ADVANCED QUANTUM MECHANICS

FREEMAN DYSON

ADVANCED QUANTUM MECHANICS

This page intentionally left blank

ADVANCED QUANTUM MECHANICS

FREEMAN DYSON

Institute for Advanced Study, Princeton, USA

TRANSCRIBED BY

DAVID DERBES

LABORATORY SCHOOLS, UNIVERSITY OF CHICAGO, USA

ADVANCED QUANTUM MECHANICS

Preface

Both Kaiser's admirable *Drawing Theories Apart* [8] and Schweber's masterful *QED and the Men Who Made It* [7] refer frequently to the famous lectures on quantum electrodynamics given by Freeman Dyson at Cornell University in 1951. Two generations ago, graduate students (and their professors) wishing to learn the new techniques of QED passed around copies of Dyson's Cornell lecture notes, then the best and fullest treatment available. Textbooks appeared a few years later, e.g. by Jauch & Rohrlich [25] and Schweber [6], but interest in Dyson's notes has never fallen to zero. Here is what the noted theorist E. T. Jaynes wrote in an unpublished article [26] on Dyson's autobiographical *Disturbing the Universe*, 1984:

> But Dyson's 1951 Cornell course notes on Quantum Electrodynamics were the original basis of the teaching I have done since. For a generation of physicists they were the happy medium: clearer and better motivated than Feynman, and getting to the point faster than Schwinger. All the textbooks that have appeared since have not made them obsolete. Of course, this is to be expected since Dyson is probably, to this day, best known among the physicists as the man who first explained the unity of the Schwinger and Feynman approaches.

As a graduate student in Nicholas Kemmer's department of theoretical physics (Edinburgh, Scotland) I had heard vaguely about Dyson's lectures (either from Kemmer or from my advisor, Peter Higgs) and had read his classic papers [27], [28] in Schwinger's collection [4]. It never occurred to me to ask Kemmer for a copy of Dyson's lectures which he almost certainly had.

My interest in the legendary notes was revived thirty years later by the Kaiser and Schweber books. Within a few minutes Google led to scans of the notes [29] at the Dibner Archive (History of Recent Science & Technology) at MIT, maintained by Karl Hall, a historian at the Central European University in Budapest, Hungary. He had gotten permission from Dyson to post scanned images of the Cornell notes. Through the efforts of Hall, Schweber and Babak Ashrafi these were uploaded to the Dibner Archive. To obtain a paper copy would require downloading almost two hundred images, expensive in time and storage. Was there a text version? Had anyone re-typed the notes? Hall did not know, nor did further searching turn anything up. I volunteered to do the job. Hall thought this a worthwhile project, as did Dyson, who sent me a copy of the second edition, edited by Michael J. Moravcsik. (This copy had originally belonged to Sam Schweber.) Dyson suggested that the second edition be retyped, not the first. Nearly all of the differences between the two editions are Moravcsik's glosses on many calculations; there is essentially no difference in text, and (*modulo* typos) all the labeled equations are identical.

Between this typed version and Moravcsik's second edition there are few differences; all are described in the added notes. (I have also added references and an index.) About half are corrections of typographical errors. Missing words or sentences have been restored by comparison with the first edition; very infrequently a word or phrase has been deleted. A few changes have been made in notation. Intermediate steps in two calculations have been corrected but change nothing. Some notes point to articles or books. No doubt new errors have been introduced. Corrections will be welcomed! The young physicists will want familiar terms and notation, occasionally changed from 1951; the historians want *no* alterations. It was not easy to find the middle ground.

I scarcely knew LATEX before beginning this project. My friend (and Princeton '74 classmate) Robert Jantzen was enormously helpful, very generous with his time and his extensive knowledge of LATEX. Thanks, Bob. Thanks, too, to Richard Koch, Gerben Wierda and their colleagues, who have made LATEX so easy on a Macintosh. George Grätzer's textbook *Math into LATEX* was never far from the keyboard. No one who types technical material should be ignorant of LATEX.

