_{YM}(\mathbf{A}) - \tfrac{1}{2}m^2\mathbf{A}_\mu\cdot\mathbf{A}_\mu - m\mathbf{A}_\nu\cdot\mathbf{P}_{,\nu} - \tfrac{1}{2}m^2\mathbf{P}^2. \quad (84)$$

Now we shall make a finite gauge transformation of $\mathbf{A}_\mu$ to $\mathbf{A}'_\mu$, using the notation of equation (16) with $a_\mu = \boldsymbol{\tau}\cdot\mathbf{A}_\mu$ with

$$\sigma = \exp(-im^{-1}\boldsymbol{\tau}\cdot\mathbf{P}). \quad (85)$$

This leaves the expression for $L_{YM}(\mathbf{A})$ unchanged, or $L_{YM}(\mathbf{A}) = L_{YM}(\mathbf{A}')$. The volume element in the path integral space remains unchanged in form, because the addition of $\mathbf{T}_\mu$ does not change the differential $\mathscr{D}\mathbf{A}$ and the rotation of σ among the isospin components at each space time point makes no change for we are integrating with respect to all of the components. The expressions $\mathbf{A}_\mu\cdot\mathbf{A}_\mu$ and $\mathbf{A}_\mu\cdot\mathbf{P}_{,\nu}$ are changed, however, and we are left in the integrand with

$$L_{YM}(\mathbf{A}') - \tfrac{1}{2}m^2\mathbf{A}'_\mu\cdot\mathbf{A}'_\mu + m^2\mathbf{A}'_\mu\cdot\sigma(\mathbf{T}_\mu - m^{-1}\mathbf{P}_{,\mu})\sigma^{-1} - \tfrac{1}{2}m^2\mathbf{T}_\mu\cdot\mathbf{T}_\mu + m\mathbf{T}_\mu\cdot\mathbf{P}_{,\mu} - \tfrac{1}{2}m^2\mathbf{P}^2. \quad (86)$$

Next we expand $\mathbf{T}_\mu$ explicitly as in (17′) but with $\boldsymbol{\alpha} = -m^{-1}\mathbf{P}$. The quantity $\sigma\mathbf{P}_{,\mu}\sigma^{-1}$ is similarly expanded just as in (15′). More strictly, we should have written the term $\mathbf{A}'_\mu\cdot\sigma(\mathbf{T}_\mu - m^{-1}\mathbf{P}_{,\mu})\sigma^{-1}$ as $Tr(a'_\mu\sigma(t_\mu - m^{-1}\boldsymbol{\tau}\cdot\mathbf{P}_\mu)\sigma^{-1})$ but we hope the meaning was clear. We obtain, after a number of integrations by parts, the result that the integrand in the exponent of (83) may be replaced by (dropping the prime on $\mathbf{A}'_\mu$)

$$L_{YM}(\mathbf{A}) - \tfrac{1}{2}m^2\mathbf{A}_\mu\cdot\mathbf{A}_\mu - \frac{m^2}{2}\mathbf{P}\cdot\mathbf{P} + \mathbf{P}_{;\nu}\cdot\left\{\tfrac{1}{2}\mathbf{P}_{,\nu} + \frac{1}{3m}\mathbf{P}\times\mathbf{P}_{,\nu} + \frac{1}{8m^2}\mathbf{P}\times(\mathbf{P}\times\mathbf{P}_{,\nu}) + \cdots\right\}. \quad (87)$$

Here $\mathbf{P}_{;\nu} = \mathbf{P}_\nu - \mathbf{A}_\nu \times \mathbf{P}$ and the series continues, with the n-th term being

$$\mathbf{P}_{;\nu} \cdot \{((n+2)n!\, m^n)^{-1} (\mathbf{P} \times)^n \mathbf{P}_{,\nu}\}. \tag{88}$$

The first two terms are the Lagrangian $L_{YMM}(\mathbf{A})$ of the usual Yang-Mills theory with mass. The remaining terms are the nonlinear Lagrangian of a spin zero isospin vector particle of mass m^2 coupled to the vector field $\mathbf{A}_\mu$ and to itself. The Lagrangian of this field is

$$L_{\text{SCALAR } M}(\mathbf{P}) = -\tfrac{1}{2}m^2 \mathbf{P}\cdot\mathbf{P} + \mathbf{P}_{;\nu} \cdot \left\{ \tfrac{1}{2}\mathbf{P}_{,\nu} + \frac{1}{3m}\mathbf{P} \times \mathbf{P}_{,\nu} + \frac{1}{8m^2}\mathbf{P} \times (\mathbf{P} \times \mathbf{P}_{,\nu}) + \cdots \right\}. \tag{89}$$

This is what we set out to prove in this section.

The expression to third order (and partly to fourth) has been confirmed by actual calculations on double loops. In fact, it was from such calculations that the form of the third and higher terms were guessed, and the proof was only found after the form was known.

When matter is present, no change in the scalar particle Lagrangian results; it couples purely to itself and $\mathbf{A}_\mu$.

To second order in P, the Lagrangian is precisely equation (61). Thus, for a single loop, where this order is sufficient, we have that *YMMD* is *YMM* plus a scalar particle (61), or *YMM* is *YMMD* minus this particle, a result which we proved before in another way. To this order, we can take the limit $m^2 \to 0$.

The theorem to all orders is not useful for the case $m^2 \to 0$, for the expression (87) is meaningless in this limit. Since *YMMD* does have a limit as $m^2 \to 0$, this suggests (but perhaps does not directly prove) that *YMM* has no such limit.

To answer such questions, it would be much more satisfactory to have an expression for *YMM* in terms of *YMMD* and modifications instead of the other way around, but this I have not derived.

At least one can prove from (89) that terms in perturbation theory having just two closed loops in the form of a θ do not have a limit in the Yang-Mills theory with mass as $m^2 \to 0$. Direct calculation of these terms have confirmed this result.

APPENDIX

The quantization of the gravitational field presents us with problems nearly completely analogous to those of the Yang-Mills theory of the text. We point out some of these analogies in outline in this appendix in which the gravitational field equations are developed following the text development for the Yang-Mills field. The equations will be numbered as (*na*) where (*n*) is the analogous text equation. We shall formulate the gravitational field in the special case that it is

in interaction with a scalar matter field (our problems are independent of the kind of matter used[3]).

$$L_{\text{matter}} = \mathscr{T}(g, \phi) = \sqrt{g}\,\tfrac{1}{2}(g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} - m^2\phi^2) \tag{1a}$$

where $g_{\mu\nu}$ is the metric tensor, g its determinant, and $g^{\mu\nu}$ its reciprocal. The field Lagrangian is that of Einstein:

$$L_G = \mathscr{R}(g) \tag{4a}$$

where $\mathscr{R}$ is the scalar curvature tensor density or $\mathscr{R} = \mathscr{R}_{\mu\nu}g^{\mu\nu}$ with $\mathscr{R}_{\mu\nu} = \sqrt{g}(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R^{\sigma}_{\sigma})$, where $R_{\mu\nu} = R^{\sigma}_{\mu\nu\sigma}$ and $R^{\sigma}_{\mu\nu\rho}$ is the Riemann curvature tensor. This Lagrangian is invariant under a general coordinate transformation $x^{\mu} \to f^{\mu}(x^{\alpha})$ in which transformation the field quantities $g_{\mu\nu}$ are transformed to

$$g'_{\alpha\beta} = (g_{\mu\nu}f^{\mu}{}_{,\alpha}f^{\nu}{}_{,\beta})_f. \tag{16a}$$

We must substitute $f^{\mu}(x)$ for x^{μ} in the functional form of $g_{\mu\nu}$. Commas denote ordinary differentiation. For infinitesimal transformations, $f^{\mu} = x^{\mu} + \eta^{\mu}(x)$, this becomes

$$g_{\mu\nu} \to g_{\mu\nu} + \eta^{\sigma}{}_{,\mu}g_{\sigma\nu} + \eta^{\sigma}{}_{,\nu}g_{\sigma\mu} + \eta^{\sigma}g_{\mu\nu,\sigma} \tag{2a}$$

to first order in η. We call (2a) a gauge transformation. We define the covariant derivative of a vector (which thereby becomes a tensor) as

$$A^{\mu}_{;\nu} = A^{\mu}_{,\nu} + \left\{ \begin{matrix} \mu \\ \nu \quad \sigma \end{matrix} \right\} A^{\sigma} \tag{3a}$$

in the usual way (including its usual generalization to differentiating higher tensors). For example, if $\mathscr{S}^{\mu\nu}$ is a tensor density,

$$\mathscr{S}^{\mu\nu}_{;\nu} = \mathscr{S}^{\mu\nu}_{,\nu} + \left\{ \begin{matrix} \mu \\ \sigma \quad \tau \end{matrix} \right\} \mathscr{S}^{\sigma\tau}.$$

Here,

$$\left\{ \begin{matrix} \mu \\ \sigma \quad \tau \end{matrix} \right\} = \tfrac{1}{2}g^{\mu\nu}[g_{\sigma\nu,\tau} + g_{\tau\nu,\sigma} - g_{\sigma\tau,\nu}].$$

The Riemann tensor can then be defined by

$$A^{\mu}_{;\nu\sigma} - A^{\mu}_{;\sigma\nu} = R^{\mu}_{\lambda\nu\sigma}A^{\lambda}. \tag{6a}$$

The variation of $\int L_G(g)\,d\tau$ with respect to $g_{\mu\nu}$ is $-\mathscr{R}^{\mu\nu}(g)$. The fact that $\int L_G\,d\tau$ is invariant under the transformation (2a) implies that the tensor density $\mathscr{R}^{\mu\nu}$ satisfies

$$\mathscr{R}^{\mu\nu}_{;\nu} = \mathscr{R}^{\mu\nu}_{,\nu} + \left\{ \begin{matrix} \mu \\ \sigma \quad \tau \end{matrix} \right\} \mathscr{R}^{\sigma\tau} = 0. \tag{8a}$$

[3] This coupling is not unique. For a full discussion of this question and for discussions of coupling to other matter fields (for example, spin $\frac{1}{2}$), see Elisha Huggins, Ph.D. thesis, California Institute of Technology.

By integrating by parts, the Lagrangian may also be written as

$$\int L_G(g)\,d\tau = \int \sqrt{g}\,g^{\mu\nu}\left[\begin{Bmatrix}\lambda\\ \sigma\ \ \nu\end{Bmatrix}\begin{Bmatrix}\sigma\\ \lambda\ \ \mu\end{Bmatrix} - \begin{Bmatrix}\lambda\\ \mu\ \ \nu\end{Bmatrix}\begin{Bmatrix}\sigma\\ \lambda\ \ \sigma\end{Bmatrix}\right] d\tau.$$

This $\mathscr{R}^{\mu\nu}$ is a nonlinear operator. If $g_{\mu\nu}$ is expanded as $\delta_{\mu\nu} + h_{\mu\nu}$, the part $\mathscr{R}_2^{\mu\nu}$ of $\mathscr{R}^{\mu\nu}$ linear in $h_{\mu\nu}$ comes from the quadratic part of the Lagrangian

$$L_{2G}(h) = +\tfrac{1}{2}(\bar{h}_{\mu\nu,\sigma}h_{\mu\nu,\sigma}) - \bar{h}_{\mu\nu,\nu}\bar{h}_{\mu\sigma,\sigma} \tag{9a}$$

where we define the bar operation on a tensor (now, in the Minkowski sense, not a tensor in the fully covariant sense) by

$$\bar{C}_{\mu\nu} = \tfrac{1}{2}(C_{\mu\nu} + C_{\nu\mu} - \delta_{\mu\nu}C_{\sigma\sigma})$$

(repeated, the bar operation on a symmetric tensor restores the original tensor). Thus

$$\bar{\mathscr{R}}_{2\mu\nu}(h) = h_{\mu\nu,\sigma\sigma} - \bar{h}_{\mu\sigma,\sigma\nu} - \bar{h}_{\nu\sigma,\sigma\mu}$$

or in momentum space

$$-\mathscr{R}_{2\mu\nu}(h) = (q^2\,\delta_{\mu\sigma}\,\delta_{\nu\tau} - q_\nu q_\tau\,\delta_{\mu\sigma} - q_\mu q_\sigma\,\delta_{\nu\tau} + \delta_{\mu\nu}q_\sigma q_\tau)\bar{h}_{\sigma\tau}. \tag{11a}$$

This operation is singular and is without inverse because identically

$$\mathscr{R}_{2\mu\nu,\nu}(h) = 0$$

or

$$q_\nu\mathscr{R}_{2\mu\nu} = 0.$$

The first variation with respect to $g_{\mu\nu}$ of L_{matter} from (1a) we shall call $\mathscr{T}^{\mu\nu}$. The full equations for the gravitational field with matter present becomes then

$$\mathscr{R}^{\mu\nu}(g) = \mathscr{T}^{\mu\nu}(g, \phi). \tag{14a}$$

The equations for the motion of the field ϕ obtained from variation of L_{matter} with respect to ϕ, (namely, $g^{\mu\nu}\phi_{;\mu\nu} + m^2\phi = 0$) imply that

$$\mathscr{T}^{\mu\nu}_{\ \ ;\nu} = 0, \tag{13a}$$

a condition without which (14a) could not make sense in view of the identity (8a).

To make a quantum theory, we would expect at first to write $\exp i\int L\,d\tau$ and integrate over all $\mathscr{D}g\mathscr{D}\phi$ as a path integral in the usual way (the integral over $g_{\mu\nu}$ at a point is $g^{-5/2}$ times the product of the ten $dg_{\mu\nu}$). But this leads to the difficulties described for the *YM* case in the text. One can easily convert it into rules for diagrams in various orders. One way is to substitute $g_{\mu\nu} = \delta_{\mu\nu} + h_{\mu\nu}$ and expand in powers of $h_{\mu\nu}$. The second order in h leads to the singular propagator (11a), so if the other terms are called sources we have

$$\mathscr{T}_{2\mu\nu}(h) = s_{\mu\nu}.$$

Other terms are represented as interactions between matter and gravity in the usual way. For example, matter is coupled in first order to $h_{\mu\nu}$ via a term $\bar{h}_{\sigma\tau}\phi_{,\tau}\phi_{,\sigma} - \frac{1}{2}m^2h_{\sigma\sigma}\phi^2$, which leads to a rule that matter, in going from momentum p_1 to p_2, is coupled to one graviton of tensor $h_{\mu\nu}$ by an amplitude

$$p_\mu^1 p_\nu^2 - \tfrac{1}{2}\,\delta_{\mu\nu}(p^1\cdot p^2 - m^2).$$

There are new coupling terms of every order; that is, of junctions of every number of lines. We have made many calculations in this way to study the nature of the divergences of this theory. Here, however, we do not discuss these, but only problems in formulating the theory which arises from the singular nature of the operator $\mathscr{R}_{2\mu\nu}$ of (11a).

A free graviton entering or leaving has a symmetrical tensor polarization $e_{\mu\nu}$ and a momentum q_μ satisfying $\mathscr{R}_{2\mu\nu}(h) = 0$ or (see equation (11a))

$$q^2\bar{e}_{\mu\nu} - q_\nu q_\tau\bar{e}_{\tau\mu} - q_\mu q_\tau\bar{e}_{\tau\nu} + \delta_{\mu\nu}q_\sigma q_\tau\bar{e}_{\sigma\tau} = 0. \tag{34a}$$

If it is physical $q^2 = 0$ and (34a) says it is also transverse,

$$q_\tau\bar{e}_{\tau\nu} = 0. \tag{35a}$$

On the other hand, (34a) has solutions even if $q^2 \neq 0$, namely a quantum partly polarized in the direction q_ν, the gradient of a vector ξ_ν:

$$e_{\mu\nu} = q_\mu\xi_\nu + q_\nu\xi_\mu, \tag{36a}$$

a form we shall call a pure gradient, and one which corresponds to a pure gauge transformation (compare equation 2a). Such a potential has no physical effect and always gives zero effect on running all diagrams containing it.

To deal now with the second order propagator, $\mathscr{R}_{2\mu\nu}(h)$, we can discuss a theory in which we modify the second order terms in the original Lagrangian. The modification we shall consider is to add $\bar{h}_{\mu\nu,\nu}\bar{h}_{\mu\sigma,\sigma}$ to the Lagrangian so that it becomes simply

$$L_{2GD} = \tfrac{1}{2}\bar{h}_{\mu\nu,\sigma}h_{\mu\nu,\sigma}. \tag{25a}$$

In all these equations by this method, we no longer have complete general relativistic covariance. We adopt the convention that *repeated indices* on the *same horizontal line are to be summed with the Minkowski metric* $\delta_{\mu\nu}$. Indices repeated one upper, one lower are summed in the conventional way.

The field produced from a source is now simply

$$h_{\mu\nu} = (1/q^2)\bar{s}_{\mu\nu}. \tag{28a}$$

Of course, if the source is divergenceless, (28a) implies

$$\bar{h}_{\mu\nu,\nu} = 0, \tag{22a}$$

so in that case our system is equivalent to the unmodified equation (14a). Further discussion exactly parallels the *YM* case.

It appears, to get the single closed loop diagrams to work out satisfactorily if we add $\bar{h}_{\mu\nu,\nu}\bar{h}_{\mu\sigma,\sigma}$ to the gravity Lagrangian so that we may use (28a) as a

propagator, we must add the effects of a vector particle, field vector η^σ going around a loop and having the Lagrangian density

$$2\eta^\sigma_{,\nu}\overline{\eta_{\sigma;\nu}} = \eta^\sigma_{,\nu}\eta^\sigma_{,\nu} + h_{\sigma\tau}\eta^\sigma_{,\nu}\eta^\tau_{,\nu} - \eta^\tau\eta^\sigma[h_{\nu\tau,\sigma} - h_{\sigma\nu,\tau}] \tag{29a}$$

where we have used the condition (22a). It is interesting that only terms of first order in $h_{\mu\nu}$ appear here.

The tree diagrams we dealt with in the following way. The Lagrangian is

$$L = \mathscr{R}(g) + \mathscr{T}(g, \phi)$$

and the equation of motion resulting from it is

$$\delta L/\delta g_{\mu\nu} = \mathscr{R}_{\mu\nu}(g) + \mathscr{T}_{\mu\nu}(g, \phi) = 0,$$

$$\delta L/\delta\phi = W(\phi, g) = -[\sqrt{g}\, g^{\mu\nu}\phi_{,\mu}]_{,\nu} + m^2\phi\sqrt{g} = 0.$$

(We shall henceforth write second derivatives like $(\phi_{,\mu})_{,\nu}$ as $\phi_{,\mu\nu}$, etc.) We think of $g_{\mu\nu}$ as $\delta_{\mu\nu} + h_{\mu\nu}$ and split off the second order operator from the rest, calling

$$\mathscr{R}^{\mu\nu}(g) = \mathscr{R}_2^{\mu\nu}(g) - K^{\mu\nu}(g), \tag{39a}$$

$$\mathscr{R}_2^{\mu\nu}(g) = \mathscr{R}_{2D}^{\mu\nu}(g) - \bar{g}_{\mu\sigma,\sigma\nu} - \bar{g}_{\nu\sigma,\sigma\mu} + \delta_{\mu\nu}\bar{g}_{\sigma\tau,\sigma\tau}. \tag{40a}$$

(We have written $\mathscr{R}_2^{\mu\nu}(h)$ as $\mathscr{R}_2^{\mu\nu}(g)$, for they are the same.) Note that $\mathscr{R}^{\mu\nu}_{2,\nu} = 0$. We write $W(\phi, g) = -\Box^2\phi + m^2\phi + s(\phi, g)$. Thus, $\mathscr{R}_{2D}^{-1}$ is just $1/\Box^2$ times the bar operation. We then construct trees by the iterative solution of

$$g_{\mu\nu} = \mathscr{R}_{2D}^{-1}\{\mathscr{T}^{\mu\nu}(g) + K^{\mu\nu}(g)\} + g^0_{\mu\nu}, \tag{41a}$$

$$\phi = (\Box^2 - m^2)^{-1}\{s(\phi, g)\} + \phi^0, \tag{43a}$$

where $g^0_{\mu\nu}$, ϕ^0 represent appropriate asymptotic waves.

$$\mathscr{R}_2^{\mu\nu}(g^0) = \mathscr{R}_{2D}^{\mu\nu}(g^0) + \bar{g}^0_{\mu\sigma,\sigma\nu} + \bar{g}^0_{\nu\sigma,\sigma\mu} - \delta_{\mu\nu}\bar{g}^0_{\sigma\tau,\sigma\tau} = 0. \tag{44a}$$

If one of the asymptotic gravitons $E_{\mu\nu}$ is left out of the asymptotic waves $g^0_{\mu\nu}$ the final tree amplitude is

$$\delta t = \int E_{\mu\nu}(\mathscr{T}^{\mu\nu}(g, \phi) + K^{\mu\nu}(g, \phi))\, d\tau. \tag{45a}$$

Multiplying (41a) by $\mathscr{R}_{2D}$ shows that $g_{\mu\nu}$ satisfies

$$\mathscr{R}^{\mu\nu}(g) = \mathscr{T}^{\mu\nu}(g) - 2\bar{C}_{\nu,\mu} + \mathscr{R}_2^{\mu\nu}(g_0) \tag{46a}$$

where the last term is zero by (44a), and we have put

$$C_\nu = \bar{g}_{\nu\sigma,\sigma} - \bar{g}^0_{\nu\sigma,\sigma}. \tag{47a}$$

Taking the covariant derivative of both sides of (46a) shows us that

$$2(\bar{C}_{\nu,\mu})_{;\nu} = 0 = C_{\mu,\nu\nu} + 2\left\{\begin{matrix}\mu\\ \sigma\ \ \tau\end{matrix}\right\}\bar{C}_{\sigma,\tau}. \tag{48a}$$

We are solving this via (49a) as we show as follows: Taking the bar and then the divergence of both sides of (41a), replacing $K^{\mu\nu}$ via (39a), we can show

$$C_\nu = \Box^{-2}\{-\mathscr{R}^{\mu\nu}_{,\nu} + \mathscr{T}^{\mu\nu}_{,\nu} + \mathscr{R}^{\mu\nu}_{2,\nu}\}.$$

The last term vanishes, and the others may be re-expressed (since their covariant derivative vanishes) as $-\left\{\begin{matrix}\mu\\ \sigma\ \ \tau\end{matrix}\right\}(\mathscr{T}^{\sigma\tau} - \mathscr{R}^{\sigma\tau})$ which by (46a) leads us to

$$C_\nu = -\Box^{-2}\{2\left\{\begin{matrix}\mu\\ \sigma\ \ \tau\end{matrix}\right\}\bar{C}_{\sigma,\tau}\}, \tag{49a}$$

so C_ν is zero everywhere as it is zero asymptotically.

To show that the tree result does not depend on $g^0_{\mu\nu}$ if it is changed by a first order gauge transformation, $g^0_{\mu\nu} \to g^0_{\mu\nu} + \xi_{\nu,\mu} + \xi_{\mu,\nu}$, we show that a pure gradient for the last line $E_{\mu\nu} = \xi_{\nu,\mu} + \xi_{\mu,\nu}$ has no effect in (45a). Integration by parts gives

$$\delta t = -2\int \xi_\mu(\mathscr{T}^{\mu\nu} + K^{\mu\nu})_{,\nu}\, d\tau, \tag{50a}$$

which can be re-expressed as

$$\delta t = 4\int \xi_\mu\left\{\begin{matrix}\mu\\ \sigma\ \ \tau\end{matrix}\right\}\bar{C}_{\sigma,\tau}\, d\tau = 0 \tag{51a}$$

since C_σ vanishes.

When $g^0_{\mu\nu}$ is changed by $\xi_{\mu,\nu} + \xi_{\nu,\mu}$, $g_{\mu\nu}$ is changed by a gauge transformation (2a) or $g_{\mu\nu} \to g_{\mu\nu} + \chi_{\mu,\nu} + \chi_{\nu,\mu} - 2\left\{\begin{matrix}\alpha\\ \mu\ \ \nu\end{matrix}\right\}\chi_\alpha$ with $\eta^\mu = g^{\mu\alpha}\chi_\alpha$. The condition $C_\mu = 0$, or $\bar{g}_{\mu\nu,\nu} = \bar{g}^0_{\mu\nu,\nu}$ means that the χ is related to the ξ via $\chi_{\mu,\nu\nu} = \xi_{\mu,\nu\nu} + 2\left(\left\{\begin{matrix}\alpha\\ \mu\ \ \nu\end{matrix}\right\}\chi_\alpha\right)_{,\nu}$ which is solved by iteration of

$$\chi_\mu = \Box^{-2}\left\{2\left(\left\{\begin{matrix}\alpha\\ \mu\ \ \nu\end{matrix}\right\}\chi_\alpha\right)_{,\nu}\right\} + \xi_\mu, \tag{52a}$$

an equation adjoint to (60a).