This project would never have been undertaken without the approval of Prof. Dyson and the efforts of Profs. Hall, Schweber and Ashrafi, who made the notes accessible. I thank Prof. Hall for his steady encouragement

through the many hours of typing. I thank Prof. Dyson both for friendly assistance and for allowing his wonderful lectures to become easier to obtain, to be read with pleasure and with profit for many years to come.

Originally, the typed version was meant to serve as an adjunct to Karl Hall's scanned images at the Dibner site. Bob Jantzen, a relativist active in research, insisted that it also go up at the electronic physics preprint site arXiv.org, and after a substantial amount of work by him, this was arranged. A few weeks later the alert and hardworking team at World Scientific[1] got in touch with Prof. Dyson, to ask if he would allow them to publish his notes. He was agreeable, but told them to talk to me. I was delighted, but did not see how I could in good conscience profit from Prof. Dyson's work, and suggested that my share be donated to the New Orleans Public Library, now struggling to reopen after the disaster of Hurricane Katrina. Prof. Dyson agreed at once to this proposal. I am very grateful to him for his contribution to the restoration of my home town.

<div align="right">

David Derbes
Laboratory Schools
University of Chicago
loki@uchicago.edu
11 July 2006

</div>

---

[1] World Scientific is very grateful to Professor Freeman Dyson and Dr David Derbes for this magnificent manuscript.

Contents

Generally Used Notation

$$
\begin{aligned}
A^* \quad &= \quad \text{complex conjugate transposed (Hermitian conjugate)} \\
A^+ \quad &= \quad \text{complex conjugate (not transposed)} \\
\bar{A} \quad &= \quad A^*\beta = A^*\gamma_4 = \text{adjoint} \\
A^{-1} \quad &= \quad \text{inverse} \\
A^{\mathrm{T}} \quad &= \quad \text{transposed} \\
\mathbb{I} \quad &= \quad \text{identity matrix or operator} \\
\operatorname{Tr} A \quad &= \quad \text{trace of matrix } A \text{ (sum of all diagonal elements)} \\
\not{a} \quad &= \quad \sum_\mu a_\mu \gamma_\mu \text{ (This slash notation is also colorfully known as the} \\
&\quad\quad \text{Feynman dagger.)}
\end{aligned}
$$

CHAPTER 1

Introduction

1.1 Books

W. Pauli, "*Die Allgemeinen Principien der Wellenmechanik*"; *Handbuch der Physik*, 2 ed., Vol. 24,

Part 1; Edwards reprint, Ann Arbor 1947. (In German) [1]

W. Heitler, *Quantum Theory of Radiation*, 2nd Edition, Oxford. 3rd edition just published. [2]

G. Wentzel, *Introduction to the Quantum Theory of Wave-Fields*, Inter-science, N.Y. 1949 [3]

I shall not expect you to have read any of these, but I shall refer to them as we go along. The later part of the course will be new stuff, taken from papers of Feynman and Schwinger mainly. [4], [5], [6], [7], [8]

1.2 Subject Matter

You have had a complete course in non-relativistic quantum theory. I assume this known. All the general principles of the non-relativistic theory are valid and true under all circumstances, in particular also when the system happens to be relativistic. What you have learned is therefore still good.

You have had a course in classical mechanics and electrodynamics including special relativity. You know what is meant by a system being relativistic; the equations of motion are formally invariant under Lorentz transformations. General relativity we shall not touch.

This course will be concerned with the development of a *Lorentz–invariant quantum theory*. That is not a general dynamical method like the non-relativistic quantum theory, applicable to all systems. We cannot

yet devise a general method of that kind, and it is probably impossible. Instead we have to find out what are the possible systems, the particular equations of motion, which can be handled by the non-relativistic quantum dynamics and which are at the same time Lorentz–invariant.

In the non-relativistic theory it was found that almost any classical system could be handled, i.e. quantized. Now on the contrary we find there are very few possibilities for a relativistic quantized system. This is a most important fact. It means that starting only from the principles of relativity and quantization, it is mathematically possible only for very special types of objects to exist. So one can *predict* mathematically some important things about the real world. The most striking examples of this are:

(i) Dirac from a study of the electron predicted the positron, which was later discovered [9].

(ii) Yukawa from a study of nuclear forces predicted the meson, which was later discovered [10].

These two examples are special cases of the general principle, which is the basic success of the relativistic quantum theory, that *A Relativistic Quantum Theory of a Finite Number of Particles is Impossible.* A relativistic quantum theory necessarily contains these features: an indefinite number of particles of one or more types, particles of each type being identical and indistinguishable from each other, possibility of creation and annihilation of particles.

Thus the two principles of relativity and quantum theory when combined lead to a world built up out of various types of elementary particles, and so make us feel quite confident that we are on the right way to an understanding of the real world. In addition, various detailed properties of the observed particles are necessary consequences of the general theory. These are for example:

(i) Magnetic moment of Electron (Dirac) [9].

(ii) Relation between spin and statistics (Pauli) [11].

1.3   Detailed Program

We shall not develop straightaway a correct theory including many particles. Instead we follow the historical development. We try to make a relativistic quantum theory of *one* particle, find out how far we can go and where we get into trouble. Then we shall see how to change the theory and get over the

trouble by introducing many particles. Incidentally, the one–particle theories are quite useful, being correct to a good approximation in many situations where creation of new particles does not occur, and where something better than a non-relativistic approximation is needed. An example is the Dirac theory of the Hydrogen atom.[1]

The non-relativistic theory gave levels correctly but no fine-structure. (Accuracy of one part in 10,000). The Dirac one-particle theory gives all the main features of the fine-structure correctly, number of components and separations good to 10% but not better. (Accuracy one part in 100,000).

The Dirac many-particle theory gives the fine-structure separations (Lamb experiment) correctly to about one part in 10,000. (Overall accuracy 1 in 10^8.)

Probably to get accuracy better than 1 in 10^8 even the Dirac many-particle theory is not enough and one will need to take all kinds of meson effects into account which are not yet treated properly. Experiments are so far only good to about 1 in 10^8.

In this course I will go through the one-particle theories first in detail. Then I will talk about their breaking down. At that point I will make a fresh start and discuss how one can make a relativistic quantum theory in general, using the new methods of Feynman and Schwinger. From this we shall be led to the many-particle theories. I will talk about the general features of these theories. Then I will take the special example of quantum electrodynamics and get as far as I can with it before the end of the course.

1.4   One-Particle Theories

Take the simplest case, one particle with no forces. Then the non-relativistic wave-mechanics tells you to take the equation $E = \frac{1}{2m}p^2$ of classical mechanics, and write

$$E \rightarrow i\hbar\frac{\partial}{\partial t} \qquad p_x \rightarrow -i\hbar\frac{\partial}{\partial x} \tag{1}$$

to get the wave-equation[2]

$$i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi = -\frac{\hbar^2}{2m}\nabla^2\psi \tag{2}$$

satisfied by the wave-function ψ.

To give a physical meaning to ψ, we state that $\rho = \psi^*\psi$ is the probability of finding the particle at the point x, y, z at time t. And the probability is

conserved because[3]

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{\jmath} = 0 \tag{3}$$

where

$$\vec{\jmath} = \frac{\hbar}{2mi}(\psi^* \nabla \psi - \psi \nabla \psi^*) \tag{4}$$

where ψ^* is the complex conjugate of ψ.

Now do this relativistically. We have classically

$$E^2 = m^2 c^4 + c^2 p^2 \tag{5}$$

which gives the wave equation

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\psi = \nabla^2 \psi - \frac{m^2 c^2}{\hbar^2}\psi. \tag{6}$$

This is an historic equation, the Klein–Gordon equation. Schrödinger already in 1926 tried to make a relativistic quantum theory out of it. But he failed, and many other people too, until Pauli and Weisskopf gave the many-particle theory in 1934 [12]. Why?