For single closed loops, the $\mathscr{R}_{2D}$ theory does not agree with expectations when the loop is opened just as for the analogous YM case. To deal with single closed loops by closing a tree, we shall have to close a tree with an external gradient $(\xi_{\nu,\mu} + \xi_{\mu,\nu})$ and another external graviton, say $b_{\mu\nu}$, where

$$\bar{b}_{\mu\nu,\nu} = \tfrac{1}{2}\xi_\mu. \tag{55a}$$

Now, if $g^0_{\mu\nu}$ contains all the physical external lines and also $b_{\mu\nu}$, we have (note $q^2 = 0$, so $\mathscr{R}^{\mu\nu}_{2D}(b) = 0$)

$$\mathscr{R}^{\mu\nu}_2(g^0) = \mathscr{R}^{\mu\nu}_2(b) = -\xi_{\mu,\nu}. \tag{56a}$$

to put into (46a). We next define a new C_ν by adding a term $\frac{1}{2}\xi_\nu$ to (47a). Now our tree is (51a) but C_μ solves by iteration the equation

$$C_\mu = \Box^{-2}\{-2\left\{\begin{matrix}\mu\\ \sigma\ \ \tau\end{matrix}\right\}\bar{C}_{\sigma,\tau}\} + \tfrac{1}{2}\xi_\mu. \tag{60a}$$

If we choose $\bar{g}_{\mu\nu,\nu} = 0$, then $\left\{ \begin{matrix} \mu \\ \sigma \quad \sigma \end{matrix} \right\} = 0$ and the bar can be left off of $\bar{C}_{\sigma,\tau}$.

It appears that this tree would be the tree coming from the closed loop diagram of a field η^μ coming from an action

$$\int \left\{ g_{\sigma\tau}\eta^\sigma_{,\nu}\eta^\tau_{,\nu} + \eta^\tau\eta^\sigma_{,\nu}[g_{\nu\sigma,\tau} - g_{\nu\tau,\sigma}] \right\} d\tau \qquad \text{(see 29a)}$$

where we explicitly assume that $\bar{g}_{\mu\nu,\nu} = 0$ and where the σ, τ are summed covariantly but the index ν is summed by Minkowski's metric. The first variation of the action with respect to η^μ is

$$-g_{\mu\tau}\left(\eta^\tau_{,\nu\nu} + 2\left\{ \begin{matrix} \tau \\ \sigma \quad \nu \end{matrix} \right\}\eta^\sigma_{,\nu}\right)$$

as required to generate equation (60a) with $C_\mu = \eta^\mu$. The Lagrangian of this action can also be written

$$L_\eta = 2\eta^\tau_{,\nu}\bar{\eta}_{\tau;\nu}. \qquad (29a)$$

We now give the equations analogous to those in the section "Formal Theory of Simple Loops." We imagine the tree fields as external fields $g_{\mu\nu}$, and work out the contribution of one extra closed loop. That is, (see 62), we replace $g_{\mu\nu}$ by $g_{\mu\nu} + H_{\mu\nu}$ in the Lagrangian and expand to second order in $H_{\mu\nu}$ (we omit the matter field here). The second order Lagrangian which results is

$$L''_G = \int \sqrt{g}\, d\tau(-\tilde{H}^{\mu\nu}_{;\nu}\tilde{H}^\sigma_{\mu;\sigma} + \tfrac{1}{2}g^{\sigma\tau}\tilde{H}^{\mu\nu}_{;\sigma}H_{\mu\nu;\tau} - \tilde{H}^{\mu\nu}H_{\mu\tau}\tilde{R}^\tau_\nu + \tilde{H}^{\mu\nu}H^{\sigma\tau}R_{\mu\tau\nu\sigma}). \qquad (63a)$$

Here, $g_{\mu\nu}$ is a fixed external field, $R_{\mu\tau\nu\sigma}$ is the curvature tensor of the field $g_{\mu\nu}$ and we define for any tensor $X_{\mu\nu}$,

$$\tilde{X}_{\mu\nu} = \tfrac{1}{2}(X_{\mu\nu} + X_{\nu\mu} - g_{\mu\nu}X^\sigma_\sigma)$$

and all raising and lowering of indices, and Christoffel symbols in the definition of the covariant derivative, is with the metric tensor $g_{\mu\nu}$. This is the Lagrangian of a tensor field in curved space. For example, it would be the proper Lagrangian for discussing weak gravity wave propagation in an otherwise strong given field; that is to say, it would describe weak gravitational effects in a curved space.

Now we should like to use the trick of adding a mass term to the original gravity Lagrangian and we need the analogue L_{GM} of (19). There is no obvious and unique way to add the term in the gravitational case, but we have found the answer (corresponding to our particular way of modifying the gravity propagator) by the supposition that the term must be of such a form that the field condition

$$\bar{g}_{\mu\nu,\nu} = 0 \qquad (22a)$$

should be a consequence of the equations of motion. This will occur if we take

$$L_{GM} = L_G - M^2 g_{\sigma\sigma} \qquad (19a)$$

(so $-M^2\int g_{\sigma\sigma}\,d\tau$ is added to the action), for then the equations of motion become

$$\mathscr{R}^{\mu\nu}(g) + M^2\,\delta_{\mu\nu} = \mathscr{T}^{\mu\nu}(g). \tag{20a}$$

So that taking the covariant divergence of both sides, we find

$$\delta_{\mu\nu,\nu} + \left\{ \begin{matrix} \mu \\ \sigma\ \ \tau \end{matrix} \right\} \delta_{\sigma\tau} = 0,$$

which implies (22a) since $\left\{ \begin{matrix} \mu \\ \sigma\ \ \sigma \end{matrix} \right\} = g^{\mu\lambda}(\bar{g}_{\lambda\sigma,\sigma})$. This mass term is first order in $g_{\mu\nu}$ and does not affect the second order Lagrangian so that $L''_{GM} = L''_G$. This is not analogous to the *YM* case, but for uniformity of notation we continue to call it L''_{GM}, although it equals L''_G.

The theory L''_{GMD} will add a divergence term to L''_{GM} so that the propagator is nonsingular,

$$L''_{GMD} = L''_G + \int d\tau(g^{\alpha\beta}\bar{H}_{\alpha\nu,\nu}\bar{H}_{\beta\sigma,\sigma}).$$

The closed loop contributions are taken as path integrals over these Lagrangians over the fields $H_{\mu\nu}$ with zero asymptotic value (vacuum to vacuum). The loop with the mass term will differ from that with the cancelled divergence terms by a factor equal to the action of a loop of a vector particle. We shall prove

$$CLGM = CLGMD/CLVM \tag{67a}$$

where *CLVM* is the closed loop of a vector particle of mass M^2 of Lagrangian

$$L_{VM} = \int d\tau[2\eta^{\sigma}_{,\nu}\bar{\eta}_{\sigma;\nu} - M^2\eta^{\sigma}\eta_{\sigma}]. \tag{68a}$$

The path integral on $\mathscr{D}\eta$ means $\sqrt{g}$ times the product of the four differentials $d\eta^{\mu}$ at each space-time point.

The right side of (67a) has a limit as $M^2 \to 0$. We take that limit, leaving out the mass terms in the Lagrangian, to define the closed loop theory for gravitation. Namely, find loops via the Lagrangian (63a), omitting the first term (the divergence term). Then correct it by subtracting the action of a massless vector particle with Lagrangian density $\eta^{\sigma}{}_{,\nu}\bar{\eta}_{\sigma;\nu}$.

(If matter were present, the extra term in the Lagrangian in second order in $H_{\mu\nu}$ is

$$\tfrac{1}{2}\int (\phi_{,\sigma}\phi_{,}[\tilde{H}^{\sigma\alpha}H^{\tau}_{\alpha} - \tfrac{1}{4}g^{\sigma\tau}\tilde{H}^{\mu\nu}H_{\mu\nu}] + \tfrac{1}{4}m^2\phi^2\tilde{H}^{\mu\nu}H_{\mu\nu})\sqrt{g}\,d\tau$$

where ϕ is the external matter field. To deal with the matter loops we need to substitute $\phi = \phi + X$ and expand to second order in X. The Lagrangian second order in X is $\frac{1}{2}\int \sqrt{g}\,(g^{\mu\nu}X_{,\mu}X_{,\nu} - m^2X^2)\,d\tau$ and the term linear in X and H is $\int (-\tilde{H}^{\mu\nu}\phi_{,\mu}X_{,\nu} - (m^2/2)H^{\sigma}_{\sigma}\phi X]\sqrt{g}\,d\tau$.)

Under an infinitesimal coordinate transformation, (2a), $g_{\mu\nu} + H^{\nu}_{\mu\nu}$ becomes

$g_{\mu\nu} + H_{\mu\nu} + \eta^{\sigma}_{,\mu} g_{\sigma\nu} + \eta^{\sigma}_{,\nu} g_{\sigma\mu} + \eta^{\sigma} g_{\mu\nu,\sigma}$ to first order in H and η. This can be taken as a change of H and of X,

$$H'_{\mu\nu} = H_{\mu\nu} + \eta_{\mu;\nu} + \eta_{\nu;\mu} \tag{69a}$$

and

$$X' = X + \eta^{\sigma}\phi_{,\sigma}.$$

Now if we make this substitution in the Lagrangian, we can show after a number of steps analogous to (70a) that the entire action is changed by

$$-2M^2\eta^{\mu}_{,\sigma}(\bar{H}_{\sigma\mu} + \bar{\eta}_{\sigma;\mu}). \tag{71a}$$

Since this is a rather complicated calculation and the result is not directly obvious by analogy to *YM*, we outline a way it can be carried out abstractly (this method would, of course, work just as well for the *YM* case). Inclusion of matter terms only complicates the algebra and does not alter the result, so we omit matter. The Lagrangian density for the field will be written $L_G - M^2 g_{\sigma\sigma}$ to include the mass term. Putting $g_{\mu\nu} \to g_{\mu\nu} + H_{\mu\nu}$ changes this by

$$H_{\mu\nu}\,\delta L_G/\delta g_{\mu\nu} - M^2 H_{\sigma\sigma}$$

in first order and by

$$L''_{GM} = \tfrac{1}{2} H_{\mu\nu} H_{\sigma\tau}\,\delta^2 L_G/\delta g_{\mu\nu}\,\delta g_{\sigma\tau}$$

in second order. The first order vanishes if we assume that the external field $g_{\mu\nu}$ satisfies the gravitational equations of motion with mass, (20a) with $\mathscr{T}^{\mu\nu} = 0$, $\delta L_G/\delta g_{\mu\nu} = -\mathscr{R}^{\mu\nu}(g) = +M^2\,\delta_{\mu\nu}$. The second order is the Lagrangian we want. It is a property of L_G that gauge transformations do not alter it for any g, so

$$(\delta L_G/\delta g_{\mu\nu})_{;\nu} = (\delta L_G/\delta g_{\mu\nu})_{,\nu} + \left\{ \begin{matrix} \mu \\ \rho \quad \sigma \end{matrix} \right\}_g \delta L_G/\delta g_{\rho\sigma} = 0$$

identically for any $g_{\mu\nu}$. If into this equation we substitute $g_{\mu\nu} \to g_{\mu\nu} + H_{\mu\nu}$ and carry to first order in $H_{\mu\nu}$, we find

$$\begin{aligned}((\delta^2 L/\delta g_{\mu\nu}\,\delta g_{\sigma\tau})H_{\sigma\tau})_{;\nu} = -\tfrac{1}{2} g^{\mu\nu}\Big(&H_{\nu\rho,\sigma} + H_{\nu\sigma,\rho} - H_{\rho\sigma,\nu} \\ &- 2\left\{ \begin{matrix} \lambda \\ \sigma \quad \rho \end{matrix} \right\} H_{\lambda\nu}\Big)\,\delta L/\delta g_{\sigma\rho}\end{aligned} \tag{90}$$

for any $H_{\mu\nu}$. In our case, here $\delta L/\delta g_{\sigma\rho} = M^2\,\delta_{\sigma\rho}$ so (90) becomes

$$((\delta^2 L/\delta g_{\mu\nu}\,\delta g_{\sigma\tau})H_{\sigma\tau})_{;\nu} = -M^2 g^{\mu\nu}\bar{H}_{\nu\sigma,\sigma}. \tag{91}$$

Next the change in L''_{GM} produced by replacing $H'_{\mu\nu}$ by $H_{\mu\nu} + \eta_{\mu;\nu} + \eta_{\nu;\mu}$ (which is what we wish to calculate) is

$$2H_{\sigma\tau}\eta_{\mu;\nu}\,\delta^2 L/\delta g_{\mu\nu}\,\delta g_{\sigma\tau} + 2\eta_{\mu;\nu}\eta_{\sigma;\tau}\,\delta^2 L/\delta g_{\mu\nu}\,\delta g_{\sigma\tau}.$$

This, by parts, is $-2\eta_{\mu}(\delta^2 L/\delta g_{\mu\nu}\,\delta g_{\sigma\tau}(H_{\sigma\tau} + \eta_{\sigma;\tau}))_{;\nu}$ which by the analysis of the last paragraph is $2\eta^{\mu}M^2(\bar{H}_{\sigma\mu} + \bar{\eta}_{\sigma;\mu})_{,\sigma}$ which by parts proves our contention (71a).

With this observation (71a), we can proceed to prove (67a) in a way exactly analogous to the *YM* case. First, multiply the path integral for *CLGM* (with H' for H) by the path integral (dependent on $g_{\mu\nu}$)

$$X = \int \exp\left\{ i \int [M^2\eta'_\mu + 2(\bar{\eta}'_{\sigma;\mu})_{,\sigma}]^2 \, d\tau \right\} \mathscr{D}\eta' \tag{72a}$$

where by $(\xi_\mu)^2$ we mean $g^{\mu\nu}\xi_\mu\xi_\nu$. Define the linear operation G,

$$G\eta'_\mu = 2(\bar{\eta}'_{\sigma;\mu})_{,\sigma} + M^2\eta'_\mu \tag{74a}$$

and substitute η'_μ by

$$\eta'_\mu = \eta_\mu - G^{-1}\bar{H}'_{\sigma\mu,\sigma} \tag{75a}$$

so that the path integral now has for the exponent the integral of

$$L''_{GM}(H') + [M^2\eta_\mu + 2(\bar{\eta}_{\sigma;\mu})_{,\sigma} - \bar{H}'_{\sigma\mu,\sigma}]^2.$$

We now replace $H'_{\sigma\mu}$ by $H_{\sigma\mu} + \eta_{\sigma;\mu} + \eta_{\mu;\sigma}$ and use our rule (71a) to find the change in $L''_{GM}(H')$ to reduce this to

$$L''_{GM}(H) - 2M^2\eta^\mu_{,\sigma}(\bar{H}_{\sigma\mu} + \bar{\eta}_{\sigma;\mu}) + [M^2\eta_\mu - \bar{H}_{\sigma\mu,\sigma}]g^{\mu\nu}[M^2\eta_\nu - \bar{H}_{\sigma\nu,\sigma}]$$

which is (convert one term by parts)

$$L''_{GM}(H) + g^{\mu\nu}\bar{H}_{\mu\sigma,\sigma}\bar{H}_{\nu\sigma,\sigma} + M^2(M^2\eta^\mu\eta_\mu - 2\eta^\mu_{,\sigma}\bar{\eta}_{\sigma;\mu}). \qquad \text{(see 76a)}$$

The fields $H_{\mu\nu}$ and η_μ are now separated. The first two terms combine together to form the Lagrangian without the divergence, the one proper to *CLGDM*. The last terms redefining a new η^μ as M times the old gives just the path integral of the Lagrangian $\int (\eta^\mu G\eta_\mu)\sqrt{g}\, d\tau$, or $(\text{Det } G)^{-1/2}$, whereas the quantity of X of (72a) was $\int (\eta^\mu GG\eta_\mu)\sqrt{g}\, d\tau$ or $(\text{Det } G^2)^{-1/2} = (\text{Det } G)^{-1}$. Therefore, we must subtract from the action of one closed loop with the divergence term left out (*CLGMD*) the action of a vector particle with Lagrangian density

$$L_{VM} = -2\eta^\mu_{,\sigma}\bar{\eta}_{\sigma;\mu} + M^2\eta^\mu\eta_\mu. \tag{61a}$$

Since the limit of each term may be taken as $M^2 \to 0$ we can omit these M^2 terms, simply putting $M^2 = 0$, and thus have a rule to make calculations for one closed loop with massless gravity.

This formulation of gravitation is not unique. One could have started with some other nonlinear function of $g_{\mu\nu}$ (like $g^{\mu\nu}$ or $\sqrt{g}\,g^{\mu\nu}$) and expanded that as $\delta_{\mu\nu} + k_{\mu\nu}$ to begin the theory. We would presumably be led to other forms of conditions on $g_{\mu\nu}$, such as $(\sqrt{g}\,g^{\mu\nu})_{,\nu} = 0$ instead of $\bar{g}_{\mu\nu,\nu} = 0$, other divergence terms to define L''_{GD} or $\mathscr{R}_{2D}$ and other vector Lagrangians to subtract. Presumably they would all give equivalent results. The method used here is very awkward with its mixture of Minkowski and covariant indices. I have not investigated other possibilities to see if a more obviously covariant formulation could be made.

There is, of course, one possible covariant definition, the analogue of (66′), (68′), but I do not believe the result is entirely satisfactory (see discussion in text). What we do here is this. Instead of adding a mass term to the Lagrangian L_G from the beginning, we suppose the external field from trees is analyzed in any covariant way and that (in the absence of matter) $\mathscr{R}_{\mu\nu}(g) = 0$. Then the loop field $H_{\mu\nu}$ is considered as an extra tensor meson in a curved space, but this time a tensor field with mass, to form the Lagrangian

$$L''_{GM''} = L''_G - \frac{M^2}{2}\int \sqrt{g}\, d\tau \tilde{H}^{\mu\nu}H_{\mu\nu}$$

by adding the obvious and covariant (in g) mass term. The technique explained above can be directly applied in an obvious way. For example, the change in $L''_{GM}(H')$ and replacing H' by H comes only from the mass terms, since (90) is 0 as $\delta L/\delta g_{\mu\nu} = 0$, and is (all indices covariant, raised by $g^{\mu\nu}$, etc.)

$$-2M^2\bar{\eta}^{\mu;\nu}(H_{\mu\nu} + \eta_{\mu;\nu})\sqrt{g}.$$

The rest of the analysis proceeds by adding

$$\int \sqrt{g}\, g^{\mu\nu}(G\eta')_\mu(G\eta')_\lambda\, d\tau$$

where now

$$(G\eta)_\mu = M^2\eta_\mu + 2\tilde{\eta}_{\sigma;\mu;\tau}g^{\tau\sigma}.$$

The result is that

$$L''_{GMD'} = L''_{GM} + \int \sqrt{g}\, d\tau(\tilde{H}^{\mu\nu}_{;\nu}\tilde{H}_{\mu\sigma;\tau}g^{\sigma\tau}).$$

(thus cancelling the first term of 63a) and the vector particle to be subtracted has the Lagrangian

$$L_{VM'} = -\int \sqrt{g}\, d\tau[2\eta^\mu_{;\sigma}\tilde{\eta}_{\mu;\tau}g^{\sigma\tau} - M^2\eta^\mu\eta_\mu]. \qquad \text{(see 68′a)}$$

This is evidently covariant, has a limit as $M^2 \to 0$, but I believe the limit may not satisfactorily agree with the tree opening theorem.

The analogy to *YM* is very close except the mass term (in 19a) is quite peculiar, and it does not affect the loop Lagrangian. This causes one to question whether it might not be easier in gravitation than in the *YM* theory to deal with closed loops of arbitrary complexity. I did not find the solution to the problem of how to deal with such loops in either theory. Formula (19a) was not known to me then as I found it only when preparing this manuscript for publication, for some of my older formulas contained errors that I had not noticed.

Note added in proof: L. D. Faddeev and Bryce DeWitt have kindly prepared for me the following list of references to related work.

I. *Yang-Mills and Gravitation*
1. B. S. DeWitt, *Phys. Rev.*, **162**, 1195, 1239 (1967).
2. B. S. DeWitt, *Phys. Rev. Letters*, **12**, 742 (1964).

3. V. N. Popov and L. D. Faddeev, "Perturbation Theory for Gauge Invariant Fields," preprint I.T.Ph., USSR, Kiev (1967).
4. S. Mandelstam, *Phys. Rev.*, **175**, 1580, 1604 (1968).
5. S. Mandelstam, *Ann. Phys.* (New York), **19**, 25 (1962).

II. *Yang-Mills only*
1. L. D. Faddeev and V. N. Popov, *Phys. Letters*, **25B**, 29 (1967).
2. L. D. Faddeev, *Theo. and Math. Phys.*, **1**, 1 (1969).

III. *Canonical approach (Hamiltonian)*
1. P. A. M. Dirac, *Can. J. Math.*, **2**, 129 (1950).
2. P. A. M. Dirac, *Proc. Roy. Soc.* (London), **A246**, 326, 333 (1958).
3. P. A. M. Dirac, *Phys. Rev.*, **114**, 924 (1959).
4. R. Arnowitt, S. Deser and C. W. Misner, "The Dynamics of General Relativity," *Gravitation, an Introduction to Current Research*, L. Witten, ed. (New York: Wiley, 1962).
5. B. S. DeWitt, *Phys. Rev.*, **160**, 1113 (1967).
6. J. Schwinger, *Phys. Rev.*, **152**, 1219 (1966); **158**, 1391 (1967); **173**, 1264 (1968).

IV. *Canonical Derivations of the Feynman-DeWitt Rules*
1. I. B. Khryslovich, *Yadernaya Fizika*, **10**, 409 (1968); English translation: *Sov. J. Nucl. Phys.*, **10**, 235 (1970).
2. A. M. Altukhov and I. B. Khryslovich, *Yadernaya Fizika*, **11**, 902 (1970).
3. E. S. Fradkin and I. V. Tyutin, *Phys. Letters*, **30B**, 562 (1969); *Phys. Rev.*, **D2**, 2841 (1970).

VII. Computer Theory

Feynman's interest in numerical computation dated back to his wartime Los Alamos days, when Bethe put him in charge of a group doing calculations to model the plutonium implosion bomb. Feynman developed a system, using persons at mechanical calculators in a system analogous to what would much later, with digital computers, be called "parallel computing." During the last ten years of his life he became fascinated by the theory and application of computers and he gave a joint course in computation at Caltech, together with his colleagues John Hopfield and Carver Mead, also using guest lecturers.[1] Feynman also published three papers on computers. (Papers [110] and [115] are the same.)

The most important of the three papers is probably [106], in which the idea of a quantum computer is suggested and in which the limitations on computers imposed by the laws of physics are discussed. In some of his ideas, Feynman had been paralleled (and sometimes anticipated) independently by others, especially Paul Benioff and Rolf Landauer. Paper [110] discusses reversible computers (first considered by C.H. Bennet) and limitations that could arise from the second law of thermodynamics (the increase of entropy with time). However, the paper concludes with the sentence "At any rate, it seems that the laws of physics present no barrier to reducing the size of computers until bits are the size of atoms, and quantum behavior holds dominant sway."

Paper [113] is a lecture delivered in Japan in 1985 as a memorial to Yoichiro Tomonaga, who shared the 1965 Nobel Prize in Physics with Feynman and Julian Schwinger. It is a "popular" presentation discussing parallel computation and the possibilities of reducing energy consumption and size.[2]

Selected Papers

[106] Simulating physics with computers. *Int. J. Theor. Phys.* **B2** (1982): 467–488.

[110] Quantum mechanical computers. *Opt. News* **11** (1985): 11–46.

[113] The computing machines in the future. Nishina Memorial Lecture (1985), Nishina Foundation and Gakushuin.

[115] Quantum mechanical computers. *Foundations of Physics* **16** (1986): 507–531.

[1] See item [124] for the *Feynman Lectures on Computation*. See also *Feynman and Computation*, edited by Anthony J.G. Hey (Reading, Massachusetts: 1999).