Because in order to interpret the wave-function as a probability we must have a continuity equation. This can only be got out of the wave-equation if we take $\vec{\jmath}$ as before, and

$$\rho = \frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial \psi}{\partial t} - \frac{\partial \psi^*}{\partial t}\psi\right). \tag{7}$$

But now since the equation is second order, ψ and $\frac{\partial \psi}{\partial t}$ are arbitrary. Hence ρ need not be positive. We have *negative probabilities*. This defeated all attempts to make a sensible one-particle theory.

The theory can be carried through quite easily, if we make ψ describe an assembly of particles of both positive and negative charge, and ρ is the *net* charge density at any point. This is what Pauli and Weisskopf did, and the theory you get is correct for π-mesons, the mesons which are made in the synchrotron downstairs. I will talk about it later.

The Dirac Theory

2.1   The Form of the Dirac Equation

Historically before the relativistic quantum theory came the one-particle theory of Dirac. This was so successful in dealing with the electron, that it was for many years the only respectable relativistic quantum theory in existence. And its difficulties are a lot less immediate than the difficulties of the one-particle Klein–Gordon theory.

Dirac said, suppose the particle can exist in several distinct states with the same momentum (different orientations of spin.) Then the wave-function ψ satisfying (6) must have several *components*; it is not a scalar but a set of numbers each giving the probability amplitude to find the particle at a given place and in a given substate. So we write for ψ a column matrix

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \qquad \text{for the components } \psi_\alpha \, ; \qquad \alpha = 1, 2, \ldots$$

Dirac assumed that the probability density at any point is still given by

$$\rho = \sum_\alpha \psi_\alpha^* \psi_\alpha \tag{8}$$

which we write

$$\rho = \psi^* \psi$$

as in the non-relativistic theory. Here ψ^* is a *row* matrix

$$[\psi_1^*, \psi_2^*, \ldots]$$

We must have (3) still satisfied. So ψ must satisfy a wave-equation of *First Order* in t. But since the equations are relativistic, the equation has to be also of first order in x, y, z. Thus the most general possible wave-equation is

$$\frac{1}{c}\frac{\partial\psi}{\partial t} + \sum_{1}^{3}\alpha^k\frac{\partial\psi}{\partial x_k} + i\frac{mc}{\hbar}\beta\psi = 0 \tag{9}$$

where x_1, x_2, x_3 are written for x, y, z and α^1, α^2, α^3, β are square matrices whose elements are numbers. The conjugate of (9) gives

$$\frac{1}{c}\frac{\partial\psi^*}{\partial t} + \sum_{1}^{3}\frac{\partial\psi^*}{\partial x_k}\alpha^{k*} - i\frac{mc}{\hbar}\psi^*\beta^* = 0 \tag{10}$$

where α^{k*} and β^* are Hermitian conjugates.

Now to get (3) out of (8), (9) and (10) we must have $\alpha^{k*} = \alpha^k$, $\beta^* = \beta$ so α^k and β are *Hermitian*; and

$$j_k = c(\psi^*\alpha^k\psi) \tag{11}$$

Next what more do we want from equation (9)? Two things. (A) it must be consistent with the second order equation (6) we started from; (B) the whole theory must be Lorentz invariant.

First consider (A). If (9) is consistent with (6) it must be possible to get exactly (6) by multiplying (9) by the operator

$$\frac{1}{c}\frac{\partial}{\partial t} - \sum_{1}^{3}\alpha^\ell\frac{\partial}{\partial x_\ell} - i\frac{mc}{\hbar}\beta \tag{12}$$

chosen so that the terms with mixed derivatives $\frac{\partial}{\partial t}$, $\frac{\partial}{\partial x_k}$ and $\frac{\partial}{\partial t}$ cancel. This gives

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \sum\sum_{k\neq\ell}\frac{1}{2}(\alpha^k\alpha^\ell + \alpha^\ell\alpha^k)\frac{\partial^2\psi}{\partial x_k\partial x_\ell} + \sum_k\alpha_k^2\frac{\partial^2\psi}{\partial x_k^2}$$

$$- \frac{m^2c^2}{\hbar^2}\beta^2\psi + i\frac{mc}{\hbar}\sum_k(\alpha^k\beta + \beta\alpha^k)\frac{\partial\psi}{\partial x_k}$$