[2] On the last topic, see also paper [44], "There's plenty of room at the bottom." "[This] paper is often credited with starting the field of nanotechnology." *Feynman and Computation*, p. xii (note 1).

International Journal of Theoretical Physics, Vol. 21, Nos. 6/7, 1982

Simulating Physics with Computers

Richard P. Feynman

Department of Physics, California Institute of Technology, Pasadena, California 91107

Received May 7, 1981

1. INTRODUCTION

On the program it says this is a keynote speech—and I don't know what a keynote speech is. I do not intend in any way to suggest what should be in this meeting as a keynote of the subjects or anything like that. I have my own things to say and to talk about and there's no implication that anybody needs to talk about the same thing or anything like it. So what I want to talk about is what Mike Dertouzos suggested that nobody would talk about. I want to talk about the problem of simulating physics with computers and I mean that in a specific way which I am going to explain. The reason for doing this is something that I learned about from Ed Fredkin, and my entire interest in the subject has been inspired by him. It has to do with learning something about the possibilities of computers, and also something about possibilities in physics. If we suppose that we know all the physical laws perfectly, of course we don't have to pay any attention to computers. It's interesting anyway to entertain oneself with the idea that we've got something to learn about physical laws; and if I take a relaxed view here (after all I'm here and not at home) I'll admit that we don't understand everything.

The first question is, What kind of computer are we going to use to simulate physics? Computer theory has been developed to a point where it realizes that it doesn't make any difference; when you get to a *universal computer*, it doesn't matter how it's manufactured, how it's actually made. Therefore my question is, Can physics be simulated by a universal computer? I would like to have the elements of this computer *locally interconnected*, and therefore sort of think about cellular automata as an example (but I don't want to force it). But I do want something involved with the

0020-7748/82/0600-0467$03.00/0

locality of interaction. I would not like to think of a very enormous computer with arbitrary interconnections throughout the entire thing.

Now, what kind of physics are we going to imitate? First, I am going to describe the possibility of simulating physics in the classical approximation, a thing which is usually described by local differential equations. But the physical world is quantum mechanical, and therefore the proper problem is the simulation of quantum physics—which is what I really want to talk about, but I'll come to that later. So what kind of simulation do I mean? There is, of course, a kind of approximate simulation in which you design numerical algorithms for differential equations, and then use the computer to compute these algorithms and get an approximate view of what physics ought to do. That's an interesting subject, but is not what I want to talk about. I want to talk about the possibility that there is to be an *exact* simulation, that the computer will do *exactly* the same as nature. If this is to be proved and the type of computer is as I've already explained, then it's going to be necessary that *everything* that happens in a finite volume of space and time would have to be exactly analyzable with a finite number of logical operations. The present theory of physics is not that way, apparently. It allows space to go down into infinitesimal distances, wavelengths to get infinitely great, terms to be summed in infinite order, and so forth; and therefore, if this proposition is right, physical law is wrong.

So good, we already have a suggestion of how we might modify physical law, and that is the kind of reason why I like to study this sort of problem. To take an example, we might change the idea that space is continuous to the idea that space perhaps is a simple lattice and everything is discrete (so that we can put it into a finite number of digits) and that time jumps discontinuously. Now let's see what kind of a physical world it would be or what kind of problem of computation we would have. For example, the first difficulty that would come out is that the speed of light would depend slightly on the direction, and there might be other anisotropies in the physics that we could detect experimentally. They might be very small anisotropies. Physical knowledge is of course always incomplete, and you can always say we'll try to design something which beats experiment at the present time, but which predicts anistropies on some scale to be found later. That's fine. That would be good physics if you could predict something consistent with all the known facts and suggest some new fact that we didn't explain, but I have no specific examples. So I'm not objecting to the fact that it's anistropic in principle, it's a question of how anistropic. If you tell me it's so-and-so anistropic, I'll tell you about the experiment with the lithium atom which shows that the anistropy is less than that much, and that this here theory of yours is impossible.

Another thing that had been suggested early was that natural laws are reversible, but that computer rules are not. But this turned out to be false; the computer rules can be reversible, and it has been a very, very useful thing to notice and to discover that. (Editors' note: see papers by Bennett, Fredkin, and Toffoli, these Proceedings). This is a place where the relationship of physics and computation has turned itself the other way and told us something about the possibilities of computation. So this is an interesting subject because it tells us something about computer rules, and *might* tell us something about physics.

The rule of simulation that I would like to have is that the number of computer elements required to simulate a large physical system is only to be proportional to the space-time volume of the physical system. I don't want to have an explosion. That is, if you say I want to explain this much physics, I can do it exactly and I need a certain-sized computer. If doubling the volume of space and time means I'll need an *exponentially* larger computer, I consider that against the rules (I make up the rules, I'm allowed to do that). Let's start with a few interesting questions.

2. SIMULATING TIME

First I'd like to talk about simulating time. We're going to assume it's discrete. You know that we don't have infinite accuracy in physical measurements so time might be discrete on a scale of less than 10^{-27} sec. (You'd have to have it at least like to this to avoid clashes with experiment—but make it 10^{-41} sec. if you like, and then you've got us!)

One way in which we simulate time—in cellular automata, for example—is to say that "the computer goes from state to state." But really, that's using intuition that involves the idea of time—you're going from state to state. And therefore the time (by the way, like the space in the case of cellular automata) is not simulated at all, it's imitated in the computer.

An interesting question comes up: "Is there a way of simulating it, rather than imitating it?" Well, there's a way of looking at the world that is called the space-time view, imagining that the points of space and time are all laid out, so to speak, ahead of time. And then we could say that a "computer" rule (now computer would be in quotes, because it's not the standard kind of computer which operates in time) is: We have a state s_i at each point i in space-time. (See Figure 1.) The state s_i at the space time point i is a given function $F_i(s_j, s_k, \ldots)$ of the state at the points j, k in some neighborhood of i:

$$s_i = F_i(s_j, , s_k, \ldots)$$

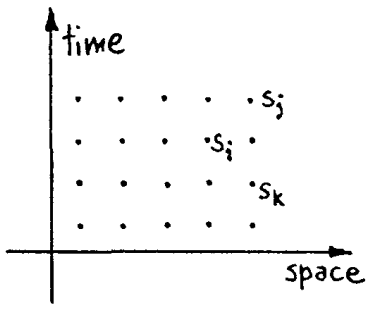


Fig. 1.

You'll notice immediately that if this particular function is such that the value of the function at i only involves the few points behind in time, earlier than this time i, all I've done is to redescribe the cellular automaton, because it means that you calculate a given point from points at earlier times, and I can compute the next one and so on, and I can go through this in that particular order. But just let's us think of a more general kind of computer, because we might have a more general function. So let's think about whether we could have a wider case of generality of interconnections of points in space-time. If F depends on *all* the points both in the future and the past, what then? That could be the way physics works. I'll mention how our theories go at the moment. It has turned out in many physical theories that the mathematical equations are quite a bit simplified by imagining such a thing—by imagining positrons as electrons going backwards in time, and other things that connect objects forward and backward. The important question would be, if this computer were laid out, is there in fact an organized algorithm by which a solution could be laid out, that is, computed? Suppose you know this function F_i and it is a function of the variables in the future as well. How would you lay out numbers so that they automatically satisfy the above equation? It may not be possible. In the case of the cellular automaton it is, because from a given row you get the next row and then the next row, and there's an organized way of doing it. It's an interesting question whether there are circumstances where you get functions for which you can't think, at least right away, of an organized way of laying it out. Maybe sort of shake it down from some approximation, or something, but it's an interesting different type of computation.

Question: "Doesn't this reduce to the ordinary boundary value, as opposed to initial-value type of calculation?"

Answer: "Yes, but remember this is the computer itself that I'm describing."

It appears actually that classical physics is causal. You can, in terms of the information in the past, if you include both momentum and position, or

the position at two different times in the past (either way, you need two pieces of information at each point) calculate the future in principle. So classical physics is *local*, *causal*, and *reversible*, and therefore apparently quite adaptable (except for the discreteness and so on, which I already mentioned) to computer simulation. We have no difficulty, in principle, apparently, with that.

3. SIMULATING PROBABILITY

Turning to quantum mechanics, we know immediately that here we get only the ability, apparently, to predict probabilities. Might I say immediately, so that you know where I really intend to go, that we always have had (secret, secret, close the doors!) we always have had a great deal of difficulty in understanding the world view that quantum mechanics represents. At least I do, because I'm an old enough man that I haven't got to the point that this stuff is obvious to me. Okay, I still get nervous with it. And therefore, some of the younger students ... you know how it always is, every new idea, it takes a generation or two until it becomes obvious that there's no real problem. It has not yet become obvious to me that there's no real problem. I cannot define the real problem, therefore I suspect there's no real problem, but I'm note sure there's no real problem. So that's why I like to investigate things. Can I learn anything from asking this question about computers—about this may or may not be mystery as to what the world view of quantum mechanics is? So I know that quantum mechanics seem to involve probability—and I therefore want to talk about simulating probability.

Well, one way that we could have a computer that simulates a probabilistic theory, something that has a probability in it, would be to calculate the probability and then interpret this number to represent nature. For example, let's suppose that a particle has a probability $P(x, t)$ to be at x at a time t. A typical example of such a probability might satisfy a differential equation, as, for example, if the particle is diffusing:

$$\frac{\partial P(x,t)}{\partial t} = -\nabla^2 P(x,t)$$

Now we could discretize t and x and perhaps even the probability itself and solve this differential equation like we solve any old field equation, and make an algorithm for it, making it exact by discretization. First there'd be a problem about discretizing probability. If you are only going to take k digits it would mean that when the probability is less that 2^{-k} of something happening, you say it doesn't happen at all. In practice we do that. If the

probability of something is 10^{-700}, we say it isn't going to happen, and we're not caught out very often. So we could allow ourselves to do that. But the real difficulty is this: If we had many particles, we have R particles, for example, in a system, then we would have to describe the probability of a circumstance by giving the probability to find these particles at points $x_1, x_2, \ldots, x_R$ at the time t. That would be a description of the probability of the system. And therefore, you'd need a k-digit number for every configuration of the system, for every arrangement of the R values of x. And therefore if there are N points in space, we'd need N^R configurations. Actually, from our point of view that at each point in space there is information like electric fields and so on, R will be of the same order as N if the number of information bits is the same as the number of points in space, and therefore you'd have to have something like N^N configurations to be described to get the probability out, and that's too big for our computer to hold if the size of the computer is of order N.

We emphasize, if a description of an isolated part of nature with N variables requires a general function of N variables and if a computer stimulates this by actually computing or storing this function then doubling the size of nature ($N \rightarrow 2N$) would require an exponentially explosive growth in the size of the simulating computer. It is therefore impossible, according to the rules stated, to simulate by calculating the probability.

Is there any other way? What kind of simulation can we have? We can't expect to compute the probability of configurations for a probabilistic theory. But the other way to simulate a probabilistic nature, which I'll call $\mathcal{N}$ for the moment, might still be to simulate the probabilistic nature by a computer $\mathcal{C}$ which itself is probabilistic, in which you always randomize the last two digit's of every number, or you do something terrible to it. So it becomes what I'll call a probabilistic computer, in which the output is not a unique function of the input. And then you try to work it out so that it simulates nature in this sense: that $\mathcal{C}$ goes from some state—initial state if you like—to some final state with the *same* probability that $\mathcal{N}$ goes from the corresponding initial state to the corresponding final state. Of course when you set up the machine and let nature do it, the imitator will not do the same thing, it only does it with the same probability. Is that no good? No it's O.K. How do you know what the probability is? You see, nature's unpredictable; how do you expect to predict it with a computer? You can't, —it's unpredictable if it's probabilistic. But what you really do in a probabilistic system is repeat the experiment in nature a large number of times. If you repeat the same experiment in the computer a large number of times (and that doesn't take any more time than it does to do the same thing in nature of course), it will give the frequency of a given final state proportional to the number of times, with approximately the same rate (plus

or minus the square root of n and all that) as it happens in nature. In other words, we could imagine and be perfectly happy, I think, with a probabilistic simulator of a probabilistic nature, in which the machine doesn't exactly do what nature does, but if you repeated a particular type of experiment a sufficient number of times to determine nature's probability, then you did the corresponding experiment on the computer, you'd get the corresponding probability with the corresponding accuracy (with the same kind of accuracy of statistics).

So let us now think about the characteristics of a local probabilistic computer, because I'll see if I can imitate nature with that (by "nature" I'm now going to mean quantum mechanics). One of the characteristics is that you can determine how it behaves in a local region by simply disregarding what it's doing in all other regions. For example, suppose there are variables in the system that describe the whole world (x_A, x_B)—the variables x_A you're interested in, they're "around here"; x_B are the whole result of the world. If you want to know the probability that something around here is happening, you would have to get that by integrating the total probability of all kinds of possibilities over x_B. If we had *computed* this probability, we would still have to do the integration

$$P_A(x_A) = \int P(x_A, x_B)\, dx_B$$

which is a hard job! But if we have *imitated* the probability, it's very simple to do it: you don't have to do anything to do the integration, you simply disregard what the values of x_B are, you just look at the region x_A. And therefore it does have the characteristic of nature: if it's local, you can find out what's happening in a region not by integrating or doing an extra operation, but merely by disregarding what happens elsewhere, which is no operation, nothing at all.

The other aspect that I want to emphasize is that the equations will have a form, no doubt, something like the following. Let each point $i = 1, 2, \ldots, N$ in space be in a state s_i chosen from a small state set (the size of this set should be reasonable, say, up to 2^5). And let the probability to find some configuration $\{s_i\}$ (a set of values of the state s_i at each point i) be some number $P(\{s_i\})$. It satisfies an equation such that at each jump in time

$$P_{t+1}(\{s\}) = \sum_{\{s'\}} \left[\prod_i m(s_i | s'_j, s'_k \ldots) \right] P_t(\{s'\})$$

where $m(s_i | s'_j, s'_k \ldots)$ is the probability that we move to state s_i at point i

when the neighbors have values $s'_j, s'_k \ldots$, where j, k etc. are points in the neighborhood of i. As j moves far from i, m becomes ever less sensitive to s'_j. At each change the state at a particular point i will move from what it was to a state s with a probability m that depends only upon the states of the neighborhood (which may be so defined as to include the point i itself). This gives the probability of making a transition. It's the same as in a cellular automaton; only, instead of its being definite, it's a probability. Tell me the environment, and I'll tell you the probability after a next moment of time that this point is at state s. And that's the way it's going to work, okay? So you get a mathematical equation of this kind of form.

Now I explicitly go to the question of how we can simulate with a computer—a universal automaton or something—the quantum-mechanical effects. (The usual formulation is that quantum mechanics has some sort of a differential equation for a function ψ.) If you have a single particle, ψ is a function of x and t, and this differential equation could be simulated just like my probabilistic equation was before. That would be all right and one has seen people make little computers which simulate the Schröedinger equation for a single particle. But the full description of quantum mechanics for a large system with R particles is given by a function $\psi(x_1, x_2, \ldots, x_R, t)$ which we call the amplitude to find the particles $x_1, \ldots, x_R$, and therefore, because it has too many variables, it *cannot be simulated* with a normal computer with a number of elements proportional to R or proportional to N. We had the same troubles with the probability in classical physics. And therefore, the problem is, how can we simulate the quantum mechanics? There are two ways that we can go about it. We can give up on our rule about what the computer was, we can say: Let the computer itself be built of quantum mechanical elements which obey quantum mechanical laws. Or we can turn the other way and say: Let the computer still be the same kind that we thought of before—a logical, universal automaton; can we imitate this situation? And I'm going to separate my talk here, for it branches into two parts.

4. QUANTUM COMPUTERS—UNIVERSAL QUANTUM SIMULATORS

The first branch, one you might call a side-remark, is, Can you do it with a new kind of computer—a quantum computer? (I'll come back to the other branch in a moment.) Now it turns out, as far as I can tell, that you can simulate this with a quantum system, with quantum computer elements. It's not a Turing machine, but a machine of a different kind. If we disregard the continuity of space and make it discrete, and so on, as an approximation (the same way as we allowed ourselves in the classical case), it does seem to

be true that all the various field theories have the same *kind* of behavior, and can be simulated in every way, apparently, with little latticeworks of spins and other things. It's been noted time and time again that the phenomena of field theory (if the world is made in a discrete lattice) are well imitated by many phenomena in solid state theory (which is simply the analysis of a latticework of crystal atoms, and in the case of the kind of solid state I mean each atom is just a point which has numbers associated with it, with quantum-mechanical rules). For example, the spin waves in a spin lattice imitating Bose-particles in the field theory. I therefore believe it's true that with a suitable class of quantum machines you could imitate any quantum system, including the physical world. But I don't know whether the general theory of this intersimulation of quantum systems has ever been worked out, and so I present that as another interesting problem: to work out the classes of different kinds of quantum mechanical systems which are really intersimulatable—which are equivalent—as has been done in the case of classical computers. It has been found that there is a kind of universal computer that can do anything, and it doesn't make much difference specifically how it's designed. The same way we should try to find out what kinds of quantum mechanical systems are mutually intersimulatable, and try to find a specific class, or a character of that class which will simulate everything. What, in other words, is the universal quantum simulator? (assuming this discretization of space and time). If you had discrete quantum systems, what other discrete quantum systems are exact imitators of it, and is there a class against which everything can be matched? I believe it's rather simple to answer that question and to find the class, but I just haven't done it.

Suppose that we try the following guess: that every finite quantum mechanical system can be described *exactly*, imitated exactly, by supposing that we have another system such that at each point in space-time this system has only two possible base states. Either that point is occupied, or unoccupied—those are the two states. The mathematics of the quantum mechanical operators associated with that point would be very simple.

$$a = \text{ANNIHILATE} = \begin{array}{c|cc} & \text{OCC} & \text{UN} \\ \hline \text{OCC} & 0 & 0 \\ \text{UN} & 1 & 0 \end{array} = \tfrac{1}{2}(\sigma_x - i\sigma_y)$$

$$a^* = \text{CREATE} = \begin{array}{|cc} 0 & 1 \\ 0 & 0 \end{array} = \tfrac{1}{2}(\sigma_x + i\sigma_y)$$

$$n = \text{NUMBER} = \begin{array}{|cc} 1 & 0 \\ 0 & 0 \end{array} = a^*a = \tfrac{1}{2}(1 + \sigma_z)$$

$$\mathbb{1} = \text{IDENTITY} = \begin{array}{|cc} 1 & 0 \\ 0 & 1 \end{array}$$

There would be an operator a which *annihilates* if the point is occupied —it changes it to unoccupied. There is a conjugate operator a^* which does the opposite: if it's unoccupied, it occupies it. There's another operator n called the *number* to ask, Is something there? The little matrices tell you what they do. If it's there, n gets a one and leaves it alone, if it's not there, nothing happens. That's mathematically equivalent to the product of the other two, as a matter of fact. And then there's the identity, $\mathbb{1}$, which we always have to put in there to complete our mathematics—it doesn't do a damn thing!

By the way, on the right-hand side of the above formulas the same operators are written in terms of matrices that most physicists find more convenient, because they are Hermitian, and that seems to make it easier for them. They have invented another set of matrices, the Pauli σ matrices:

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

And these are called *spin*—spin one-half—so sometimes people say you're talking about a spin-one-half lattice.

The question is, if we wrote a Hamiltonian which involved only these operators, locally coupled to corresponding operators on the other space-time points, could we imitate every quantum mechanical system which is discrete and has a finite number of degrees of freedom? I know, almost certainly, that we could do that for any quantum mechanical system which involves Bose particles. I'm not sure whether Fermi particles could be described by such a system. So I leave that open. Well, that's an example of what I meant by a general quantum mechanical simulator. I'm not sure that it's sufficient, because I'm not sure that it takes care of Fermi particles.

5. CAN QUANTUM SYSTEMS BE PROBABILISTICALLY SIMULATED BY A CLASSICAL COMPUTER?

Now the next question that I would like to bring up is, of course, the interesting one, i.e., Can a quantum system be probabilistically simulated by a classical (probabilistic, I'd assume) universal computer? In other words, a computer which will give the same probabilities as the quantum system does. If you take the computer to be the classical kind I've described so far, (not the quantum kind described in the last section) and there're no changes in any laws, and there's no hocus-pocus, the answer is certainly, No! This is called the hidden-variable problem: it is impossible to represent the results of quantum mechanics with a classical universal device. To learn a little bit about it, I say let us try to put the quantum equations in a form as close as

possible to classical equations so that we can see what the difficulty is and what happens. Well, first of all we can't simulate ψ in the normal way. As I've explained already, there're too many variables. Our only hope is that we're going to simulate probabilities, that we're going to have our computer do things with the same probability as we observe in nature, as calculated by the quantum mechanical system. Can you make a cellular automaton, or something, imitate with the same probability what nature does, where I'm going to suppose that quantum mechanics is correct, or at least after I discretize space and time it's correct, and see if I can do it. I must point out that you must directly generate the probabilities, the results, with the correct quantum probability. Directly, because we have no way to store all the numbers, we have to just imitate the phenomenon directly.

It turns out then that another thing, rather than the wave function, a thing called the *density matrix*, is much more useful for this. It's not so useful as far as the mathematical equations are concerned, since it's more complicated than the equations for ψ, but I'm not going to worry about mathematical complications, or which is the easiest way to calculate, because with computers we don't have to be so careful to do it the very easiest way. And so with a slight increase in the complexity of the equations (and not very much increase) I turn to the density matrix, which for a single particle of coordinate x in a pure state of wave function $\psi(x)$ is

$$\rho(x, x') = \psi^*(x)\psi(x')$$

This has a special property that is a function of two coordinates x, x'. The presence of two quantities x and x' associated with each coordinate is analogous to the fact that in classical mechanics you have to have two variables to describe the state, x and $\dot{x}$. States are described by a second-order device, with two informations ("position" and "velocity"). So we have to have two pieces of information associated with a particle, analogous to the classical situation, in order to describe configurations. (I've written the density matrix for one particle, but of course there's the analogous thing for R particles, a function of $2R$ variables).

This quantity has many of the mathematical properties of a probability. For example if a state $\psi(x)$ is not certain but is ψ_α with the probability p_α then the density matrix is the appropriate weighted sum of the matrix for each state α:

$$\rho(x, x') = \sum_\alpha p_\alpha \psi_\alpha^*(x)\psi a(x').$$

A quantity which has properties even more similar to classical probabilities is the Wigner function, a simple reexpression of the density matrix; for a

single particle

$$W(x, p) = \int \rho\left(x + \frac{y}{2}, x - \frac{y}{2}\right) e^{ipy} dy$$

We shall be emphasizing their similarity and shall call it "probability" in quotes instead of Wigner function. Watch these quotes carefully, when they are absent we mean the real probability. If "probability" had all the mathematical properties of a probability we could remove the quotes and simulate it. $W(x, p)$ is the "probability" that the particle has position x and momentum p (per dx and dp). What properties does it have that are analogous to an ordinary probability?

It has the property that if there are many variables and you want to know the "probabilities" associated with a finite region, you simply disregard the other variables (by integration). Furthermore the probability of finding a particle at x is $\int W(x, p)dp$. If you can interpret W as a probability of finding x and p, this would be an expected equation. Likewise the probability of p would be expected to be $\int W(x, p)dx$. These two equations are correct, and therefore you would hope that maybe $W(x, p)$ is the probability of finding x and p. And the question then is can we make a device which simulates this W? Because then it would work fine.

Since the quantum systems I noted were best represented by spin one-half (occupied versus unoccupied or spin one-half is the same thing), I tried to do the same thing for spin one-half objects, and it's rather easy to do. Although before one object only had two states, occupied and unoccupied, the full description—in order to develop things as a function of time—requires twice as many variables, which mean two slots at each point which are occupied or unoccupied (denoted by $+$ and $-$ in what follows), analogous to the x and $\dot{x}$, or the x and p. So you can find four numbers, four "probabilities" $\{f_{++}, f_{+-}, f_{-+}, f_{--}\}$ which act just like, and I have to explain why they're not exactly like, but they act just like, probabilities to find things in the state in which both symbols are up, one's up and one's down, and so on. For example, the sum $f_{++} + f_{+-} + f_{-+} + f_{--}$ of the four "probabilities" is 1. You'll remember that one object now is going to have two indices, two plus/minus indices, or two ones and zeros at each point, although the quantum system had only one. For example, if you would like to know whether the first index is positive, the probability of that would be

$$\text{Prob(first index is } +) = f_{++} + f_{+-} \qquad [\text{spin } z \text{ up}]$$

i.e., you don't care about the second index. The probability that the first index is negative is

$$\text{Prob(first index is } -) = f_{-+} + f_{--}, \qquad [\text{spin } z \text{ down}]$$

These two formulas are exactly correct in quantum mechanics. You see I'm hedging on whether or not "probability" f can really be a probability without quotes. But when I write probability without quotes on the left-hand side I'm not hedging; that really is the quantum mechanical probability. It's interpreted perfectly fine here. Likewise the probability that the second index is positive can be obtained by finding

$$\text{Prob(second index is } +) = f_{++} + f_{-+} \qquad [\text{spin } x \text{ up}]$$

and likewise

$$\text{Prob(second index is } -) = f_{+-} + f_{--} \qquad [\text{spin } x \text{ down}]$$

You could also ask other questions about the system. You might like to know, What is the probability that both indices are positive? You'll get in trouble. But you could ask other questions that you won't get in trouble with, and that get correct physical answers. You can ask, for example, what is the probability that the two indices are the same? That would be

$$\text{Prob(match)} = f_{++} + f_{--} \qquad [\text{spin } y \text{ up}]$$

Or the probability that there's no match between the indices, that they're different,

$$\text{Prob(no match)} = f_{+-} + f_{-+} \qquad [\text{spin } y \text{ down}]$$

All perfectly all right. All these probabilities are correct and make sense, and have a precise meaning in the spin model, shown in the square brackets above. There are other "probability" combinations, other linear combinations of these f's which also make physically sensible probabilities, but I won't go into those now. There are other linear combinations that you can ask questions about, but you don't seem to be able to ask questions about an individual f.

6. NEGATIVE PROBABILITIES

Now, for many interacting spins on a lattice we can give a "probability" (the quotes remind us that there is still a question about whether it's a probability) for correlated possibilities:

$$F(s_1, s_2, \ldots, s_N) \qquad (s_i \in \{++, +-, -+, --\})$$

Next, if I look for the quantum mechanical equation which tells me what the changes of F are with time, they are exactly of the form that I wrote above for the classical theory:

$$F_{t+1}(\{s\}) = \sum_{\{s'\}} \left[\prod_i M(s_i | s'_j, s'_k \ldots) \right] F_t(\{s'\})$$

but now we have F instead of P. The $M(s_i|s'_j, s'_k \ldots)$ would appear to be interpreted as the "probability" per unit time, or per time jump, that the state at i turns into s_i when the neighbors are in configuration s'. If you can invent a probability M like that, you write the equations for it according to normal logic, those are the correct equations, the real, correct, quantum mechanical equations for this F, and therefore you'd say, Okay, so I can imitate it with a probabilistic computer!

There's only one thing wrong. These equations unfortunately cannot be so interpreted on the basis of the so-called "probability", or this probabilistic computer can't simulate them, because the F is not necessarily positive. Sometimes it's negative! The M, the "probability" (so-called) of moving from one condition to another is itself not positive; if I had gone all the way back to the f for a single object, it again is not necessarily positive.

An example of possibilities here are

$$f_{++} = 0.6 \qquad f_{+-} = -0.1 \qquad f_{-+} = 0.3 \qquad f_{--} = 0.2$$

The sum $f_{++} + f_{+-}$ is 0.5, that's 50% chance of finding the first index positive. The probability of finding the first index negative is the sum $f_{-+} + f_{--}$ which is also 50%. The probability of finding the second index positive is the sum $f_{++} + f_{-+}$ which is nine tenths, the probability of finding it negative is $f_{+-} + f_{--}$ which is one-tenth, perfectly alright, it's either plus or minus. The probability that they match is eight-tenths, the probability that they mismatch is plus two-tenths; every physical probability comes out positive. But the original f's are not positive, and therein lies the great difficulty. The only difference between a probabilistic classical world and the equations of the quantum world is that somehow or other it appears as if the probabilities would have to go negative, and that we do not know, as far as I know, how to simulate. Okay, that's the fundamental problem. I don't know the answer to it, but I wanted to explain that if I try my best to make the equations look as near as possible to what would be imitable by a classical probabilistic computer, I get into trouble.

7. POLARIZATION OF PHOTONS—TWO-STATES SYSTEMS

I would like to show you why such minus signs cannot be avoided, or at least that you have some sort of difficulty. You probably have all heard this example of the Einstein-Podolsky-Rosen paradox, but I will explain this little example of a physical experiment which can be done, and which has been done, which does give the answers quantum theory predicts, and the answers are really right, there's no mistake, if you do the experiment, it actually comes out. And I'm going to use the example of polarizations of photons, which is an example of a two-state system. When a photon comes, you can say it's either x polarized or y polarized. You can find that out by putting in a piece of calcite, and the photon goes through the calcite either out in one direction, or out in another—actually slightly separated, and then you put in some mirrors, that's not important. You get two beams, two places out, where the photon can go. (See Figure 2.)

If you put a polarized photon in, then it will go to one beam called the ordinary ray, or another, the extraordinary one. If you put detectors there you find that each photon that you put in, it either comes out in one or the other 100% of the time, and not half and half. You either find a photon in one or the other. The probability of finding it in the ordinary ray plus the probability of finding it in the extraordinary ray is always 1—you have to have that rule. That works. And further, it's never found at both detectors. (If you might have put two photons in, you could get that, but you cut the intensity down—it's a technical thing, you don't find them in both detectors.)

Now the next experiment: Separation into 4 polarized beams (see Figure 3). You put two calcites in a row so that their axes have a relative angle ϕ, I happen to have drawn the second calcite in two positions, but it doesn't make a difference if you use the same piece or not, as you care. Take the ordinary ray from one and put it through another piece of calcite and look at its ordinary ray, which I'll call the ordinary–ordinary ($O-O$) ray, or look at its extraordinary ray, I have the ordinary–extraordinary ($O-E$) ray. And then the extraordinary ray from the first one comes out as the $E-O$ ray, and then there's an $E-E$ ray, alright. Now you can ask what happens.

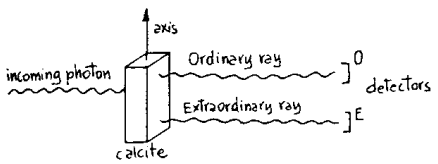


Fig. 2.

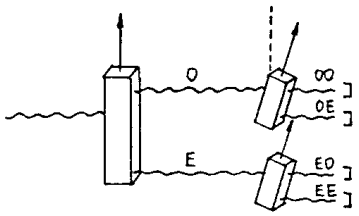


Fig. 3.

You'll find the following. *When a photon comes in, you always find that only one of the four counters goes off.*

If the photon is O from the first calcite, then the second calcite gives $O-O$ with probability $\cos^2\phi$ or $O-E$ with the complementary probability $1-\cos^1\phi=\sin^2\phi$. Likewise an E photon gives a $E-O$ with the probability $\sin^2\phi$ or an $E-E$ with the probability $\cos^2\phi$.

8. TWO-PHOTON CORRELATION EXPERIMENT

Let us turn now to the two photon correlation experiment (see Figure 4).

What can happen is that an atom emits two photons in opposite direction (e.g., the $3s\rightarrow 2p\rightarrow 1s$ transition in the H atom). They are observed simultaneously (say, by you and by me) through two calcites set at ϕ_1 and ϕ_2 to the vertical. Quantum theory and experiment agree that the probability P_{OO} that both of us detect an ordinary photon is

$$P_{OO}=\tfrac{1}{2}\cos^2(\phi_2-\phi_1)$$

The probability P_{EE} that we both observe an extraordinary ray is the same

$$P_{EE}=\tfrac{1}{2}\cos^2(\phi_2-\phi_1)$$

The probability P_{OE} that I find O and you find E is

$$P_{OE}=\tfrac{1}{2}\sin^2(\phi_2-\phi_1)$$

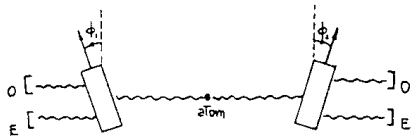


Fig. 4.

and finally the probability P_{EO} that I measure E and you measure O is

$$P_{EO} = \tfrac{1}{2}\sin^2(\phi_2 - \phi_1)$$

Notice that you can always predict, from your own measurement, what I shall get, O or E. For any axis ϕ_1 that I chose, just set your axis ϕ_2 to ϕ_1, then

$$P_{OE} = P_{EO} = 0$$

and I must get whatever you get.

Let us see now how it would have to be for a *local* probabilistic computer. Photon 1 must be in some condition α with the probability $f_\alpha(\phi_1)$, that determines it to go through as an ordinary ray [the probability it would pass as E is $1 - f_\alpha(\phi_1)$]. Likewise photon 2 will be in a condition β with probability $g_\beta(\phi_2)$. If $p_{\alpha\beta}$ is the conjoint probability to find the condition pair α, β, the probability P_{OO} that both of us observe O rays is

$$P_{OO}(\phi_1, \phi_2) = \sum_{\alpha\beta} p_{\alpha\beta} f_\alpha(\phi_1) g_\beta(\phi_2) \qquad \sum_{\alpha\beta} p_{\alpha\beta} = 1$$

likewise

$$P_{OE}(\phi_1, \phi_2) = \sum_{\alpha\beta} p_{\alpha\beta}(1 - f_\alpha(\phi_1)) g_\beta(\phi_2) \quad \text{etc.}$$

The conditions α determine how the photons go. There's some kind of correlation of the conditions. Such a formula cannot reproduce the quantum results above for any $p_{\alpha\beta}$, $f_\alpha(\phi_1)$, $g_\beta(\phi_2)$ if they are real probabilities—that is all positive, although it is easy if they are "probabilities"—negative for some conditions or angles. We now analyze why that is so.

I don't know what kinds of conditions they are, but for any condition the probability $f_\alpha(\phi)$ of its being extraordinary or ordinary in any direction must be either one or zero. Otherwise you couldn't predict it on the other side. You would be unable to predict with certainty what I was going to get, unless, every time the photon comes here, which way it's going to go is absolutely determined. Therefore, whatever condition the photon is in, there is some hidden inside variable that's going to determine whether it's going to be ordinary or extraordinary. This determination is done deterministically, not probabilistically; otherwise we can't explain the fact that you could predict what I was going to get *exactly*. So let us suppose that something like this happens. Suppose we discuss results just for angles which are multiples of 30°.

On each diagram (Figure 5) are the angles 0°, 30°, 60°, 90°, 120°, and 150°. A particle comes out to me, and it's in some sort of state, so what it's going to give for 0°, for 30°, etc. are all predicted—determined—by the state. Let us say that in a particular state that is set up the prediction for 0° is that it'll be extraordinary (black dot), for 30° it's also extraordinary, for 60° it's ordinary (white dot), and so on (Figure 5a). By the way, the outcomes are complements of each other at right angles, because, remember, it's always either extraordinary or ordinary; so if you turn 90°, what used to be an ordinary ray becomes the extraordinary ray. Therefore, whatever condition it's in, it has some predictive pattern in which you either have a prediction of ordinary or of extraordinary—three and three—because at right angles they're not the same color. Likewise the particle that comes to you when they're separated must have the same pattern because you can determine what I'm going to get by measuring yours. Whatever circumstances come out, the patterns must be the same. So, if I want to know, Am I going to get white at 60°? You just measure at 60°, and you'll find white, and therefore you'll predict white, or ordinary, for me. Now each time we do the experiment the pattern may not be the same. Every time we make a pair of photons, repeating this experiment again and again, it doesn't have to be the same as Figure 5a. Let's assume that the next time the experiment my photon will be O or E for each angle as in Figure 5c. Then your pattern looks like Figure 5d. But whatever it is, your pattern has to be my pattern exactly—otherwise you couldn't predict what I was going to get exactly by measuring the corresponding angle. And so on. Each time we do the experiment, we get different patterns; and it's easy: there are just six dots and three of them are white, and you chase them around different way—everything can happen. If we measure at the same angle, we always find that with this kind of arrangement we would get the same result.

Now suppose we measure at $\phi_2 - \phi_1 = 30°$, and ask, With what probability do we get the same result? Let's first try this example here (Figure 5a, 5b). With what probability would we get the same result, that they're

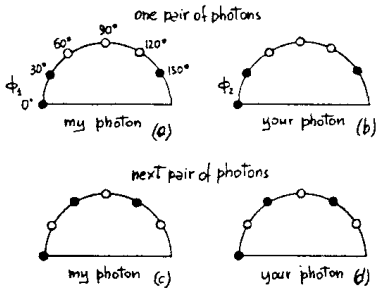


Fig. 5.

both white, or they're both black? The thing comes out like this: suppose I say, After they come out, I'm going to choose a direction at random, I tell you to measure 30° to the right of that direction. Then whatever I get, you would get something different if the neighbors were different. (We would get the same if the neighbors were the same.) What is the chance that you get the same result as me? The chance is the number of times that the neighbor is the same color. If you'll think a minute, you'll find that two thirds of the time, in the case of Figure 5a, it's the same color. The worst case would be black/white/black/white/black/white, and there the probability of a match would be zero (Figure 5c,d). If you look at all eight possible distinct cases, you'll find that the biggest possible answer is two-thirds. You cannot arrange, in a classical kind of method like this, that the probability of agreement at 30° will be bigger than two-thirds. But the quantum mechanical formula predicts $\cos^2 30°$ (or 3/4)—and experiments agree with this—and therein lies the difficulty.

That's all. That's the difficulty. That's why quantum mechanics can't seem to be imitable by a local classical computer.

I've entertained myself always by squeezing the difficulty of quantum mechanics into a smaller and smaller place, so as to get more and more worried about this particular item. It seems to be almost ridiculous that you can squeeze it to a numerical question that one thing is bigger than another. But there you are—it is bigger than any logical argument can produce, if you have this kind of logic. Now, we say "this kind of logic;" what other possibilities are there? Perhaps there may be no possibilities, but perhaps there are. Its interesting to try to discuss the possibilities. I mentioned something about the possibility of time—of things being affected not just by the past, but also by the future, and therefore that our probabilities are in some sense "illusory." We only have the information from the past, and we try to predict the next step, but in reality it depends upon the near future which we can't get at, or something like that. A very interesting question is the origin of the probabilities in quantum mechanics. Another way of puttings things is this: we have an illusion that we can do any experiment that we want. We all, however, come from the same universe, have evolved with it, and don't really have any "real" freedom. For we obey certain laws and have come from a certain past. Is it somehow that we are correlated to the experiments that we do, so that the apparent probabilities don't look like they ought to look if you assume that they are random. There are all kinds of questions like this, and what I'm trying to do is to get you people who think about computer-simulation possibilities to pay a great deal of attention to this, to digest as well as possible the real answers of quantum mechanics, and see if you can't invent a different point of view than the physicists have had to invent to describe this. In fact the physicists have no

good point of view. Somebody mumbled something about a many-world picture, and that many-world picture says that the wave function ψ is what's real, and damn the torpedos if there are so many variables, N^R. All these different worlds and every arrangement of configurations are all there just like our arrangement of configurations, we just happen to be sitting in this one. It's possible, but I'm not very happy with it.

So, I would like to see if there's some other way out, and I want to emphasize, or bring the question here, because the discovery of computers and the thinking about computers has turned out to be extremely useful in many branches of human reasoning. For instance, we never really understood how lousy our understanding of languages was, the theory of grammar and all that stuff, until we tried to make a computer which would be able to understand language. We tried to learn a great deal about psychology by trying to understand how computers work. There are interesting philosophical questions about reasoning, and relationship, observation, and measurement and so on, which computers have stimulated us to think about anew, with new types of thinking. And all I was doing was hoping that the computer-type of thinking would give us some new ideas, if any are really needed. I don't know, maybe physics is absolutely OK the way it is. The program that Fredkin is always pushing, about trying to find a computer simulation of physics, seem to me to be an excellent program to follow out. He and I have had wonderful, intense, and interminable arguments, and my argument is always that the real use of it would be with quantum mechanics, and therefore full attention and acceptance of the quantum mechanical phenomena—the challenge of explaining quantum mechanical phenomena —has to be put into the argument, and therefore these phenomena have to be understood very well in analyzing the situation. And I'm not happy with all the analyses that go with just the classical theory, because nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy. Thank you.

9. DISCUSSION

Question: Just to interpret, you spoke first of the probability of A given B, versus the probability of A and B jointly—that's the probability of one observer seeing the result, assigning a probability to the other; and then you brought up the paradox of the quantum mechanical result being 3/4, and this being 2/3. Are those really the same probabilities? Isn't one a joint probability, and the other a conditional one?

Answer: No, they are the same. P_{OO} is the *joint probability* that both you and I observe an ordinary ray, and P_{EE} is the *joint* probability for two

extraordinary rays. The probability that our observations match is

$$P_{OO} + P_{EE} = \cos^2 30° = 3/4$$

Question: Does it in some sense depend upon an assumption as to how much information is accessible from the photon, or from the particle? And second, to take your question of prediction, your comment about predicting, is in some sense reminiscent of the philosophical question, Is there any meaning to the question of whether there is free will or predestination? namely, the correlation between the observer and the experiment, and the question there is, Is it possible to construct a test in which the prediction could be reported to the observer, or instead, has the ability to represent information already been used up? And I suspect that you may have already used up all the information so that prediction lies outside the range of the theory.

Answer: All these things I don't understand; deep questions, profound questions. However physicists have a kind of a dopy way of avoiding all of these things. They simply say, now look, friend, you take a pair of counters and you put them on the side of your calcite and you count how many times you get this stuff, and it comes out 75% of the time. Then you go and you say, Now can I imitate that with a device which is going to produce the same results, and which will operate locally, and you try to invent some kind of way of doing that, and if you do it in the ordinary way of thinking, you find that you can't get there with the same probability. Therefore some new kind of thinking is necessary, but physicists, being kind of dull minded, only look at nature, and don't know how to think in these new ways.

Question: At the beginning of your talk, you talked about discretizing various things in order to go about doing a real computation of physics. And yet it seems to me that there are some differences between things like space and time, and probability that might exist at some place, or energy, or some field value. Do you see any reason to distinguish between quantization or discretizing of space and time, versus discretizing any of the specific parameters or values that might exist?

Answer: I would like to make a few comments. You said quantizing or discretizing. That's very dangerous. Quantum theory and quantizing is a very specific type of theory. Discretizing is the right word. Quantizing is a different kind of mathematics. If we talk about discretizing... of course I pointed out that we're going to have to change the laws of physics. Because the laws of physics as written now have, in the classical limit, a continuous variable everywhere, space and time. If, for example, in your theory you were going to have an electric field, then the electric field could not have (if it's going to be imitable, computable by a finite number of elements) an

infinite number of possible values, it'd have to be digitized. You might be able to get away with a theory by redescribing things without an electric field, but supposing for a moment that you've discovered that you can't do that and you want to describe it with an electric field, then you would have to say that, for example, when fields are smaller than a certain amount, they aren't there at all, or something. And those are very interesting problems, but unfortunately they're not good problems for classical physics because if you take the example of a star a hundred light years away, and it makes a wave which comes to us, and it gets weaker, and weaker, and weaker, and weaker, the electric field's going down, down, down, how low can we measure? You put a counter out there and you find "clunk," and nothing happens for a while, "clunk," and nothing happens for a while. It's not discretized at all, you never can measure such a tiny field, you don't find a tiny field, you don't have to imitate such a tiny field, because the world that you're trying to imitate, the physical world, is not the classical world, and it behaves differently. So the particular example of discretizing the electric field, is a problem which I would not see, as a physicist, as fundamentally difficult, because it will just mean that your field has gotten so small that I had better be using quantum mechanics anyway, and so you've got the wrong equations, and so you did the wrong problem! That's how I would answer that. Because you see, if you would imagine that the electric field is coming out of some 'ones' or something, the lowest you could get would be a full one, but that's what we see, you get a full photon. All these things suggest that it's really true, somehow, that the physical world is representable in a discretized way, because every time you get into a bind like this, you discover that the experiment does just what's necessary to escape the trouble that would come if the electric field went to zero, or you'd never be able to see a star beyond a certain distance, because the field would have gotten below the number of digits that your world can carry.

Nishina Memorial Lecture

The Computing Machines in the Future

Richard P. Feynman
California Institute of Technology
Pasadena, California, U.S.A.

August 9, 1985

Contents

The Computing Machines in the Future

Nishina Memorial Lecture
given at Gakushuin, Tokyo, August 9, 1985

Richard P. Feynman
California Institute of Technology, Pasadena, California, U.S.A.

It's a great pleasure and an honor to be here as a speaker in memorial for a scientist that I have respected and admired as much as Prof. Nishina. To come to Japan and talk about computers is like giving a sermon to Buddha. But I have been thinking about computers and this is the only subject I could think of when invited to talk.

The first thing I would like to say is what I am not going to talk about. I want to talk about the future computing machines. But the most important possible developments in the future, are things that I will not speak about. For example, there is a great deal of work to try to develop smarter machines, machines which have a better relationship with the humans so that input and output can be made with less effort than the complex programing that's necessary today. This goes under the name often of artificial intelligence, but I don't like that name. Perhaps the unintelligent machines can do even better than the intelligent ones. Another problem is the standardization of programing languages. There are too many languages today, and it would be a good idea to choose just one. (I hesitate to mention that in Japan, for what will happen will be that there will simply be more standard languages; you already have four ways of writing now and attempts to standardize anything here result apparently in more standards and not fewer.) Another interesting future problem that is worth working on but I will not talk about, is automatic debugging programs; debugging means to fix errors in a program or in a machine. It is surprisingly difficult to debug programs as they get more

complicated. Another direction of improvement is to make physical machines three dimensional instead of all on a surface of a chip. That can be done in stages instead of all at once; you can have several layers and then many more layers as the time goes on. Another important device would be a way of detecting automatically defective elements on a chip, then this chip itself automatically rewiring itself so as to avoid the defective elements. At the present time when we try to make big chips there are flaws, bad spots in the chips, and we throw the whole chip away. But of course if we could make it so that we could use the part of the chip that was effective, it would be much more efficient. I mention these things to try to tell you that I am aware of what the real problems are for future machines. But what I want to talk about is simple, just some small technical, physically good things that can be done in principle according to the physical laws; I would like in other words to discuss the machinery and not the way we use the machines.

I will talk about some technical possibilities for making machines. There will be three topics really. One is parallel processing machines which is something of the very near future, almost present, that is being developed now. Further in the future are questions of the energy consumption of machines which seems at the moment to be a limitation, but really isn't. Finally I will talk about the size; it is always better to make the machines smaller, and the question is how much smaller is it still possible to make machines according to the laws of Nature, in principle. I will not discuss which and what of these things will actually appear in the future. That depends on economic problems and social problems and I am not going to try to guess at those.

1. Parallel Computers

First about parallel programing, parallel computers, rather. Almost all the present computers, conventional computers, work on a layout or an architecture invented by von Neumann, in which there is a

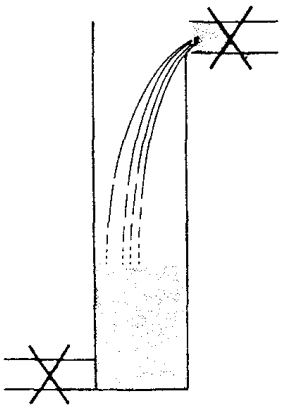
very large memory that stores all the information, and one central location that does simple calculations. We take a number from this place in the memory and a number from that place in the memory, send the two to the central arithmetical unit to add them and then send the answer to some other place in the memory. There is, therefore, effectively one central processor which is working very very fast and very hard while the whole memory sits out there like a vast filing cabinet of cards which are very rarely used. It is obvious that if there were more processors working at the same time we ought to be able to do calculations faster. But the problem is that some one who might be using one processor may be using some information from the memory that another one needs, and it gets very confusing. And so it has been said that it is very difficult to work many processors in parallel. Some steps in that direction have been taken in the larger conventional machines, what they call "vector processors". When sometimes you want to do exactly the same step on many different items you can do that perhaps at the same time. The ordinary hope is that the regular program can be written, and then an interpreter will discover automatically when it is useful to use this vector possibility. That idea is used in the Cray and in the super-computers in Japan. Another plan is to take what is effectively a large number of relatively simple (but not very simple) computers, and connect them all together in some pattern. Then they can all work on a part of the problem. Each one is really an independent computer, and they will transfer information to each other as one or another needs it. This kind of a scheme is in a machine for example called Cosmic Cube, and is one of the possibilities; many people are making such machines. Another plan is to distribute very large numbers of very simple central processors all over the memory. Each one deals with just a small part of the memory and there is an elaborate system of interconnections between them. An example of such a machine is the Connection Machine made at M.I.T. It has 64,000 processors and a system of routing in which every 16 can talk to any other 16 and thus 4000 routing connection possibilities. It would appear that scientific questions like the propagation of waves in some material

might be very easily handled by parallel processing, because what happens in this part of space at a moment can be worked out locally and only the pressures and the stresses from the neighbor needs to be known for each section can be worked out at the same time, and communicate boundary conditions across. That's why this type of design is built for such a thing. But it has turned out that very large number of problems of all kinds can be dealt with in parallel. As long as the problem is big enough so that a lot of calculating has to be done, it turns out that a parallel computation can speed this up enormously, not just scientific problems.