This agrees with (6) if and only if

$$\begin{aligned}
\alpha^k\alpha^\ell + \alpha^\ell\alpha^k &= 0 \quad k\neq\ell \\
\alpha^k\beta + \beta\alpha^k &= 0 \\
\alpha^{k2} &= \beta^2 = \mathbb{I}, \text{(identity matrix)}
\end{aligned} \tag{13}$$

Thus we could not possibly *factorize* the 2nd order equation into two first-order operators involving ordinary numbers. But we can do it with *matrices*.

Consider the Pauli spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (14)$$

you are familiar with. They satisfy

$$\sigma_k \sigma_\ell + \sigma_\ell \sigma_k = 2\delta_{\ell k}$$

But we cannot make 4 matrices of this type all anti-commuting. They must be *at least* 4×4.

One possible set of α^k and β is

$$\alpha^k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \qquad \beta = \begin{pmatrix} \begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} & 0 \\ 0 & \begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix} \end{pmatrix} \qquad (15)$$

In particular

$$\alpha^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad \alpha^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$

$$\alpha^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

These are hermitian as required. Of course if α^k and β are any set satisfying (13) then $S\alpha^k S^{-1}$ and $S\beta S^{-1}$ are another set, where S is any *unitary* matrix $SS^* = 1$. And conversely it can be proved that every possible 4×4 matrices α^k and β are of this form with *some* such matrix S. We do not prove this here.

The Dirac equation is thus a set of 4 simultaneous linear partial differential[4] equations in the four functions ψ_α.

2.2 Lorentz Invariance of the Dirac Equation

What does this mean? Consider a general Lorentz transformation: If x'_μ are the new coordinates:

$$x'_\mu = \sum_{\nu=0}^{3} a_{\mu\nu} x_\nu \qquad (x_o = ct) \qquad (16)$$

In the new coordinate system the wave-function will be ψ'. Clearly we do not expect that $\psi' = \psi$. Example: in the Maxwell theory which is relativistic, the magnetic field H is no longer a pure magnetic field in a moving system. Instead it transforms like a tensor. So we have to find *some* transformation law for the ψ which will leave invariant the physical consequences of the equations.

We need in fact two things: (i) the interpretation of $\psi^*\psi$ as a probability density must be preserved, (ii) the validity of the Dirac equation must be preserved in the new system.

First consider (i). The quantity which can be directly observed and must be invariant is the quantity

$$(\psi^*\psi) \times V$$

where V is a volume. Now in going to a new Lorentz system with relative velocity v the volume V changes by Fitzgerald contraction to the value

$$V' = V\sqrt{1 - \frac{v^2}{c^2}}$$

Therefore

$$(\psi^{*\prime}\psi') = \frac{\psi^*\psi}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{17}$$

and so $(\psi^*\psi) = \rho$ transforms like an *energy*, i.e. like the fourth component of a vector. This shows incidentally that $\psi' \neq \psi$. Since ρ and $\vec{j}$ are related by the equation of continuity, the space-components of the 4-vector are

$$(S_1, S_2, S_3) = \psi^*\alpha^k\psi = \frac{1}{c}j_k \tag{18}$$

So we require that the 4 quantities

$$(S_1, S_2, S_3, S_0) = (\psi^*\alpha^k\psi, \psi^*\psi) \tag{19}$$

transform like a *4-vector*. This will be enough to preserve the interpretation of the theory.

Assume that

$$\psi' = S\psi \tag{20}$$

where S is a *linear* operator. Then

$$\psi'^* = \psi^*S^* \tag{21}$$

So we require

$$\psi^{*\prime}\alpha^k\psi' = \psi^* S^* \alpha^k S\psi = \sum_{\nu=0}^{3} a_{k\nu}\psi^*\alpha^\nu\psi$$

$$\psi^{*\prime}\psi' = \psi^* S^* S\psi = \sum_{\nu=0}^{3} a_{0\nu}\psi^*\alpha^\nu\psi \tag{22}$$

writing $\alpha^0 = \mathbb{I}$.