And what happened to the prejudice of 2 years ago, which was that the parallel programing is difficult? It turns out that what was difficult, and almost impossible, is to take an ordinary program and automatically figure out how to use the parallel computation effectively on that program. Instead, one must start all over again with the problem, appreciating that we have parallel possibility of calculation, and rewrite the program completely with a new attitude to what is inside the machine. It is not possible to effectively use the old programs. They must be rewritten. That is a great disadvantage to most industrial applications and has met with considerable resistance. But, the big programs belong usually to scientists or others, unofficial intelligent programmers who love computer science and are willing to start all over again and rewrite the program if they can make it more efficient. So what's going to happen is that the hard programs, vast big ones, will first be programed by experts in the new way, and then gradually everybody will have to come around, and more and more programs will be programed that way, and programmers will just have to learn how to do it.

2. Reducing the Energy Loss

The second topic I want to talk about is energy loss in computers. The fact that they must be cooled is the limitation apparently to the largest computers; a good deal of the effort is



Now

Fig. 1

spent in cooling the machine. I would like to explain that this is simply a result of very poor engineering and is nothing fundamental at all. Inside the computer a bit of information is controlled by a wire which either has a voltage of one value or another value. It is called "one bit", and we have to change the voltage of the wire from one value to the other and have to put charge on or take charge off. I make an analogy with water: We have to fill a vessel with water and get one level or to empty it to get to the other level. It's just an analogy. If you like electricity better you can think more accurately electrically. What we do now is analogous, in the water case, to filling the vessel by pouring water in from a top level (Fig.1), and lowering the level by opening the valve at the bottom and letting it all run out. In both cases there is a loss of energy because of the drop of the water, suddenly, through a height say from top level where it comes in to the low bottom level when you start pouring it in to fill it up again. In the cases of voltage and charge, there occurs the same thing.

It's like, as Mr. Bennett has explained, operating an automobile which has to start and stop by turning on the engine and putting on the brakes, turning on the engine and putting on the brakes; each time you lose power. Another way with a car would be to connect the wheels

6

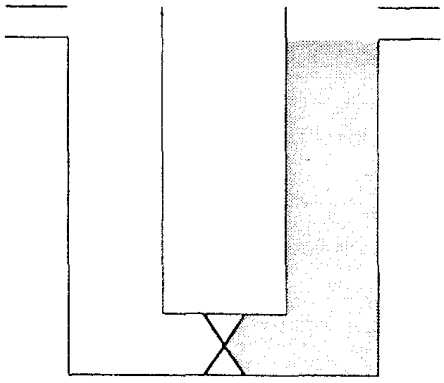


Fig. 2 Inertia (Inductance)

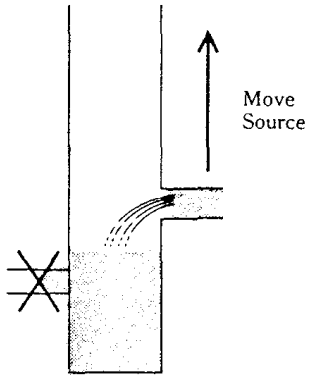
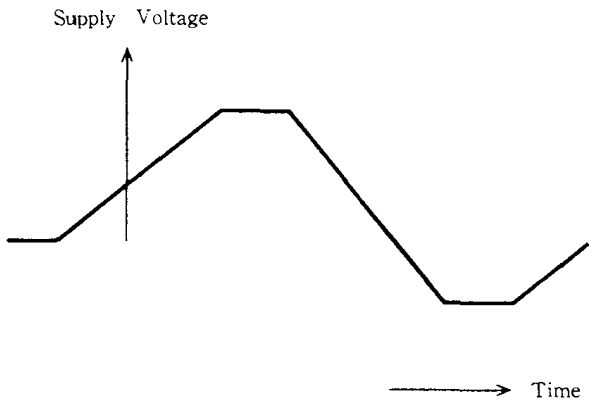


Fig. 3

to flywheels. Stop the car and speed up the flywheel saving the energy, which can then be reconnected to start the car again. The analogy electrically or in the water would be to have a U-shaped tube with a valve at the bottom in the center connecting the two arms of the U (Fig.2). When it is full here on the right but empty on the left with the valve closed, if we open that valve the water will slip out to the other side, and we close it just in time to catch it. Then when we want to go the other way we open the valve again and it slips to the other side and we catch it. There is some loss and it doesn't climb as high as it did before, but all we have to do is to put a little water in to correct the little loss, a much smaller energy loss than the direct fill method. But such a thing uses the inertia of the water and the analogue in the electricity, is inductance. However it is very difficult with the silicon transistors that we use today to make up inductance on the chips. So this is not particularly practical with the present technology.

Another way would be to fill the tank by a supply which stays only a little bit above the level lifting the water supply in time as we fill it up (Fig.3), because then the dropping of water is always small during the entire effort. In the same way, we could use an



Energy Loss · Time = Constant

Fig. 4

outlet to lower it but just taking off the top and lowering the tube, so that the heat loss would not appear at the position of the transistor, or would be small; it will depend on how high the distance is between the supply and the surface as we fill it up. This method corresponds to changing the voltage supply with time (Fig.4). So if we would use a time varying voltage supply, we could use this method. Of course, there is energy loss in the voltage supply, but that is all located in one place, that is simple, and there we can make one big inductance. This scheme is called "hot clocking", because the voltage supply operates also as the clock which times everything. And we don't need an extra clock signal to time the circuits as we do in conventional designs.

Both of these last two devices use less energy if they go slower. If I try to move the water supply level too fast, the water in the tube doesn't keep up with it and I have a big drop. So to work I must go slowly. Again, the U-tube scheme will not work unless that central valve can open and close faster than the time it takes for the water in the U-tube to slip back and forth. So my devices are slower. I've saved an energy loss but I've made the devices slower. The energy loss multiplied by the time it takes for the circuit to operate is

8

constant. But nevertheless, this turns out to be very practical because the clock time is usually much larger than the circuit time for the transistors, and we can use that to decrease the energy. Also if we went, let us say, three times slower with our calculations, we could use one third the energy over three times the time, which is nine times less power that has to be dissipated. Maybe it is worth it. Maybe by redesigning using parallel computations or other devices, we can spend a little longer than we could do at maximum circuit speed, in order to make a larger machine that is practical and from which we could still get the energy out.

For a transistor, the energy loss multiplied by the time it takes to operate is a product of several factors (Fig.5): (1) the thermal energy propotional to temperature, kT; (2) the length of the transistor between source and drain, divided by the velocity of the electrons inside (the thermal velocity $\sqrt{3kT/m}$); (3) the length of the transistor in units of the mean free path for collisions of electrons in the transistor; and finally (4) the total number of the electrons that are inside the transistor when it operates. All of these numbers come out to tell us that the energy used in the transistor today is somewhere around a billion or ten billions or more times the thermal energy kT. When it switches we use that much energy. It is very large amount of energy. It is obviously a good idea to decrease the size of the transistor. We decrease the length between source and drain, and we can decrease the number of the electrons, and use much

ENERGY · TIME FOR TRANSISTOR

$$= kT \cdot \frac{\text{LENGTH}}{\text{THERMAL VELOCITY}} \cdot \frac{\text{LENGTH}}{\text{MEAN FREE PATH}} \cdot \text{NUMBER OF ELECT. PAIRS}$$

ENERGY $\sim 10^{9-11}\, kT$.

$\therefore$ DECREASE SIZE. FASTER
LESS ENERGY

Fig. 5

less energy. It also turns out that a smaller transistor is much faster, because the electrons can cross it and make their decisions to switch faster. For every reason, it is a good idea to make the transistor smaller, and everybody is always trying to do that.

But suppose we come to a circumstance in which the mean free path is longer than the size of the transistor, then we discover that the transistor doesn't work right any more. It does not behave the way we expected. This reminds me, years ago there was something called the sound barrier. Airplanes cannot go faster than the speed of sound because, if you design them normally and try to put them in that speed, the propeller wouldn't work and the wings don't lift and nothing works correctly. Nevertheless, airplanes can go faster than the speed of sound. You just have to know what the right laws are under the right circumstances, and design the device with the correct laws. You cannot expect old designs to work in new circumstances. But new designs can work in new circumstances, and I assert that it is perfectly possible to make transistor systems, that is to say more correctly, switching systems, computing devices in which the dimensions are smaller than the mean free path. I speak of course in principle and I am not speaking about actual manufacture. Therefore, let us discuss what happens if we try to make the devices as small as possible.

3. Reducing the Size

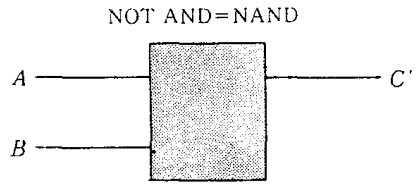
So, my third topic is the size of computing elements and now I speak entirely theoretically. The first thing that you would worry about when things get very small, is Brownian motion; everything is shaking and nothing stays in place, and how can you control the circuits then? And if the circuits did work, it has a chance of accidentally jumping back. But, if we use two volts for the energy of this electric system which is what we ordinarily use, that is eighty times the thermal energy (kT=1/40 volt) and the chance that something jumps backward against 80 times thermal energy is e, the base of the

10

natural logarithm, to minus eighty power, or 10^{-43}. What does that mean? If we had a billion transistors in a computer (which we don't have, we don't have that many at all), working all of them 10^{10} times a second, that is, tenth of a nanosecond switching perpetually, operating for 10^{9} seconds, which is 30 years, the total number of switching operations in that machine is 10^{28} and the chance of one of them going backward is only 10^{-43}; there will be no error produced by thermal oscillations whatsoever in 30 years. If you don't like that, use 2.5 volts and then it gets smaller. Long before that, the real failure will come when a cosmic ray accidentally goes through the transistor, and we don't have to be more perfect than that.

However, much more is in fact possible and I would like to refer you to an article in a most recent Scientific American by Bennett and Landauer. It is possible to make a computer in which each element, each transistor, can go forward and accidentally reverse and still the computer will operate. All the operation in succession in the computer go forward or backward. The computation proceeds for a while this way and then it undoes itself, uncalculates, and then goes forward again and so on. If we just pull it along a little, we can make it go through and finish the calculation by making it just a little bit more likely that it goes forward than backward.

It is known that all the computations can be made by putting together some simple elements like transistors; or, if we be more logically abstract, a thing for instance called NAND gate (NAND means NOT-AND). It has two "wires" in and one out (Fig.6). Forget the NOT first. What is AND? AND is: The output is 1 only if both input wires are 1, otherwise the output is 0. NOT-AND means the opposite. The output wire reads 1 (i.e. has the voltage level corresponding to 1) unless both input wires read 1, if both input wires read 1 then the output wire reads 0 (i.e. has the voltage level corresponding to 0). Here is a little table of inputs and outputs. A and B are inputs and C is the output. Unless A and B are both 1, the output is 1 otherwise 0. But such a device is irreversible. Information is lost. If I only know the output, I cannot recover the input. The device can't be expected to flip forward and then come back and compute



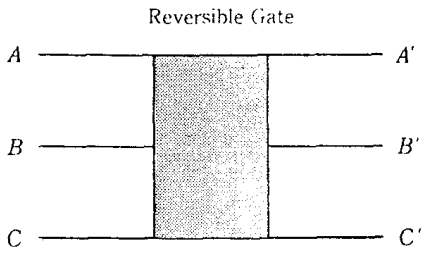
A	B	C'
0	0	1
0	1	1
1	0	1
1	1	0

Not Reversible
Information Lost

Fig. 6

correctly anymore. Because if we know for instance that the output is now 1, we don't know whether it came from A=0, B=1 or A=1, B=0 or A=0, B=0 and it cannot go back. Such a device is an irreversible gate. The great discovery of Bennett and, independently, of Fredkin is that it is possible to do computation with a diffferent kind of fundamental gate unit, a reversible gate unit. I have illustrated their idea -- with this unit which I could call a reversible NAND or whatever. It has three inputs and three outputs (Fig.7). Of the outputs, two, A' and B', are the same as two of the inputs, A and B, but the third input works this way: C' is the same as C unless A and B are both 1. Then it changes whatever C is. For instance, if C is 1 it is changed to 0, if C is 0 it is changed to 1 only if both A and B are 1. If you put two in succession, you see A and B will go through, and if C is not changed in both it stays the same or if C is changed twice it stays the same. So this gate reverses itself. No information has been lost. It is possible to discover what

12



$A' = A$

$B' = B$

$C' = C$, unless $A = 1$ and $B = 1$

$C' = 1 - C$

$=$ NOT C, if $A = 1$ and $B = 1$

No Information Lost

Some Force Needed to Push Calculation
Predominantly Forward
Energy Loss · Time Used = Constant

Fig. 7

went in if you know what went out.

A device made entirely with such gates will make calculations if everything moves forward, but if things go back and forth for a while and then eventually go forward enough it still operates correctly. If the things flip back and then go forward later it is still all right. It's very much the same as a particle in a gas which is bombarded by the atoms around it, usually goes nowhere, but with just a little pull, a little prejudice that makes a chance to move one way a little higher than the other way, the thing will slowly drift forward and reach from one end to the other, in spite of the Brownian motion that is made. So our computer will compute provided we apply a force of drift to pull the thing more likely across the calculation. Although it is not doing the calculation in a smooth way, but calculating like this, forward and backward, it eventually finishes the job. As with the particle in the gas, if we pull it very slightly, we lose very

little energy, but it takes a long time to get to one side from the other. If we are in a hurry, and we pull hard, then we loose a lot of energy. And the same with this computer. If we are patient and go slowly, we can make the computer operate with practically no energy, even less than kT per step, any amount as small as you like if you have enough time. But if you are in a hurry, you must dissipate energy, and again it's true that the energy lost to pull the calculation forward to complete it multiplied by the time you are allowed to make the calculation is a constant.

With these possibilities how small can we make a computer? How big must a number be? We all know we can write numbers in base 2 as strings of "bits" each a one or a zero. But how small can I write? Surely only one atom is needed to be in one state or another to determine if it represents a one or a zero. And the next atom could be a one or a zero, so a little string of atoms are enough to hold a number, one atom for each bit. (Actually since an atom can be in more states than just two we could use even fewer atoms, but enough is little enough!)

So now for intellectual entertainment we consider whether we could make a computer in which the bits writing is of atomic size, in which a bit is for example whether the spin in the atom is up for 1 or down for 0. And then our transistor changing the bits in different places would correspond to some interaction between some atoms, which will change their states. The simplest would be a kind of 3-atom interaction to be the fundamental element or gate in such a device. But again, it won't work right if we design it with the laws appropriate for large objects. We must use the new laws of physics, quantum mechanical laws, the laws that they are appropriate to atomic motion. And so we have to ask whether the principles of quantum mechanics permit an arrangements of atoms so small in number as a few times the number of gates in a computer that could still be put together and operate as a computer. This has been studied in principle, and such an arrangement has been found. The laws of quantum mechanics are reversible and therefore we must use the invention of reversible gates, that principle, that idea of Bennett and Fredkin,

but we know that's alright now. When the quantum mechanical situation is studied it is found that quantum mechanics adds no further limitations to anything that Mr. Bennet has said from thermodynamic considerations. Of course there is a limitation, the practical limitation anyway, that the bits must be of the size of an atom and a transistor 3 or 4 atoms; the quantum mechanical gate I used has 3 atoms. (I would not try to write my bits on to nuclei, I'll wait till the technological development reaches the atoms before I need to go any further!) That leads us just with (a) the limitations in size to the size of atoms, (b) the energy requirements depending on the time as worked out by Bennett, (c) and the feature that I did not mention concerning the speed of light; we can't send the signals any faster than the speed of light. Those are the only physical limitations that I know on computers.

If we make an atomic size computer, somehow, it would mean that the dimension, the <u>linear</u> dimension is a thousand to ten thousands times smaller than those very tiny chips that we have now. It means that the volume of the computer is 100 billionth, 10^{-11} of the present volume, because the transistor is that much smaller 10^{-11}, than the transistors that we make today. The energy requirement for a single switch is also about eleven orders of magnitude smaller than the energy required to switch the transistor today, and the time to make the transisions will be at least ten thousands times faster per step of calculation. So there is plenty of room for improvement in the computer and I leave you, practical people who work on computers, this as an aim to get to. I underestimated how long it would take for Mr. Ezawa to translate what I said, and I have no more to say that I have prepared for today. Thank you!

I will answer questions if you'd like.

Questions and Answers

Q: You mentioned that one bit of information can be stored in one atom, and I wonder if you can store the same amount of information in one quark.

A: Yes. But we don't have control of the quarks and that becomes a really impractical way to deal with things. You might think that what I am talking about is impractical, but I don't believe so. When I am talking about atoms, I believe that someday we will be able to handle and control them individually. There would be so much energy involved in the quark interactions it would be very dangerous to handle because of the radioactivity and so on. But the atomic energies that I am talking about are very familiar to us in chemical energies, electrical energies, and those, that I am speaking of, are numbers that are within the realm of reality, I believe, however absurd it may seem at the moment.

Q: You said that the smaller the computing element is the better. But, I think equipments have to be larger, because....

A: You mean that your finger is too big to push the buttons? Is that what you mean?

Q: Yes, it is.

A: Of course, you are right. I am talking about internal computers perhaps for robots or other devices. The input and output is something that I didn't discuss, whether the input comes from looking at pictures, hearing voices, or buttons being pushed. I am discussing how the computation is done in principle, and not what form the output should take. It is certainly true that the input and the output cannot be reduced in most cases effectively beyond human dimension. It is already too difficult to push the buttons on some of the computers with our big fingers. But with elaborate computing problems that take hours and hours, they could be done very rapidly on the very small machines with low energy consumption. That's the kind of

machine I was thinking of. Not the simple applications of adding two numbers but the elaborate calculations.

Q: I would like to know your method to transform the information from one atomic scale element to another atomic scale element. If you will use a quantum mechanical or natural interaction between the two elements then such a device will become very close to Nature itself. For example, if we make a computer simulation, a Monte Carlo simulation of a magnet to study critical phenomena, then your atomic scale computer will be very close to the magnet itself. What are your thoughts about that?

A: Yes. All things that we make are Nature. We arrange it in a way to suit our purpose, to make a calculation for a purpose. In a magnet there is some kind of relation, if you wish, there are some kind of computations going on just like there is in the solar system in a way of thinking. But, that might not be the calculation we want to make at the moment. What we need to make is a device for which we can change the programs and let it compute the problem that we want to solve, not just its own magnet problem that it likes to solve for itself. I can't use the solar system for a computer unless it just happens that the problem that someone gave me was to find the motion of the planets, in which case all I have to do is to watch.

There was an amusing article as a joke. Far in the future the "article" appears discussing a new method of making aerodynamical calculations: Instead of using the elaborate computers of the day, the author invents a simple device to blow air past the wing. (He reinvents the wind tunnel.)

Q: I have recently read in an newspaper article that operations of the nerve system in a brain are much slower than present day computers and the unit in the nerve system is much smaller. Do you think that the computers you have talked about today have something in common with the nerve system in the brain?

A: There is an analogy between the brain and the computer in that there are apparently elements that can switch under the control of

others. Nerve impulses controling or exciting other nerves, in a way that often depends upon whether more than one impulse comes in; something like an AND or its generalization. The amount of energy used in the brain cell for one of these transitions? I don't know the number. The time it takes to make a switching in the brain is very much longer than it is in our computers even today, never mind the fancy business of some atomic computer. But, interconnection system is much more elaborate. Each nerve is connected to thousand other nerves, whereas we connect transistors to two or three others.

Some people look at the activity of the brain in action and see that in many respects it surpasses the computer of today, and in many other respects the computer surpasses ourselves. This inspires people to design machines that can do more. What often happens is that an engineer makes up how the brain works in his opinion, and then designs a machine that behaves that way. This new machine may in fact work very well. But, I must warn you that that does not tell us anything about how the brain actually works, nor is it necessary to ever really know that in order to make a computer very capable. It is not necessary to understand the way birds flap their wings and how the feathers are designed in order to make a flying machine. It is not necessary to understand the lever system in the legs of a cheetah, that is an animal that runs fast, in order to make an automobile with wheels that goes very fast. It is therefore not necessary to imitate the behavior of Nature in detail in order to engineer a device which can in many respects surpass Nature's abilities.

It is an interesting subject and I like to talk about it. Your brain is very weak compared to a computer. I will give you a series of numbers, one, three, seven, oh yes, ichi, san, shichi, san, ni, go, ni, go, ichi, hachi, ichi, ni, ku, san, go. I want you to repeat them back. But, a computer can take ten thousands numbers and give me them back in reverse every other one, or sum them or lots of things that we cannot do. On the other hand, if I look at a face, in a glance I can tell you who it is if I know that person, or that I don't know that person. But, we do not know how to make a computer system so that if we give it a pattern of a face it can tell us who he is,

even if it has seen many faces and you try to teach it. We do not know how to make computers do that, yet.

Another interesting example is chess playing machines. It is quite a surprise that we can make machines that play chess better than almost everybody in the room. But, they do it by trying many many possibilities. If he moves here, then I could move here and he can move there and so forth. They look at each alternative and choose the best. Now, millions of alternatives are looked at. But, a master chess player, a human, does it differently. He recognizes patterns. He looks at only thirty or forty positions before deciding what move to make. Therefore, although the rules are simpler in Go, machines that play Go are not very good, because in each position there are too many possibilities to move and there are too many things to check and the machines cannot look deeply. Therefore the problem of recognizing patterns and what to do under the circumstances is the thing that the computer engineers (they like to call themselves computer scientists) still find very difficult, and it is certainly one of the important things for future computers, perhaps more important than the things I spoke about. Make a machine to play Go effectively.

Q: I think that any method of computation would not be fruitful unless it would give a kind of provision on how to compose such devices or programs. I thought the Fredkin paper on conservative logic was very intriguing, but once I came to think of making a simple program using such devices I came to a halt because thinking out such a program is far more complex than the program itself. I think we could easily get into a kind of infinite regression because the process of making out a certain program would be much more complex than the program itself and in trying to automate the process the automating program would be more complex and so on. Especially in this case where the program is hard wired rather than being separated as a software, I think it is fundamental to think of the ways of composition.

A: We have some different experiences. There is no infinite

regression; it stops at a certain level of complexity. The machine that Fredkin ultimately is talking about and the one that I was talking about in the quantum mechanical case are both universal computers in the sense that they can be programed to do various jobs; this is not a hard-wired program; they are no more hard-wired than an ordinary computer that you can put information in, that the program is a part of the input, and the machine does the problem that it is assigned to do. It is hard-wired but it is universal like an ordinary computer. These things are very uncertain but I found a minimum. If you have a program written for an irreversible machine, the ordinary program, then I can convert it to a reversible machine program by a direct translation scheme, which is very inefficient and uses many more steps. Then in real situations, the number of steps can be much less. But at least I know that I can take a program with a 2^n steps where it is irreversible, convert it to 3^n steps of a reversible machine. That is many more steps. I did it very inefficiently; I did not try to find the minimum. Just a way. I don't really think that we'll find this regression that you speak of, but you might be right. I am uncertain.

Q: Won't we be sacrificing many of the merits we were expecting of such devices, because those reversible machines run so slow? I am very pessimistic about this point.