Thus we need

$$S^*\alpha^\mu S = \sum_{\nu=0}^{3} a_{\mu\nu}\alpha^\nu , \qquad \mu = 0, 1, 2, 3 \tag{23}$$

Next consider (ii). The Dirac equation for ψ' is

$$\sum_{0}^{3} \alpha^\nu \frac{\partial}{\partial x'_\nu}\psi' + i\frac{mc}{\hbar}\beta\psi' = 0 \tag{24}$$

Now the original Dirac equation for ψ expressed in terms of the new coordinates is

$$\sum_{\mu=0}^{3}\sum_{\nu=0}^{3} \alpha^\mu \frac{\partial}{\partial x'_\nu}a_{\nu\mu}S^{-1}\psi' + i\frac{mc}{\hbar}\beta S^{-1}\psi' = 0 \tag{25}$$

The sets of equations (24) and (25) have to be equivalent, not identical. Thus (25) must be the same as (24) multiplied by $\beta S^{-1}\beta$. The condition for this is

$$\beta S^{-1}\beta\alpha^\nu = \sum_{0}^{3} \alpha^\lambda a_{\nu\lambda}S^{-1} \tag{26}$$

But (23) and (26) are identical if

$$\beta S^{-1}\beta = S^* \quad \text{which means} \quad S^*\beta S = \beta \tag{27}$$

Thus β transforms like a scalar, α^ν like a 4-vector when multiplied by S^*S.

2.3   To Find the S

Given two coordinate transformations in succession, with matrices already found, the combined transformation will correspond to the product of these

matrices. Hence we have to consider only 3 simple types of transformation.

(1) Pure rotations
$$x'_0 = x_0 \qquad\qquad\qquad\qquad x'_3 = x_3$$
$$x'_1 = x_1 \cos\theta + x_2 \sin\theta$$
$$x'_2 = -x_1 \sin\theta + x_2 \cos\theta$$

(2) Pure Lorentz transformations
$$x'_1 = x_1 \qquad\qquad\qquad\qquad x'_2 = x_2$$
$$x'_3 = x_3 \cosh\theta + x_0 \sinh\theta$$
$$x'_0 = x_3 \sinh\theta + x_0 \cosh\theta$$

(3) Pure reflections
$$x'_1 = -x_1 \qquad x'_2 = -x_2 \qquad x'_3 = -x_3 \qquad x'_0 = x_0$$

Case 1. Then
$$S = \cos\frac{1}{2}\theta + i\sigma_3 \sin\frac{1}{2}\theta \tag{28}$$

Here
$$\sigma_3 = \begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}$$

commutes with α_3 and β.

$$\sigma_3\alpha_1 = i\alpha_2, \qquad \sigma_3\alpha_2 = -i\alpha_1$$
$$S^* = \cos\frac{1}{2}\theta - i\sigma_3 \sin\frac{1}{2}\theta$$

Then
$$S^*\beta S = \beta$$
$$S^*\alpha^0 S = \alpha^0$$
$$S^*\alpha^3 S = \alpha^3$$

as required.
$$S^*\alpha^1 S = \cos\theta\alpha^1 + \sin\theta\alpha^2$$
$$S^*\alpha^2 S = -\sin\theta\alpha^1 + \cos\theta\alpha^2$$

Case 2.
$$S = S^* = \cosh\frac{1}{2}\theta + \alpha_3 \sinh\frac{1}{2}\theta \tag{29}$$

Here

$$S^* \beta S = \beta$$
$$S^* \alpha^1 S = \alpha^1$$
$$S^* \alpha^2 S = \alpha^2$$
$$S^* \alpha^3 S = \cosh \theta \alpha^3 + \sinh \theta \alpha^0$$
$$S^* \alpha^0 S = \sinh \theta \alpha^3 + \cos \theta \alpha^0$$

Case 3.

$$S = S^* = \beta \tag{30}$$

Note that in all cases S is ambiguous by a factor ± 1. So in case 1 a rotation though $360°$ gives $S = -1$.

Problem 1. Find the S corresponding to a general infinitesimal coordinate transformation. Compare and show that it agrees with the exact solutions given here.

The ψ_α's transforming with these S-transformations are called *spinors*. They are a direct extension of the non-relativistic 2-component spin-functions. Mathematical theory of spinors is not very useful. In fact we find always in practice, calculations can be done most easily if one avoids any explicit representation of the spinors. *Use only formal algebra and commutation relations of the matrices.*

2.4   The Covariant Notation

In order to avoid distinction between covariant and contravariant vectors (which we have also unjustifiably ignored in the previous discussion) it is useful to use the imaginary 4th coordinate

$$x_4 = i x_0 = ict \tag{31}$$

In this coordinate system the four matrices[5]

$$\gamma_{1,2,3,4} = (-i\beta\alpha^{1,2,3}, \beta) \quad \text{i.e.} \tag{32}$$

$$\gamma_1 = \begin{pmatrix} 0 & \begin{smallmatrix} 0 & -i \\ -i & 0 \end{smallmatrix} \\ \begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix} & 0 \end{pmatrix} \; ; \qquad \gamma_2 = \begin{pmatrix} 0 & \begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix} \\ \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & 0 \end{pmatrix} \; ;$$

$$\gamma_3 = \begin{pmatrix} 0 & \begin{smallmatrix} -i & 0 \\ 0 & i \end{smallmatrix} \\ \begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix} & 0 \end{pmatrix} \; ; \qquad \gamma_4 = \begin{pmatrix} \begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} & 0 \\ 0 & \begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix} \end{pmatrix}$$

are a 4-vector. They are all Hermitian and satisfy

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\delta_{\mu\nu} \tag{33}$$

The Dirac equation and its conjugate may now be written

$$\sum_1^4 \gamma_\mu \frac{\partial \psi}{\partial x_\mu} + \frac{mc}{\hbar}\psi = 0$$

$$\sum_1^4 \frac{\partial \overline{\psi}}{\partial x_\mu}\gamma_\mu - \frac{mc}{\hbar}\overline{\psi} = 0 \tag{34}$$

with

$$\overline{\psi} = \psi^* \beta \qquad \text{and} \tag{35}$$

$$s_\mu = i\left(\overline{\psi}\,\gamma_\mu\,\psi\right) = \left(\frac{1}{c}\,\boldsymbol{j}, i\rho\right) \tag{36}$$

These notations are the most convenient for calculations.

2.5 Conservation Laws. Existence of Spin

The Hamiltonian in this theory is[6]

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \tag{37}$$

$$H = -i\hbar c \sum_1^3 \alpha^k \frac{\partial}{\partial x_k} + mc^2\beta = -i\hbar c\,\boldsymbol{\alpha}\cdot\nabla + mc^2\beta \tag{38}$$

This commutes with the momentum $\boldsymbol{p} = -i\hbar\nabla$. So the momentum $\boldsymbol{p}$ is a constant of motion.

However the angular momentum operator

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p} = -i\hbar \boldsymbol{r} \times \nabla \tag{39}$$

is not a constant. For
$$[H, \boldsymbol{L}] = -\hbar^2 c \, \boldsymbol{\alpha} \times \nabla \tag{40}$$

But
$$[H, \boldsymbol{\sigma}] = -i\hbar c \, \nabla \cdot [\boldsymbol{\alpha}, \boldsymbol{\sigma}] \qquad \text{where} \qquad \boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$$

while
$$\left[\alpha^1, \sigma_1\right] = 0, \qquad \left[\alpha^1, \sigma_2\right] = 2i\alpha^3, \qquad \left[\alpha^1, \sigma_3\right] = -2i\alpha^2, \quad \text{etc.}$$