A: They run slower, but they are very much smaller. I don't make it reversible unless I need to. There is no point in making the machine reversible unless you are trying very hard to decrease the energy enormously, rather ridiculously, because with only 80 times kT the irreversible machine functions perfectly. That 80 is much less than the present day 10^9 or 10^{10}, so I have at least 10^7 improvement in energy to make, and can still do it with irreversible machines! That's true. That's the right way to go, for the present. I entertain myself intellectually for fun, to ask how far could we go in principle, not in practice, and then I discover that I can go to a fraction of a kT of energy and make the machines microscopic, atomically microscopic. But to do so, I must use the reversible

physical laws. Irreversibility comes because the heat is spread over a large number of atoms and can't be gathered back again. When I make the machine very small, unless I allow a cooling element which is lots of atoms, I have to work reversibly. In practice there probably will never come a time when we will be unwilling to tie a little computer to a big piece of lead which contains 10^{10} atoms (which is still very small indeed), making it effectively irreversible. Therefore I agree with you that in practice, for a very long time and perhaps forever, we will use irreversible gates. On the other hand it is a part of the adventure of science to try to find a limitations in all directions and to stretch a human imagination as far as possible everywhere. Although at every stage it has looked as if such an activity was absurd and useless, it often turns out at least not to be useless.

Q: Are there any limitations from the uncertainty principle? Are there any fundamental limitations on the energy and the clock time in your reversible machine scheme?

A: That was my exact point. There is no further limitation due to quantum mechanics. One must distinguish carefully between the energy lost or consumed irreversibly, the heat generated in the operation of the machine, and the energy content of the moving parts which might be extracted again. There is a relationship between the time and the energy which might be extracted again. But that energy which can be extracted again is not of any importance or concern. It would be like asking whether we should add the mc^2, rest energy, of all the atoms which are in the device. I only speak of the energy lost times the time, and then there is no limitation. However it is true that if you want to make a calculation at a certain extremely high speed, you have to supply to the machine parts which move fast and have energy but that energy is not necessarily lost at each step of the calculation; it coasts through by inertia.

A (to no Q): Could I just say with regard to the question of useless ideas? I'd like to add one more. I waited, if you would ask me, but you didn't. So I answer it anyway. How would we make a machine of

such small dimension where we have to put the atoms in special places? Today we have no machinery with moving parts whose dimension is extremely small or atomic or hundreds of atoms even, but there is no physical limitation in that direction either. And there is no reason why, when we lay down the silicon even today, the pieces cannot be made into little islands so that they are movable. And we could arrange small jets so we could squirt the different chemicals on certain locations. We can make machinery which is extremely small. Such machinery will be easy to control by the same kind of computer circuits that we make. Ultimately, for fun again and intellectual pleasure, we could imagine machines tiny like few microns across with wheels and cables all interconnected by wires, silicon connections, so that the thing as a whole, a very large device, moves not like the awkward motion of our present stiff machines but in a smooth way of the neck of a swan, which after all is a lot of little machines, the cells all interconnected and all controlled in a smooth way. Why can't we do that ourselves?

Foundations of Physics, Vol. 16, No. 6, 1986

Quantum Mechanical Computers[1]

Richard P. Feynman[2]

Received March 15, 1985

The physical limitations, due to quantum mechanics, on the functioning of computers are analyzed.

1. INTRODUCTION

This work is a part of an effort to analyze the physical limitations of computers due to the laws of physics. For example, Bennett[1] has made a careful study of the free energy dissipation that must accompany computation. He found it to be virtually zero. He suggested to me the question of the limitations due to quantum mechanics and the uncertainty principle. I have found that, aside from the obvious limitation to size if the working parts are to be made of atoms, there is no fundamental limit from these sources either.

We are here considering ideal machines; the effects of small imperfections will be considered later. This study is one of principle; our aim is to exhibit some Hamiltonian for a system which could serve as a computer. We are not concerned with whether we have the most efficient system, nor how we could best implement it.

Since the laws of quantum physics are reversible in time, we shall have to consider computing engines which obey such reversible laws. This problem already occurred to Bennett,[1] and to Fredkin and Toffoli,[2] and a great deal of thought has been given to it. Since it may not be familiar to

[1] *Editor's note*: This article, which is based on the author's plenary talk presented at the CLEO/IQEC Meeting in 1984, originally appeared in the February 1985 issue of *Optics News*. It is here reprinted with kind permission of Professor Feynman and *Optics News*.

[2] Department of Physics, California Institute of Technology, Pasadena, California 91125.

you here, I shall review this, and in doing so, take the opportunity to review, very briefly, the conclusions of Bennett,[2] for we shall confirm them all when we analyze our quantum system.

It is a result of computer science that a universal computer can be made by a suitably complex network of interconnected primitive elements. Following the usual classical analysis we can imagine the interconnections to be ideal wires carrying one of two standard voltages representing the local 1 and 0. We can take the primitive elements to be just two, NOT and AND (actually just the one element NAND = NOT AND suffices, for if one input is set at 1 the output is the NOT of the other input). They are symbolized in Fig. 1, with the logical values resulting on the outgoing wires, resulting from different combinations of input wires.

From a logical point of view, we must consider the wires in detail, for in other systems, and our quantum system in particular, we may not have wires as such. We see we really have two more logical primitives, FAN OUT when two wires are connected to one, and EXCHANGE, when wires are crossed. In the usual computer the NOT and NAND primitives are implemented by transistors, possibly as in Fig. 2.

What is the minimum free energy that must be expended to operate an ideal computer made of such primitives? Since, for example, when the AND operates the output line, c' is being determined to be one of two values, no matter what it was before, the entropy change is ln 2 units. This represents a heat generation of kT ln 2 at temperature T. For many years it was thought that this represented an absolute minimum to the quantity of heat per primitive step that had to be dissipated in making a calculation.

The question is academic at this time. In actual machines we are quite concerned with the heat dissipation question, but the transistor system used actually dissipates about $10^{10}kT$! As Bennett[3] has pointed out, this arises because to change a wire's voltage we dump it to ground through a resistance; and to build it up again we feed charge, again through a resistance, to the wire. It could be greatly reduced if energy could be stored in an inductance, or other reactive element.

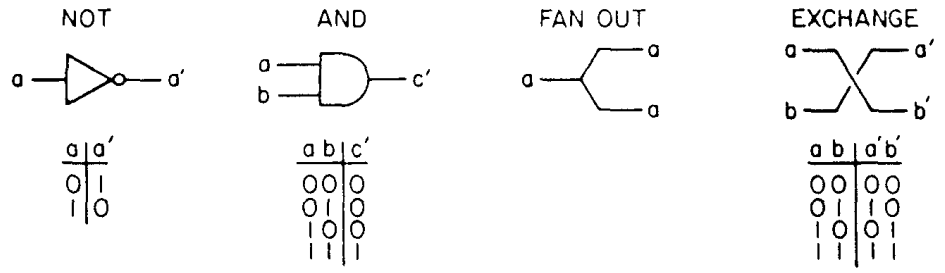


Fig. 1. Primitive elements.

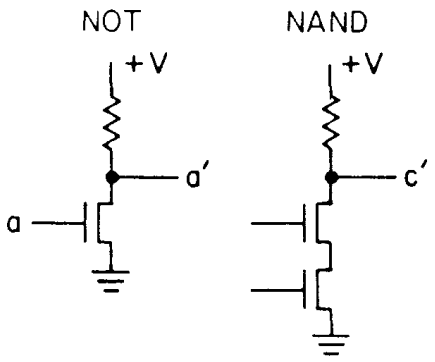


Fig. 2. Transistor circuits for NOT and NAND.

However, it is apparently very difficult to make inductive elements on silicon wafers with present techniques. Even Nature, in her DNA copying machine, dissipates about $100kT$ per bit copied. Being, at present, so very far from this $kT \ln 2$ figure, it seems ridiculous to argue that even this is too high and the minimum is really essentially zero. But, we are going to be even more ridiculous later and consider bits written on one atom instead of the present 10^{11} atoms. Such nonsense is very entertaining to professors like me. I hope you will find it interesting and entertaining also.

What Bennett pointed out was that this former limit was wrong because it is not necessary to use irreversible primitives. Calculations can be done with reversible machines containing only reversible primitives. If this is done the minimum free energy required is independent of the complexity or number of logical steps in the calculation. If anything, it is kT per bit of the output answer.

But even this, which might be considered the free energy needed to clear the computer for further use, might also be considered as part of what you are going to do with the answer—the information in the result if you transmit it to another point. This is a limit only achieved ideally if you compute with a reversible computer at infinitesimal speed.

2. COMPUTATION WITH A REVERSIBLE MACHINE

We will now describe three reversible primitives that could be used to make a universal machine (Toffoli[4]). The first is the NOT which evidently loses no information, and is reversible, being reversed by acting again with NOT. Because the conventional symbol is not symmetrical we shall use an X on the wire instead (see Fig. 3a).

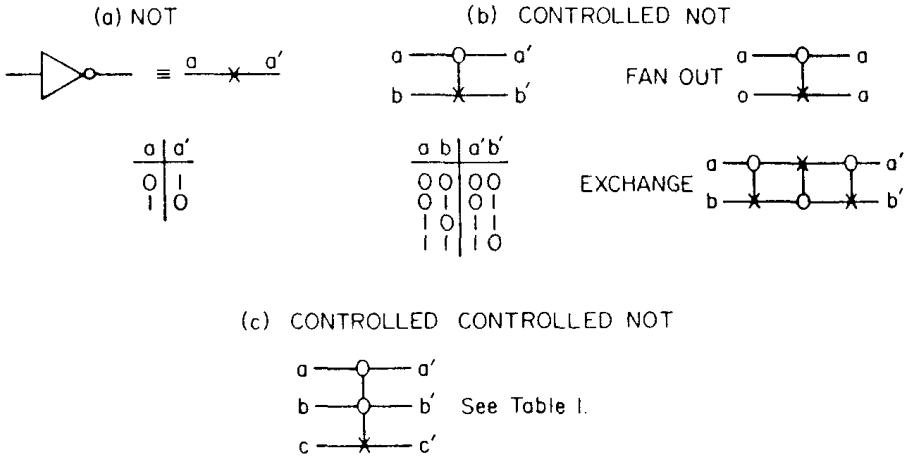


Fig. 3. Reversible primitives.

Next is what we shall call the CONTROLLED NOT (see Fig. 3b). There are two entering lines, a and b, and two exiting lines, a' and b'. The a' is always the same as a, which is the control line. If the control is activated $a = 1$ then the out b' is the NOT of b. Otherwise b is unchanged, $b' = b$. The table of values for input and output is given in Fig. 3. The action is reversed by simply repeating it.

The quantity b' is really a symmetric function of a and b called XOR, the exclusive or; a or b but not both. It is likewise the sum modulo 2 of a and b, and can be used to compare a and b, giving a 1 as a signal that they are different. Please notice that this function XOR is itself not reversible. For example, if the value is zero we cannot tell whether it came from $(a, b) = (0, 0)$ or from $(1, 1)$ but we keep the other line $a' = a$ to resolve the ambiguity.

We will represent the CONTROLLED NOT by putting a 0 on the control wire, connected with a vertical line to an X on the wire which is controlled.

This element can also supply us with FAN OUT, for if $b = 0$ we see that a is copied onto line b'. This COPY function will be important later on. It also supplies us with EXCHANGE, for three of them used

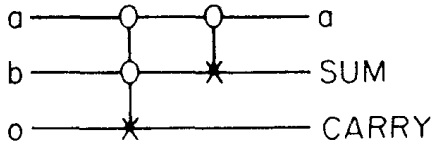


Fig. 4. Adder.

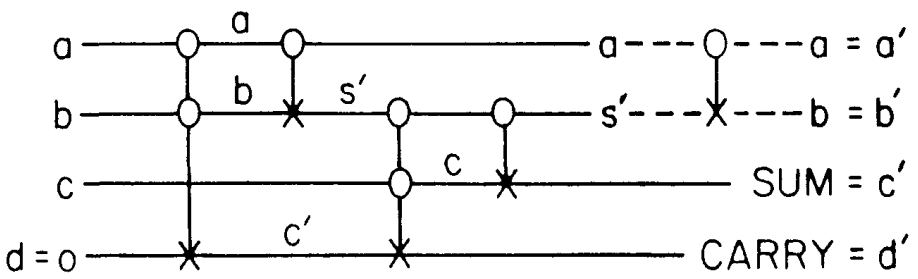


Fig. 5. Full adder.

successively on a pair of lines, but with alternate choice for control line, accomplishes an exchange of the information on the lines (Fig. 3b).

It turns out that combinations of just these two elements alone are insufficient to accomplish arbitrary logical functions. Some element involving three lines is necessary. We have chosen what we can call the CONTROLLED CONTROLLED NOT. Here (see Fig. 3c) we have two control lines a, b, which appear unchanged in the output and which change the third line c to NOT c only if both lines are activated ($a = 1$ and $b = 1$). Otherwise $c' = c$. If the third line input c is set to 0, then evidently it becomes $1 (c' = 1)$ only if both a and b are 1 and therefore supplies us with the AND function (see Table I).

Three combinations for (a, b), namely (0, 0), (0, 1), and (1, 0), all give the same value, 0, to the AND (a, b) function so the ambiguity requires two bits to resolve it. These are kept in the lines a, b in the output so the function can be reversed (by itself, in fact). The AND function is the carry bit for the sum of a and b.

From these elements it is known that any logical circuit can be put together by using them in combination, and in fact, computer science

Table I.

a	b	c	a'	b'	c'
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	1
1	1	1	1	1	0

shows that a universal computer can be made. We will illustrate this by a little example. First, of course, as you see in Fig. 4, we can make an adder, by first using the CONTROLLED CONTROLLED NOT and then the CONTROLLED NOT in succession, to produce from a and b and 0, as input lines, the original a on one line, the sum on the second line, and the carry on the third.

A more elaborate circuit is a full adder (see Fig. 5), which takes a carry c (from some previous addition) and adds it to the two lines a and b and has an additional line d with a 0 input. It requires four primitive elements to be put together. Besides this total sum, the total of the three, a, b, and c and the carry, we obtain on the other two lines two pieces of information. One is the a that we started with, and the other is some intermediary quantity that we calculated on route.

This is typical of these reversible systems; they produce not only what you want in output, but also a certain amount of garbage. In this particular case, and as it turns out in all cases, the garbage can be arranged to be, in fact, just the input, if we would just add the extra CONTROLLED NOT on the first two lines, as indicated by the dotted lines in Fig. 5; we see that the garbage would become a and b, which were the inputs of at least two of the lines. (We know this circuit can be simplified but we do it this way for illustrative purposes.)

In this way, we can by various combinations produce a general logic unit that transforms n bits to n bits in a reversible manner. If the problem you are trying to do is itself reversible, then there might be no extra garbage, but in general, there are some extra lines needed to store up the information which you would need to be able to reverse the operation. In other words, we can make any function that the conventional system can, plus garbage. The garbage contains the information you need to reverse the process.

And how much garbage? It turns out in general that if the output data that you are looking for has k bits, then starting with an input and k bits containing 0, we can produce, as a result, just the input and the output and no further garbage. This is reversible because knowing the output and the input permits you, of course, to undo everything. This proposition is always reversible. The argument for this is illustrated in Fig. 6.

Suppose we began with a machine M, which, starting with an input, and some large number of 0's, produces the desired outut plus a certain amount of extra data which we call garbage. Now we have seen that the copy operation which can be done by a sequence of CONTROLLED NOT's is possible, so if we have originally an empty register, with the k bits ready for the output, we can, after the processor M has operated, copy the output from the M onto this new register.

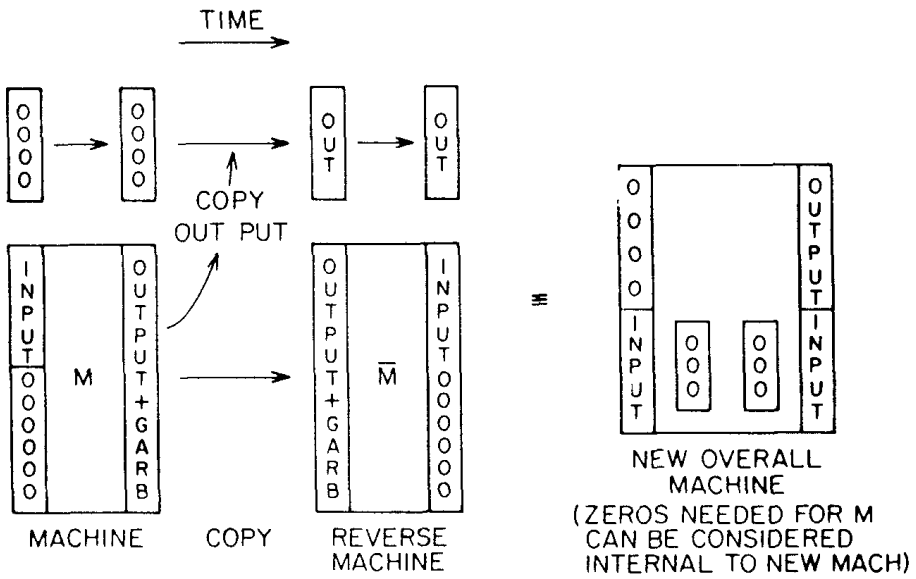


Fig. 6. Clearing garbage.

After that, we can build the opposite machine, the M in reverse, the reverse machine, which would take this output of M and garbage and turn it into the input and 0's. Thus, seen as an overall machine, we would have started with the k 0's of the register for the output, and the input, and ended up with those k 0's occupied by the output data, and repeat the inut as a final product. The number of 0's that was originally needed in the M machine in order to hold the garbage is restored again to 0, and can be considered as internal wires inside the new complete machine (M, $\bar{M}$ and copy).

Overall, then, we have accomplished what we set out to do, and therefore garbage need never be any greater than a repetition of the input data.

3. A QUANTUM MECHANICAL COMPUTER

We now go on to consider how such a computer can also be built using the laws of quantum mechanics. We are going to write a Hamiltonian, for a system of interacting parts, which will behave in the same way as a large system in serving as a universal computer. Of course the large system also obeys quantum mechanics, but it is in interaction with the heat baths and other things that could make it effectively irreversible.

What we would like to do is make the computer as small and as

simple as possible. Our Hamiltonian will describe in detail all the internal computing actions, but not, of course, those interactions with the exterior involved in entering the input (preparing the initial state) and reading the output.

How small can such a computer be? How small, for instance, can a number be? Of course a number can be represented by bits of 1's and 0's. What we are going to do is imagine that we have two-state systems, which we will call "atoms." An n bit number is then represented by a state of a "register," a set of n two-state systems.

Depending upon whether or not each atom is in one or another of its two states, which we call $|1\rangle$ and $|0\rangle$, we can of course, represent any number. And the number can be read out of such a register by determining, or measuring, in which state each of the atoms are at a given moment. Therefore one bit will be represented by a single atom being in one of two states, the states we will call $|1\rangle$ and $|0\rangle$.

What we will have to do then can be understood by considering an example; the example of a CONTROLLED CONTROLLED NOT. Let G be some sort of an operation on three atoms a, b, and c, which converts the original state of a, b, and c into a nex appropriate state, a', b', c', so that the connection between a', b', and c' and a, b, c, are just what we would have expected if a, b, and c represented wires, and the a', b', and c' were the output wires of a CONTROLLED CONTROLLED NOT.

It must be appreciated here that at the moment we are not trying to move the data from one position to another; we are just going to change it. Unlike the situation in the actual wired computer in which the voltages on one wire then go over to voltages on another, what we are specifically making is something simpler, that the three atoms are in some particular state, and that an operation is performed, which changes the state to new values, a', b', c'.

What we would have then is that the state, in the mathematical form $|a', b', c'\rangle$, is simply some operation G operating on $|a, b, c\rangle$. In quantum mechanics, state changing operators are linear operators, and so we'll suppose that G is linear. Therefore, G is a matrix, and the matrix elements of G, $G_{a',b',c',a,b,c}$ are all 0 except those in Table I, which are of course 1.

This table is the same table that represents the truth value table for the CONTROLLED CONTROLLED NOT. It is apparent that the operation is reversible, and that can be represented by saving that $G^*G = 1$, where the * means Hermitian adjoint. That is to say, G is a unitary matrix. (In fact G is also a real matrix $G^* = G$, but that's only a special case.) To be more specific, we are going to write $A_{ab,c}$ for this special G. We shall use the same matrix A with different numbers of subscripts to represent the other primitive elements.

To take a simple example, the NOT, which would be represented by A_a, is the simple matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

This is a 2×2 matrix and can be represented in many ways, in different notations, but the particular one we will use to define these is by the method of creation and annihilation operators. Consider operating in this case, on a single line a. In order to save alphabets, let us call $\underline{a}$ the matrix

$$\underline{a} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

which annihilates the 1 on atom a and converts it to 0; $\underline{a}$ is an operator which converts the state $|1\rangle$ to $|0\rangle$. But, if the state of the atom were originally $|0\rangle$, the operator $\underline{a}$ produces the number 0. That is, it doesn't change the state, it simply produces the numerical value zero when operating on that state. The conjugate of this thing, of course, is

$$\underline{a^*} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

which creates, in the sense that operating on the 0 state, it turns it to the 1 state. In other words, it moves from $|0\rangle$ to $|1\rangle$. When operating on the $|1\rangle$ state there is no further state above that which you can create, and therefore it gives it the number zero. Every other operator 2×2 matrix can be represented in terms of these $\underline{a}$ and $\underline{a^*}$. For example, the product $\underline{a^*}\underline{a}$ is equal to the matrix

$$\underline{a^* a} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

which you might call N_a. It is 1 when the state is $|1\rangle$ and 0 when the state is $|0\rangle$. It gives the number that the state of the atom represents. Likewise the product

$$\underline{aa^*} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

is $1 - N_a$, and gives 0 for the up-state and 1 for the down-state. We'll use 1 to represent the diagonal matrix,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

As a consequence of all this, $\underline{aa^*} + \underline{a^*}\underline{a} = 1$.

It is evident then that our matrix for NOT, the operator that produces NOT, is $A_a = \underline{a} + \underline{a}^*$ and further of course, that's reversible, $A_a{}^* A_a = 1$, A_a is unitary.

In the same way the matrix $A_{a,b}$ for the CONTROLLED NOT can be worked out. If you look at the table of values for CONTROLLED NOT you see that it can be written this way:

$$\underline{a}^* \underline{a}(\underline{b} + \underline{b}^*) + \underline{a}\underline{a}^*$$

In the first term, the $\underline{a}^* \underline{a}$ selects the condition that the line $a = 1$, in which case we want $\underline{b} + \underline{b}^*$ the NOT to apply to b. The second term selects the condition that the line a is 0, in which case we want nothing to happen to b and the unit matrix on the operators of b is implied. This can also be written as $1 + \underline{a}^* \underline{a}(\underline{b} + \underline{b}^* - 1)$, the 1 representing all the lines coming through directly, but in the case that a is 1, we would like to correct that by putting in a NOT instead of leaving the line b unchanged.

The matrix for the CONTROLLED CONTROLLED NOT is

$$A_{ab,c} = 1 + \underline{a}^* \underline{a}\underline{b}^* \underline{b}(\underline{c} + \underline{c}^* - 1)$$

as, perhaps, you may be able to see.

The next question is what the matrix is for a general logic unit which consists of a sequence of these. As an example, we'll study the case of the full adder which we described before (see Fig. 5). Now we'll have, in the general case, four wires represented by a, b, c, and d; we don't necessarily have to have d as 0 in all cases, and we would like to describe how the object operates in general (if d is changed to 1 d' is changed to its NOT). It produces new numbers a', b', c', and d', and we could imagine with our system that there are four atoms labeled a, b, c, d in a state labeled $|a, b, c, d\rangle$ and that a matrix M operates which changes these same four atoms so that they appear to be in the state $|a', b', c', d'\rangle$ which is appropriate for this logic unit. That is, if $|\psi_{\text{in}}\rangle$ represents the incoming state of the four bits, M is a matrix which generates an outgoing state $|\psi_{\text{out}}\rangle = M|\psi_{\text{in}}\rangle$ for the four bits.

For example, if the input state were the state $|1, 0, 1, 0\rangle$, then, as we know, the output state should be $|1, 0, 0, 1\rangle$; the first two a', b' should be 1, 0 for those two first lines come streight through, and the last two c', d' should be 0, 1 because that represents the sum and carry of the first three, a, b, c, bits in the first input, as $d = 0$. Now the matrix M for the adder can easily be seen as the result of five successive primitive operations, and therefore becomes the matrix product of the five successive matrices representing these primitive objects.

$$M = A_{a,b} A_{b,c} A_{bc,d} A_{a,b} A_{ab,d}$$

The first, which is the one written farthest to the right, is $A_{ab,d}$ for that represents the CONTROLLED CONTROLLED NOT in which a and b are the CONTROL lines, and the NOT appears on line d. By looking at the diagram in Fig. 5 we can immediately see what the remaining factors in the sequence represent. The last factor, for example, $A_{a,b}$, means that there's a CONTROLLED NOT with a CONTROL on line a and NOT on line b. This matrix will have the unitary property $M^*M = 1$ since all of the A's out of which it is a product are unitary. That is to say, M is a reversal operation, and M^* is its inverse.

Our general problem, then, is this. Let $A_1, A_2, A_3,,,,' A_k$ be the succession of operations wanted, in some logical unit, to operate on n lines. The $2'' \times 2''$ matrix M needed to accomplish the same goal is a product $A_k \cdots A_3 A_2 A_1$, where each A is a simple matrix. How can we generate this M in a physical way if we know how to make the simpler elements?

In general, in quantum mechanics, the outgoing state at time t is $e^{iHt}\psi_{\text{in}}$, where ψ_{in} is the input state, for a system with Hamiltonian H. To try to find, for a given special time t, the Hamiltonian which will produce $M = e^{iHt}$ when M is such a product of noncommuting matrices, from some simple property of the matrices themselves, appears to be very difficult.

We realize however, that at any particular time, if we expand the e^{iHt} out (as $1 + iHt - H^2t^2/2 - \cdots$) we'll find the operator H operating an innumerable arbitrary number of times, once, twice, three times, and so forth, and the total state is generated by a superposition of these possibilities. This suggests that we can solve this problem of the composition of these A's in the following way.

We add to the n atoms, which are in our register, an entirely new set of $k+1$ atoms, which we'll call "program counter sites." Let us call q_i and q_i^* the annihilation and creation operators for the program site i for $i = 0$ to k. A good thing to think of, as an example, is an electron moving from one empty site to another. If the site i is occupied by the electron, its state is $|1\rangle$, while if the site is empty, its state is $|0\rangle$.

We write, as our Hamiltonian

$$H = \sum_{i=0}^{k-1} q_{i+1}^* q_i A_{i+1} + \text{complex conjugate}$$

$$= q_1^* q_0 A_1 + q_2^* q_1 A_2 + q_3^* q_2 A_3 + \cdots + q_0^* q_1 A_1^*$$

$$+ q_1^* q_2 A_2^* + q_2^* q_3 A_3^* + \cdots$$

The first thing to notice is that if all the program sites are unoccupied, that is, all the program atoms are initially in the state 0, nothing happens

because every term in the Hamiltonian starts with an annihilation operator and it gives 0 therefore.

The second thing we notice is that if only one or another of the program sites is occupied (in state $|1\rangle$), and the rest are not (state $|0\rangle$), then this is always true. In fact the number of program sites that are in state $|1\rangle$ is a conserved quantity. We will suppose that in the operation of this computer, either no sites are occupied (in which case nothing happens) or just one site is occupied. Two or more program sites are never both occupied during normal operation.

Let us start with an initial state where site 0 is occupied, is in the $|1\rangle$ state, and all the others are empty, $|0\rangle$ state. If later, at some time, the final site k is found to be in the $|1\rangle$ state, (and therefore all the others in $|0\rangle$) then, we claim, the n register has been multiplied by the matrix M, which is $A_k \cdots A_2 A_1$ as desired.

Let me explain how this works. Suppose that the register starts in any initial state, ψ_{in}, and that the site, 0, of the program counter is occupied. Then the only term in the entire Hamiltonian that can first operate, as the Hamiltonian operates in successive times, is the first term, $q_1^* q_0 A_1$. The q_0 will change site number 0 to an unoccupied site, while q_1^* will change the site number 0 to an occupied site. Thus the term $q_1^* q_0$ is a term which simply moves the occupied site from the location 0 to the location 1. But this is multiplied by the matrix A_1 which operates only on the n register atoms, and therefore multiplies the initial state of the n register atoms by A_1.

Now, if the Hamiltonian happens to operate a second time, this first term will produce nothing because q_0 produces 0 on the number 0 site because it is now unoccupied. The term which can operate now is the second term, $q_2^* q_1 A_2$, for that can move the occupied point, which I shall call a "cursor." The cursor can move from site 1 to site 2 but the matrix A_2 now operates on the register; therefore the register has now got the matrix $A_2 A_1$ operating on it.

So, looking at the first line of the Hamiltonian, if that is all there was to it, as the Hamiltonian operates in successive orders, the cursor would move successively from 0 to k, and you would acquire, one after the other, operating on the n register atoms, the matrices, A, in the order that we would like to construct the total M.

However, a Hamiltonian must be hermitian, and therefore the complex conjugate of all these operators must be present. Suppose that at a given stage, we have gotten the cursor on site number 2, and we have the matrix $A_2 A_1$ operating on the register. Now the q_2 which intends to move that occupation to a new position need not come from the first line, but may have come from the second line. It may have come, in fact, from

$q_1^* q_2 A_2^*$ which would move the cursor back from the position 2 to the position 1.

But note that when this happens, the operator A_2^* operates on the register, and therefore the total operator on the register is $A_2^* A_2 A_1$ in this case. But $A_2^* A_2$ is 1 and therefore the operator is just A_1. Thus we see that when the cursor is returned to the position 1, the net result is that only the operator A_1 has really operated on the register. Thus it is that as the various terms of the Hamiltonian move the cursor forwards and backwards, the A's accumulate, or are reduced out again.

At any stage, for example, if the cursor were up to the j site, the matrices from A_1 to A_j have operated in succession on the n register. It does not matter whether or not the cursor on the j site has arrived there, by going directly from 0 to j, or going further and returning, or going back and forth in any pattern whatsoever, as long as it finally arrived at the state j.

Therefore it is true that if the cursor is found at the site k, we have the net result for the n register atoms that the matrix M has operated on their initial state as we desired.

How then could we operate this computer? We begin by putting the input bits onto the register, and by putting the cursor to occupy the site 0. We then check at the site k, say, by scattering electrons, that the site k is empty, or that the site k has a cursor. The moment we find the cursor at site k we remove the cursor so that it cannot return down the program line, and then we know that the register contains the output data. We can then measure it at our leisure. Of course, there are external things involved in making the measurements, and determining all of this, which are not part of our computer. Surely a computer has eventually to be in interaction with the external world, both for putting data in and for taking it out.

Mathematically it turns out that the propagation of the cursor up and down this program line is exactly the same as it would be if the operators A were not in the Hamiltonian. In other words, it represents just the waves which are familiar from the propagation of the tight binding electrons or spin waves in one dimension, and are very well known. There are waves that travel up and down the line and you can have packets of waves and so forth.

We could improve the action of this computer and make it into a ballistic action in the following way: by making a line of sites in addition to the ones inside, that we are actually using for computing, a line of say, many sites, both before and after. It's just as though we had values of the index i for q_i, which are less than 0 and greater than k, each of which has no matrix A, just a 1 multiplying there. Then we had have a longer spin chain, and we could have started, instead of putting a cursor exactly at the

beginning site 0, by putting the cursor with different amplitudes on different sites representing an initial incoming spin wave, a wide packet of nearly definite momentum.

This spin wave would then go through the entire computer in a ballistic fashion and out the other end into the outside tail that we have added to the line of program sites, and there it would be easier to determine if it is present and to steer it away to some other place, and to capture the cursor. Thus, the logical unit can act in a ballistic way.

This is the essential point and indicates, at least to a computer scientist, that we could make a universal computer, because he knows if we can make any logical unit we can make a universal computer. That this could represent a universal computer for which composition of elements and branching can be done is not entirely obvious unless you have some experience, but I will discuss that to some further extent later.

4. IMPERFECTIONS AND IRREVERSIBLE FREE ENERGY LOSS

There are, however, a number of questions that we would like to discuss in more detail such as the question of imperfections.

There are many sources of imperfections in this machine, but the first one we would like to consider is the possibility that the coefficients in the couplings, along the program line, are not exactly equal. The line is so long that in a real calculation little irregularities would produce a small probability of scattering, and the waves would not travel exactly ballistically, but would go back and forth.

If the system, for example, is built so that these sites are built on a substrate of ordinary physical atoms, then the thermal vibrations of these atoms would change the couplings a little bit and generate imperfections. (We should even need such noise for with small fixed imperfections there are shallow trapping regions where the cursor may get caught.) Suppose then, that there is a certain probability, say p per step of calculation (that is, per step of cursor motion, $i \to i+1$), for scattering the cursor momentum until it is randomized ($1/p$ is the transport mean free path). We will suppose that the p is fairly small.

Then in a very long calculation, it might take a very long time for the wave to make its way out the other end, once started at the beginning —because it has to go back and forth so many times due to the scattering. What one then could do would be to pull the cursor along the program line with an external force. If the cursor is, for example, an electron moving from one vacant site to another, this would be just like an electric field trying to pull the electron along a wire, the resistance of which is generated

by the imperfection or the probability of scattering. Under these circumstances we can calculate how much energy will be expended by this external force.

This analysis can be made very simply: it is an almost classical analysis of an electron with a mean free path. Every time the cursor is scattered, I am going to suppose it is randomly scattered forward and backward. In order for the machine to operate, of course, it must be moving forward at a higher probability than it is moving backward. When a scattering occurs therefore, the loss in entropy is the logarithm of the probability that the cursor is moving forward, divided by the probability the cursor was moving backward.

This can be approximated by (the probability forward − the probability backward)/(the probability forward + the probability backward). That was the entropy lost per scattering. More interesting is the entropy lost per net calculational step, which is, of course, simply p times that number. We can rewrite the entropy cost per calculational step as

$$pv_D/v_R$$

where v_D is the drift velocity of the cursor and v_R its random velocity.

Or if you like, it is p times the minimum time that the calculation could be done in (that is, if all the steps were always in the forward direction), divided by the actual time allowed.

The free energy loss per step then, is $kT \times p \times$ the minimum time that the calculation could be done, divided by the actual time that you allow yourself to do it. This is a formula that was first derived by Bennett. The factor p is a coasting factor, to represent situations in which not every site scatters the cursor randomly, but it has only a small probability to be thus scattered.

It will be appreciated that the energy loss per step is not kT but is that divided by two factors. One, $(1/p)$, measures how perfectly you can build the machine and the other is proportional to the length of time that you take to do the calculation. It is very much like a Carnot engine, in which in order to obtain reversibility, one must operate very slowly. For the ideal machine where p is 0, or where you allow an infinite time, the mean energy loss can be 0.

The uncertainty principle, which usually relates some energy and time uncertainty, is not directly a limitation. What we have in our computer is a device for making a computation, but the time of arrival of the cursor and the measurement of the output register at the other end (in other words, the time it takes in which to complete the calculation) is not a define time. It's a question of probabilities, and so there is a considerable uncertainty in the time at which a calculation will be done.

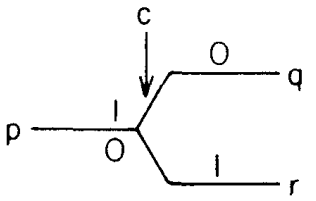
There is no loss associated with the uncertainty of cursor energy; at least no loss depending on the number of calculational steps. Of course, if you want to do a ballistic calculation on a perfect machine, some energy would have to be put into the original wave, but that energy, of course, can be removed from the final wave when it comes out of the tail of the program line. All questions associated with the uncertainty of operators and the irreversibility of measurements are associated with the input and output functions.

No further limitations are generated by the quantum nature of the computer *per se*, nothing that is proportional to the number of computational steps.

In a machine such as this, there are very many other problems, due to imperfections. For example, in the registers for holding the data, there will be problems of cross-talk, interactions between one atom and another in that register, or interaction of the atoms in that register directly with things that are happening along the program line, that we did not exactly bargain for. In other words, there may be small terms in the Hamiltonian besides the ones we have written.

Until we propose a complete implementation of this, it is very difficult to analyze. At least some of these problems can be remedied in the usual way by techniques such as error correcting codes, and so forth, that have been studied in normal computers. But until we find a specific implementation for this computer, I do not know how to proceed to analyze these effects. However, it appears that they would be very important, in practice. This computer seems to be very delicate and these imperfections may produce considerable havoc.

The time needed to make a step of calculation depends on the strength or the energy of the interactions in the terms of the Hamiltonian. If each of the terms in the Hamiltonian is supposed to be of the order of 0.1 electron volts, then it appears that the time for the cursor to make each step, if done in a ballistic fashion, is of the order 6×10^{-15} sec. This does not represent



$H = q^* cp + r^* c^* p + p^* c^* q + p^* cr$

IF c = 1 GO p TO q AND PUT c = 0
IF c = 0 GO p TO r AND PUT c = 1
IF c ≠ 1 GO r TO p AND PUT c = 0
IF c ≠ 0 GO q TO p AND PUT c = 1

Fig. 7. Switch.

an enormous improvement, perhaps only about four orders of magnitude over the present values of the time delays in transistors, and is not much shorter than the very short times possible to achieve in many optical systems.

5. SIMPLIFYING THE IMPLEMENTATION

We have completed the job we set out to do—to find some quantum mechanical Hamiltonian of a system that could compute, and that is all that we need say. But it is of some interest to deal with some questions about simplifying the implementation. The Hamiltonian that we have written involves terms which can involve a special kind of interaction between five atoms. For example, three of them in the register, for a CONTROLLED CONTROLLED NOT, and two of them as the two adjacent sites in the program counter.

This may be rather complicated to arrange. The question is, can we do it with simpler parts. It turns out that we can indeed. We can do it so that in each interaction there are only three atoms. We are going to start with new primitive elements, instead of the ones we began with. We'll have the NOT all right, but we have in addition to that simply a "switch" (see also Priese[5]).

Supposing that we have a term, $q^*cp + r^*c^*p +$ its complex conjugate in the Hamiltonian (in all cases we'll use letters in the earlier part of the alphabet for register atoms and in the latter part of the alphabet for program sites). See Fig. 7. This is a switch in the sense that, if c is originally in the $|1\rangle$ state, a cursor at p will move to q, whereas if c is in the $|0\rangle$ state, the cursor at p will move to r.

During this operation the controlling atom c changes its state. (It is possible also to write an expression in which the control atom does not change its state, such as $q^*c^*cp + r^*cc^*p$ and its complex conjugate but, there is no particular advantage or disadvantage to this, and we will take the simpler form.) The complex conjugate reverses this.

If, however, the cursor is at q and c is in the state $|1\rangle$ (or cursor at r, c in $|0\rangle$), the H gives 0, and the cursor gets reflected back. We shall build all our circuits and choose initial states so that this circumstance will not arise in normal operation, and the ideal ballistic mode will work.

With this switch we can do a number of things. For example, we could produce a CONTROLLED NOT as in Fig. 8. The switch a controls the operation. Assume the cursor starts at s. If $a = 1$ the program cursor is carried along the top line, whereas if $a = 0$ it is carried along the bottom line, in either case terminating finally in the program site t.

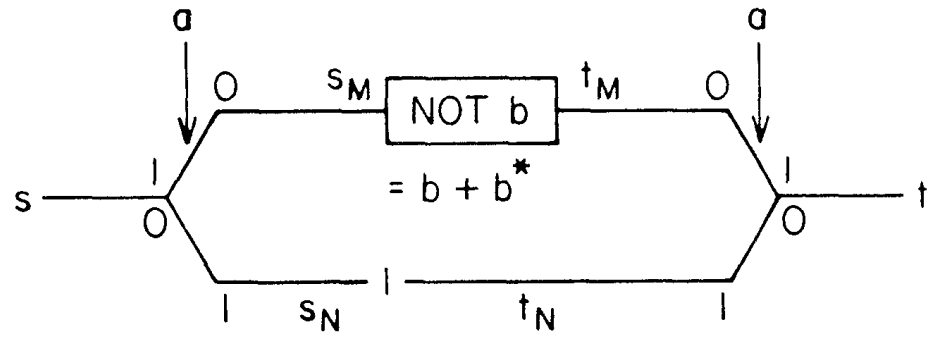


Fig. 8. CONTROLLED NOT by switches.

In these diagrams, horizontal or vertical lines will represent program atoms. The switches are represented by diagonal lines and in boxes we'll put the other matrices that operate on registers such as the NOT b. To be specific, the Hamiltonian for this little section of a CONTROLLED NOT, thinking of it as starting at s and ending at t, is given below:

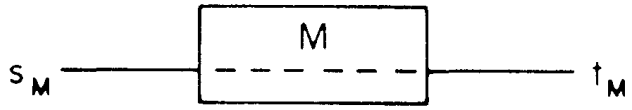
$$H_c(s, t) = s_M^* a s + t^* a^* t_M + t_M^*(b + b^*)\, s_M + s_N^* a^* s + t^* a t_N + t_N^* s_N + \text{c.c}$$

(The c.c means to add the complex conjugate of all the previous terms.)

Although there seem to be two routes here which would possibly produce all kinds of complications characteristic of quantum mechanics, this is not so. If the entire computer system is started in a definite state for a by the time the cursor reaches s, the atom a is still in some definite state (although possibly different from its initial state due to previous computer operations on it). Thus only one of the two routes is taken. The expression may be simplified by omitting the $s_N^* t_N$ term and putting $t_N = s_N$.

One need not be concerned in that case, that one route is longer (two cursor sites) than the other (one cursor site) for again there is no interference. No scattering is produced in any case by the insertion into a chain of coupled sites, an extra piece of chain of any number of sites with the same mutual coupling between sites (analogous to matching impedances in transmission lines).

To study these things further, we think of putting pieces together. A piece (see Fig. 9) M might be represented as a logical unit of interacting parts in which we only represent the first input cursor site as s_M and the final one at the other end as t_M. All the rest of the program sites that are between s_M and t_M are considered internal parts of M, and M contains its registers. Only s_M and t_M are sites that may be coupled externally.



s_M = Starting program site for piece

t_M = Terminal program site for piece

$H_M(s_M, t_M)$ is the part of the Hamiltonian representing all the "atoms" and program sites within the box M, and their interactions with s_M, t_M.

Fig. 9. One "piece."

The Hamiltonian for this subsection we'll call H_M and we'll identify s_M and t_M, as the name of the input and output program sites by writing $H_M(s_M, t_M)$. So therefore H_M is that part of the Hamiltonian representing all the atoms in the box and their external start and terminator sites.

An especially important and interesting case to consider is when the input data (in the regular atoms) comes from one logical unit, and we would like to transfer it to another (see Fig. 10). Suppose that we imagine that the box M starts with its input register with 0 and its output (which may be the same register) also with 0. Then we could use it in the following way. We could make a program line, let's say starting with s'_M whose first job is to exchange the data in an external register which contains the input, with M's input register which at the present time contains 0's.

Then the first step in our calculation, starting, say, at s'_M, would be to make an exchange with the register inside of M. That puts zero's into the original input register and puts the input where it belongs inside the box M. The cursor is now at s_M. (We have already explained how exchange can be made of controlled NOTs.) Then as the program goes from s_M to t_M we find the output now in the box M. Then the output register of M is now cleared as we write the results into some new external register provided for that purpose, originally containing 0's. This we do from t_M to t'_M by exchanging data in the empty external register with the M's output register.

We can now consider connecting such units in different ways. For example, the most obvious way is succession. If we want to do first M and then N we can connect the terminal side of one to the starting side of the other as in Fig. 11, to produce a new effective operator K, and the Hamiltonian then for H_K is

$$H_K(s_K, t_K) = H_M(s_K, t) + H_N(t, t_K)$$

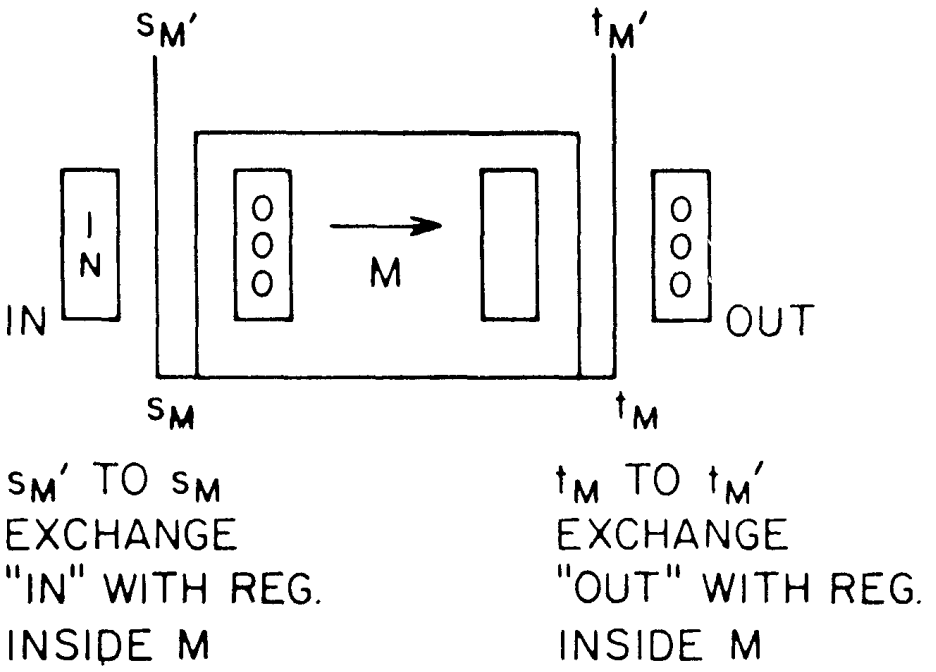


Fig. 10. Piece with external input and output.

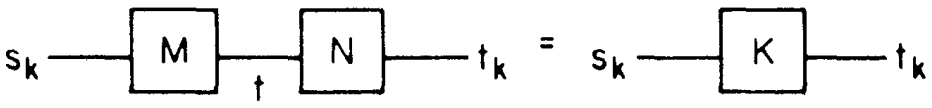
The general conditional, if $a = 1$ do M, but if $a = 0$ do N, can be made, as in Fig. 12. For this

$$H_{\text{cond}}(s_c, t_c) = (s_M^* a s_c + t_c^* a^* t_M + s_N^* a^* s_c + t_c^* a t_N + \text{c.c.}) + H_M(s_M, t_M) + H_N(s_N, t_N)$$

The CONTROLLED NOT is the special case of this with $M = \text{NOT } b$ for which H is

$$H_{\text{NOT}\,b}(s, t) = s^*(b + b^*)\, t + \text{c.c.}$$

and N is no operation $s^* t$.



$$H_k(s_k, t_k) = H_M(s_k, t) + H_N(t, t_k)$$

Fig. 11. Operations in succession.

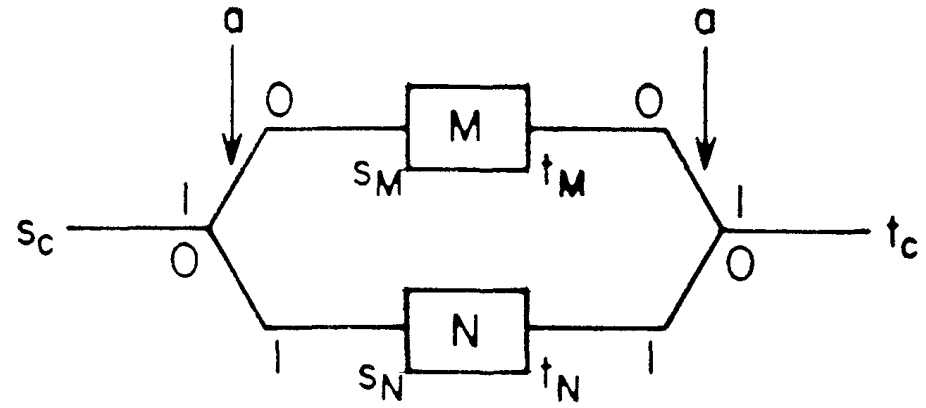


Fig. 12. Conditional if $a = 1$ then M, else N.

As another example, we can deal with a garbage clearer (previously described in Fig. 6) not by making two machines, a machine and its inverse, but by using the same machine and then sending the data back to the machine in the opposite direction, using our switch (see Fig. 13).

Suppose in this system we have a special flag which is originally always set to 0. We also suppose we have the input data in an external register, an empty external register available to hold the output, and the machine registers all empty (containing 0's). We come on the starting line s.

The first thing we do is to copy (using CONTROLLED NOT's) our external input into M. Then M operates, and the cursor goes on the top line in our drawing. It copies the output out of M into the external output register. M now contains garbage. Next it changes f to NOT f, comes down on the other line of the switch, backs out through M clearing the garbage, and uncopies the input again.

When you copy data and do it again, you reduce one of the registers to 0, the register into which you coied the first time. After the coying, it goes out (since f is now changed) on the other line where we restore f to 0

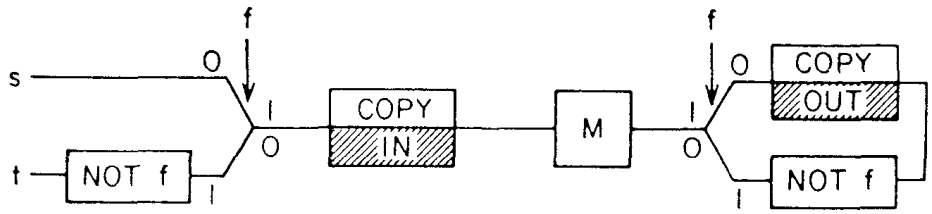


Fig. 13. Garbage clearer.

and come out at t. So between s and t we have a new piece of equipment, which has the following properties.

When its starts, we have, in a register called IN, the input data. In an external register which we call OUT, we have 0's. There is an internal flag set at 0, and the box, M, is empty of all data. At the termination of this, at t, the input register still contains the input data, and the output register contains the output of the effort of the operator M. M, however, is still empty, and the flag f is reset to 0.

Also important in computer programs is the ability to use the same subroutine several times. Of course, from a logical point of view, that can be done by writing that bit of program over and over again, each time it is to be used, but in a practical computer, it is much better if we could build that section of the computer which does a particular operation, just once, and use that section again and again.

To show the possibilities, here, first just suppose we have an operation we simply wish to repeat twice in succession (see Fig. 14). We start at s with the flag a in the condition 0, and thus we come along the line, and the first thing that happens is we change the value of a. Next we do the operation M. Now, because we changed a, instead of coming out at the top line where we went in, we come out at the bottom line, which recirculates the program back into changing a again; it restores it.

This time as we go through M, we come out and we have the a to follow on the uper line, and thus come out at the terminal, t. The Hamiltonian for this is

$$H_{MM}(s, t) = (s_N^* a^* s + s_M^*(a^* + a)\, s_N + x^* a^* t_M + s_N^* a x + t^* a t_M + \text{c.c.}) + H_M(s_M, t_M)$$

Using this switching circuit a number of times, of course, we can repeat an operation several times. For example, using the same idea three

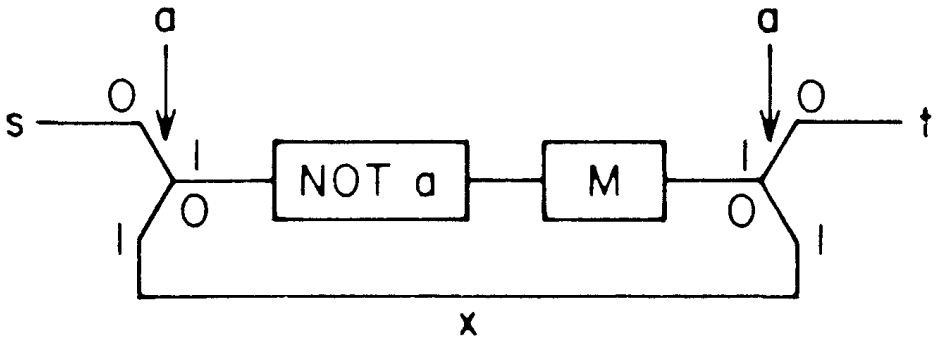


Fig. 14. Do M twice.

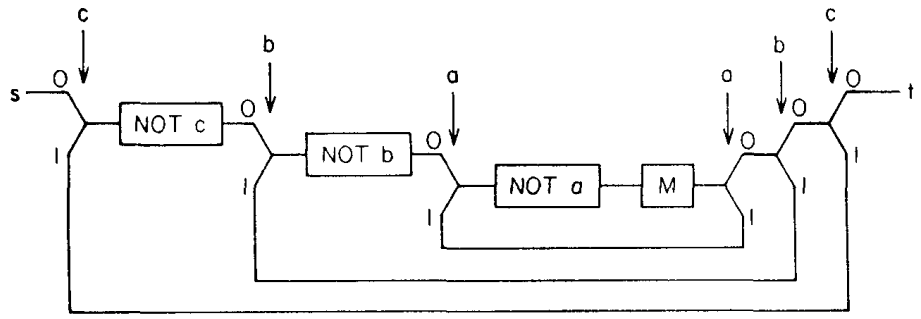


Fig. 15. Do M eight times.

times in succession, a nested succession, we can do an operation eight times, by the apparatus indicated in Fig. 15. In order to do so, we have three flags, a, b, and c. It is necessary to have flags when operations are done again for the reason that we must keep track of how many times its done and where we are in the program or we'll never be able to reverse things.

A subroutine in a normal computer can be used and emptied and used again without any record being kept of what happened. But here we have to keep a record and we do that with flags, of exactly where we are in the cycle of the use of the subroutine. If the subroutine is called from a certain place and has to go back to some other place, and another time is called, its origin and final destination are different, we have to know and keep track of where it came from and where it's supposed to go individually in each case, so more data have to be kept. Using a subroutine over and over in a reversible machine is only slightly harder than in a general machine. All these considerations appear in papers by Fredkin, Toffoli, and Bennett.

It is clear by the use of this switch, and successive uses of such switches in trees, that we would be able to steer data to any oint in a memory. A memory would simply be a place where there are registers into which you could copy data and then return the program. The cursor will

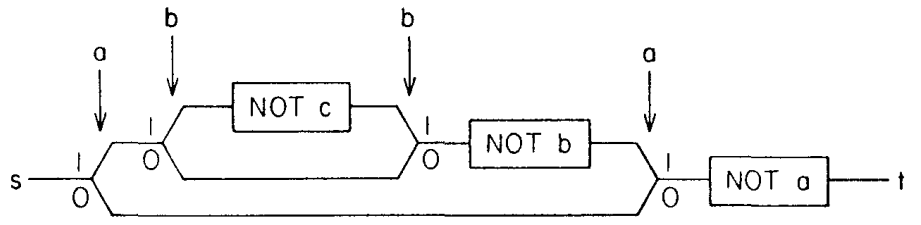


Fig. 16. Increment counter (3-bit).

have to follow the data along. I suppose there must be another set of tree switches set the oposite direction to carry the cursor out again, after copying the data so that the system remains reversible.

In Fig. 16 we show an incremental binary counter (of three bits a, b, c with c the most significant bit) which keeps track of how many net times the cursor has passed from s to t. These few examples should be enough to show that indeed we can construct all computer functions with our SWITCH and NOT. We need not follows this in more detail.

6. CONCLUSIONS

It's clear from these examples that this quentum machine has not really used many of the specific qualities of the differential equations of quantum mechanics.

What we have done is only to try to imitate as closely as possible the digital machine of conventional sequential architecture. It is analogous to the use of transistors in conventional machines, where we do not properly use all the analog continuum of the behavior of transistors, but just try to run them as saturated on or off digital devices so the logical analysis of the system behavior is easier. Furthermore, the system is absolutely sequential—for example, even in the comparison (exclusive or) of two k bit numbers, we must do each bit successively. What can be done, in these reversible quantum systems, to gain the speed available by concurrent operation has not been studied here.

Although, for theoretical and academic reasons, I have studied complete and reversible systems, if such tiny machines could become practical there is no reason why irreversible and entropy creating interactions cannot be made frequently during the course of operations of the machine.

For example, it might prove wise, in a long calculation, to ensure that the cursor has surely reached some oint and cannot be allowed to reverse again from there. Or, it may be found practical to connect irreversible memory storage (for items less frequently used) to reversible logic or short-term reversible storage registers, etc. Again, there is no reason we need to stick to chains of coupled sites for more distant communication where wires or light may be easier and faster.

At any rate, it seems that the laws of physics present no barrier to reducing the size of computers until bits are the size of atoms, and quantum behavior holds dominant sway.

REFERENCES[3]

1. C. H. Bennett, "Logical Reversibility of Computation," *IBM J. Res. Dev.* **6**, 525–532 (1979).
2. E. Fredkin and T. Toffoli, "Conservative Logic," *Int. J. Theor. Phys.* **21**, 219–253 (1982).
3. C. H. Bennett, "Thermodynamics of Computation—A Review," *Int. J. Theor. Phys.* **21**, 905–940 (1982).
4. T. Toffoli, "Bicontinuous Extensions of Invertible Combinatorial Functions," *Math. Syst. Theory* **14**, 13–23 (1981).
5. L. Priese, "On a Simple Combinatorial Structure Sufficient for Sublying Nontrivial Self-Reproduction," *J. Cybern.* **6**, 101–137 (1976).

[3] I would like to thank T. Toffoli for his help with the references.

BIBLIOGRAPHY

1938

[1] With M.S. Vallarta. Scattering of cosmic rays by the stars of a galaxy. *Phys. Rev.* **55**: 340–343.

1939

[2] Forces in molecules, *Phys. Rev.* **56**: 340–343.

1942

[3] *The Principle of Least Action in Quantum Mechanics* (thesis), Princeton. Ann Arbor: University Microfilms.

1945

[4] With J.A. Wheeler. Interaction with the absorber as the mechanism of radiation. *Rev. Mod. Phys.* **17**: 157–181.

1946

[5] Amplifier response. Edited and declassified work from the Manhattan Project. 15pp.
[6] With R.F. Bacher. *Introduction to Atomic Energy.* Dept. of State Publication 2661 USGPO.

1948

[7] Space–time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.* **20**: 367–387.
[8] A relativistic cut-off for classical electrodynamics. *Phys. Rev.* **74**: 939–946.
[9] Relativistic cut-off for quantum electrodynamics. *Phys. Rev.* **74**: 1430–1438.

1949

[10] With J.A. Wheeler. Classical electrodynamics in terms of direct interparticle action. *Rev. Mod. Phys.* **21**: 425–433.
[11] With N. Metropolis and E. Teller. Equations of state of elements based on the generalized Fermi–Thomas theory. *Phys. Rev.* **75**: 1561–1573.
[12] The theory of positrons. *Phys. Rev.* **76**: 749–759.
[13] Space–time approach to quantum electrodynamics. *Phys. Rev.* **76**: 769–789.

1950

[14] Mathematical formulation of the quantum theory of electromagnetic interaction. *Phys. Rev.* **80**: 440–457.

1951

[15] An operator calculus having application in quantum electrodynamics. *Phys. Rev.* **84**: 108–128.

[16] With Carl W. Helstrom, Malvin A. Ruderman, and William Karzas. *High Energy Phenomena and Meson Theories*: notes on course at CIT, Jan.–Mar. 1951.
[17] The concept of probability in quantum mechanics, *Second Berkeley Symposium on Mathematical Statistics and Probability*, 1950, University of California, Berkeley, pp. 553–541.

1952

[18] With Laurie M. Brown. Radiative corrections to Compton scattering. *Phys. Rev.* **85**: 231–244.

1953

[19] The lambda transition in helium. *Phys. Rev.* **90**: 1116–1117.
[20] Atomic theory of the lambda transition in liquid helium. *Phys. Rev.* **91**: 1291–1301.
[21] Atomic theory of liquid helium near absolute zero. *Phys. Rev.* **91**: 1301–1308.
[22] Atomic theory of liquid helium. *Proc. Int. Conf. on Theoretical Physics*, Kyoto and Tokyo, Japan, pp. 895–901.
[23] With M. Barranger and H.A. Bethe. Relativistic corrections to the Lamb shift. *Phys. Rev.* **92**: 482–501.

1954

[24] Atomic theory of the two-fluid model of liquid helium. *Phys. Rev.* **94**: 262–277.
[25] With G. Speisman. Proton–neutron mass difference. *Phys. Rev.* **94**: 500.
[26] The present situation in fundamental theoretical physics. *Academia Brasiliera de Ciencias* **26**: 51–59.

1955

[27] Application of quantum mechanics to liquid helium. *Progress in Low Temperature Physics*, Vol. 1. Amsterdam: North-Holland, pp. 17–53.
[28] Slow electrons in a polar crystal. *Phys. Rev.* **97**: 660–665.
[29] With M. Cohen. The character of the roton state in liquid helium. *Prog. Theor. Phys.* **14**: 261–262.
[30] The value of science. *Engineering and Science* **19**: 13–15 (Dec.).
[31] Helium II in rotational flow. *Science* **121**: 622.

1956

[32] With M. Cohen. Energy spectrum of the excitations in liquid helium. *Phys. Rev.* **102**: 1189–1204.
[33] With F. de Hoffmann and R. Serber. Dispersion of the neutron emission in U235-fission. *High Energy* **3**: 64–69.
[34] With J.E. Mercereau. Physical conditions for ferromagnetic resonance, *Phys. Rev.* **104**: 63.
[35] The relation of science and religion. *Engineering and Science* **19**: 20–23 (June).

1957

[36] Superfluidity and superconductivity. *Rev. Mod. Phys.* **29**: 205–212.

[37] The role of science in the world today. *Proc. Institute of World Affairs* **33**: 17–31.
[38] With M. Cohen. Theory of inelastic scattering of cold neutrons from liquid helium. *Phys. Rev.* **107**: 13–24.
[39] With F.L. Vernon, Jr. and R.W. Hellwarth. Geometrical representation of the Schrödinger equation for solving maser problems. *J. Appl. Phys.* **28**: 49–52.
[40] An historic moment in physics. *Engineering and Science* **20**: 17–18.

1958

[41] With M. Gell-Mann. Theory of the Fermi interaction. *Phys. Rev.* **109**: 193–198.
[42] Excitation in liquid helium. *Physics* **24** (Kammerling Onnes Conference, Leiden), pp. 18–26.

1959

[43] *Series of Lectures on the Theory of Fundamental Processes.* Pasadena: CIT.

1960

[44] There's plenty of room at the bottom. *Engineering and Science* **23**: 22–36 (Feb.).

1961

[45] The present status of quantum electrodynamics. *Extrait des rapports et discussions, Solvay, Institut International de Physique.*
[46] *Quantum Electrodynamics*; a lecture note and reprint volume. New York: W.A. Benjamin. In *Frontiers in Physics* (German edition, Munich, 1989).
[47] *The Theory of Fundamental Processes*; a lecture note volume. New York: W.A. Benjamin (Japanese edition, 1992).
[48] There's plenty of room at the bottom. In *Miniaturization*, ed. Horace D. Gilbert. (See [44].)

1962

[49] With R.W. Hellwarth, C.K. Iddings, and P.M. Platzman. Mobility of slow electrons in a polar crystal. *Phys. Rev.* **127**: 1004–1017.
[50] With R.S. Edgar, S. Klein, I. Lielausis, and C.M. Steinberg. Mapping experiments with r mutants of bacteriophage T4D1. *Genetics* **47**: 179–185.
[51] *Lectures in Elementary Physics.* Pasadena: CIT.
[52] *Quantum Electrodynamics*, 2nd printing with revisions. (See 1961.)
[53] *The Theory of Fundamental Processes.* 2nd printing, corrected. (See 1961.)

1963

[54] The problem of teaching physics in Latin America. *Engineering and Science* **27**: 21–30.
[55] With F.L. Vernon, Jr. The theory of a general quantum mechanical system interacting with a linear dissipative system. *Ann. Phys.* **24**: 118–173.
[56] With Robert B. Leighton and Matthew Sands. *The Feynman Lectures on Physics*, Vol. 1. Reading, Massachusetts: Addison-Wesley.

[57] The quantum theory of gravitation. *Acta Physica Polonica* **24**: 697–722.

1964

[58] With Robert B. Leighton and Matthew Sands. *The Feynman Lectures on Physics*, Vol. 2. Reading, Massachusetts: Addison-Wesley.
[59] The quantum theory of gravitation. In *Proc. on Theory of Gravitation.* Paris: Gauthier-Villars.
[60] The quantum theory of gravitation. In *Proc. on Theory of Gravitation.* PWN-Polish Scientific Publisher, pp. 207–208.
[61] With M. Gell-Mann and G. Zweig. Group $U(6) \times U(6)$ generated by current components. *Phys. Rev. Lett.* **13**: 678–680.
[62] *Lectures on Physics: Exercises.* Reading, Massachusetts: Addison-Wesley.

1965

[63] With Robert B. Leighton and Matthew Sands. *The Feynman Lectures on Physics*, Vol. 3. Reading, Massachusetts: Addison-Wesley.
[64] With Albert R. Hibbs. *Quantum Mechanics and Path Integrals.* New York: McGraw-Hill.
[65] Consequences of SU(3) symmetry in weak interactions. In *Symmetries in Elementary Particle Physics* (Ettore Majorana). New York: Academic, pp. 111–174.
[66] Present status in strong, electromagnetic and weak interactions. In *Symmetries in Elementary Particle Physics* (Ettore Majorana). New York: Academic, pp. 366–399 and 400–418.
[67] New textbooks for the "new" mathematics. *Engineering and Science* **28**: 9–15 (Mar.).
[68] New textbooks for the "new" mathematics. *The California Institute of Technology Quarterly* **6**: 2–9 (Spring).
[69] *The Character of Physical Law.* Cambridge: MIT Press. The Messenger Lectures, Cornell University, 1964.

1966

[70] The development of the space–time view of quantum electrodynamics. *Science* **153**: 699–708.
[71] The development of the space–time view of quantum electrodynamics. *Physics Today* **19** (August): 31–44.
[72] What is and what should be the role of scientific culture in modern society? *Supp. al Nuovo Cimento* **4**: 492–524.
[73] The development of the space–time view of quantum electrodynamics. In *Les Prix Nobel 1965.* Stockholm: Imprimerie Royale P.A. Norstedt & Soner, pp. 172–191.

1967

[74] Field theory as a guide to the strong interactions. In *Proc. 1967 Int. Conf. on Particles and Fields*, edited by C.R. Hagen, G. Guralnik, and V.S. Mathur. New York: Interscience.

1969

[75] Very high-energy collisions of hadrons. *Phys. Rev. Lett.* **23**: 1415–1417.

[76] The behavior of hadron collisions at extreme energies. In *High Energy Collisions.* London: Gordon and Breach, pp. 237–256.
[77] Present status of strong, electromagnetic, and weak interactions. *Ceskoslovensky Casopis pro Fysiku* **A19**: 47–59.
[78] With R.B. Leighton and R.E. Vogt. *Exercises in Introductory Physics.* Reading, Massachusetts: Addison-Wesley.
[79] What is science? *The Physics Teacher* **9/69**: 313–320.
[80] The application of mathematics to mathematics. *American Mathematics Monthly* **76**: 1178–1179.

1970

[81] With S. Pakvasa and S.F. Tuan. Some comments on baryonic states. *Phys. Rev.* **D2**: 1267–1270.
[82] With K.K. Thorber. Velocity acquired by an electron in a finite electric field in a polar crystal. *Phys. Rev.* **B1**: 4099–4114. *Erratum*: *Phys. Rev.* **B4**: 674.
[83] Partons. In *The Past Decade in Particle Theory.* London: Gordon and Breach, pp. 773–813. Eds. E.C.G. Sudarshan and Y. Ne'eman.

1971

[84] With M. Kislinger and F. Ravndal. Current matrix elements from a relativistic quark model. *Phys. Rev.* **D3**: 2706–2732.
[85] Closed loop and tree diagrams, in *Magic Without Magic: John Archibald Wheeler*, edited by John R. Klauder. New York: W.H. Freeman.
[86] Problems in quantizing the gravitational field and the massless Yang–Mills field. In *Magic Without Magic: John Archibald Wheeler*, edited by John R. Klauder. New York: W.H. Freeman.
[87] *Lectures on Gravitation.* Pasadena: CIT.

1972

[88] *Statistical Mechanics*; a set of lectures. Reading, Massachusetts: W.A. Benjamin.
[89] *Photon–Hadron Interactions.* Reading, Massachusetts: W.A. Benjamin.
[90] What neutrinos can tell us about partons. In *Proc. Neutrino '72 Europhysics Conference*, Vol. II. Budapest: OMKD Technoinform, pp. 75–96.
[91] *Fisica de Altas Energias: Cursos de Verano 1972.* Mexico: COPAA-SEDICT.

1973

[92] Quarks. *Fizikai Szemle* **23**: 1–7.

1974

[93] Take the world from another point of view. *Engineering and Science* **37**: 11–13 (Feb.).
[94] Cargo cult science. *Engineering and Science* **37**: 10–13 (June).
[95] Partons. In *Proc. 5th Hawaii Topical Conference in Particle Physics.* Honolulu: Univ. Press of Hawaii, pp. 1–97.

[96] Structure of the proton. *Science* **183**: 601–610.

1976

[97] Los Alamos from below. *Engineering and Science* **39**: 11–30 (Jan.–Feb.).

1977

[98] With R.D. Field and Geoffrey Fox. Correlations among particles and jets produced with large transverse momenta. *Nucl. Phys.* **B128**: 1–65.
[99] With R.D. Field. Quark elastic scattering as a source of high-transverse-momentum mesons. *Phys. Rev.* **D15**: 2590–2616.
[100] Correlations in hadron collisions at high transverse momentum. In *Orbis Scientiae* (Univ. of Miami) and *Deeper Pathways in High Energy Physics*. New York: Plenum, pp. 461–488.
[101] Quark jets. In *Proc. Int. Symp. on Multiparticle Dynamics 8th*, Kayserberg, June 1977. Strasbourg.
[102] Gauge theories. In *Weak and Electromagnetic Interactions at High Energy*. Amsterdam: North-Holland, pp. 121–204.

1978

[103] With R.D. Field and Geoffrey Fox. A quantum-chromodynamic approach for the large-transverse-momentum production of particles and jets. *Phys. Rev.* **D18**: 3320–3343.
[104] With R.D. Field. A parameterization of the properties of quark jets. *Nucl. Phys.* **B136**: 1–76.

1981

[105] The qualitative behavior of Yang–Mills theory in 2+1 dimensions. *Nucl. Phys.* **B188**: 479–512.

1982

[106] Simulating physics with computers. *Int. J. Theor. Phys.* **B2**: 6–7.
[107] A qualitative discussion of quantum chromodynamics in 2+1 dimensions. In *Proc. Int. Conf. on High Energy Physics*, Lisbon, July 9–15, 1981. Geneva: European Physical Society, pp. 660–683.
[108] Partons. In *Hawaii Topical Conference in Particle Physics*, Vol. I: Selected Lectures. Singapore: World Scientific, pp. 229–424.

1984

[109] With Ralph Leighton. The dignified professor. *Engineering and Science* **48**: 4–10 (Nov.).

1985

[110] Quantum mechanical computers. *Opt. News* **11**: 11–46.
[111] *QED: The Strange Theory of Light and Matter*. Princeton: Princeton University Press. In *Alix A. Mautner Memorial Lectures.*

[112] With Ralph Leighton and Edward Hutchings. *"Surely You're Joking, Mr. Feynman": Adventures of a Curious Character.* New York: W.W. Norton.
[113] The computing machines in the future. Nishina Memorial Lecture, Nishina Foundation and Gakushuin.

1986

[114] With H. Kleinert. Effective classical partition functions. *Phys. Rev.* **A34**: 5080–5084.
[115] Quantum mechanical computers. *Found. Phys.* **16**: 507–531.

1987

[116] With S. Weinberg. *Elementary Particles and the Laws of Physics: The 1986 Dirac Memorial Lectures.* New York: Cambridge University Press.
[117] Mr. Feynman goes to Washington. *Engineering and Science* **51**: 6–22 (Fall).
[118] Was ist Naturwissenschaft? *Physik und Didaktik* **2**: 105–116.
[119] Negative probability. In *Quantum Implications: Essays in Honor of David Bohm.* London: Routledge and Kegan Paul, pp. 235–248.

1988

[120] Difficulties in applying the variational principle to quantum field theories. In *Proc. Int. Workshop on Variational Calculus in Quantum Field Theory*, Wangerooge, West Germany, Sept. 1–4, 1987. Singapore: World Scientific.
[121] With Ralph Leighton. *"What Do You Care What Other People Think?": Further Adventures of a Curious Character.* New York: W.W. Norton.
[122] An outsider's inside view of the *Challenger* inquiry. *Physics Today* **41**(2): 26–37.

1995

[123] *Feynman Lectures on Gravitation*, edited by Brian Hatfield; notes taken by Fernando B. Morinigo and William G. Wagner. Reading, Massachusetts: Addison-Wesley.

1996

[124] *Feynman Lectures on Computation*, edited by Anthomy J.G. Hey and Robin W. Allen. Reading, Massachusetts: Addison-Wesley.

1999

[125] *The Pleasure of Finding Things Out.* Reading, Massachusetts: Perseus.