So
$$[H, \sigma_3] = 2\hbar c \left(\alpha^1 \nabla_2 - \alpha^2 \nabla_1\right) \quad \text{and thus}$$
$$[H, \boldsymbol{\sigma}] = 2\hbar c \, \boldsymbol{\alpha} \times \nabla \tag{41}$$

Thus
$$\boldsymbol{L} + \tfrac{1}{2}\hbar\boldsymbol{\sigma} = \hbar\boldsymbol{J} \tag{42}$$

is a constant, the total angular momentum, because by (40), (41) and (42)

$$[H, \boldsymbol{J}] = 0$$

$\boldsymbol{L}$ is the orbital angular momentum and $\tfrac{1}{2}\hbar\boldsymbol{\sigma}$ the spin angular momentum
This agrees with the non-relativistic theory. But in that theory the spin and
L of a free particle were *separately* constant. This is no longer the case.

When a central force potential $V(r)$ is added to H, the operator $\boldsymbol{J}$ still
is constant.

2.6   Elementary Solutions

For a particle with a particular momentum $\boldsymbol{p}$ and energy E, the wave func-
tion will be
$$\psi(x, t) = u \exp\left(i\frac{\boldsymbol{p} \cdot \boldsymbol{x}}{\hbar} - i\frac{Et}{\hbar}\right) \tag{43}$$

where u is a constant spinor. The Dirac equation then becomes an equation
for u only
$$Eu = \left(c\boldsymbol{\alpha} \cdot \boldsymbol{p} + mc^2\beta\right)u \tag{44}$$

We write now
$$p_+ = p_1 + ip_2 \qquad p_- = p_1 - ip_2 \tag{45}$$

Then (44) written out in full becomes

$$
\begin{aligned}
\left(E - mc^2\right) u_1 &= c \left(p_3 u_3 + p_- u_4\right) \\
\left(E - mc^2\right) u_2 &= c \left(p_+ u_3 - p_3 u_4\right) \\
\left(E + mc^2\right) u_3 &= c \left(p_3 u_1 + p_- u_2\right) \\
\left(E + mc^2\right) u_4 &= c \left(p_+ u_1 - p_3 u_2\right)
\end{aligned}
\tag{46}
$$

These 4 equations determine u_3 and u_4 given u_1 and u_2, or vice-versa. And either u_1 and u_2, or u_3 and u_4, can be chosen arbitrarily provided that[7]

$$
E^2 = m^2 c^4 + c^2 p^2
\tag{47}
$$

Thus given p and $E = +\sqrt{m^2 c^4 + c^2 p^2}$, there are two independent solutions of (46); these are, in non-normalized form:

$$
\begin{pmatrix}
1 \\
0 \\
\dfrac{c p_3}{E + mc^2} \\
\dfrac{c p_+}{E + mc^2}
\end{pmatrix}
\qquad
\begin{pmatrix}
0 \\
1 \\
\dfrac{c p_-}{E + mc^2} \\
\dfrac{-c p_3}{E + mc^2}
\end{pmatrix}
\tag{48}
$$

This gives the two spin-states of an electron with given momentum, as required physically.

But there are also solutions with $E = -\sqrt{m^2 c^4 + c^2 p^2}$. In fact again two independent solutions, making 4 altogether. These are the famous *negative energy states*. Why cannot we simply agree to ignore these states, say they are physically absent? Because when fields are present the theory gives transitions from positive to negative states, e.g. the Hydrogen atom should decay to a negative state in 10^{-10} secs or less.

Certainly negative energy particles are not allowed physically. They can for example never be stopped by matter at rest, with every collision they move faster and faster. So Dirac was driven to . . .

2.7   The Hole Theory

All negative-energy states are normally filled by one electron each. Because of the exclusion principle transitions of ordinary electrons to these states are forbidden. If sometimes a negative energy state of momentum $-p$ energy

$-E$ is *empty*, this appears as a particle of momentum p energy $+E$, and the opposite charge to an electron, i.e. an ordinary positron.

Thus we are led at once to a many-particle theory in order to get sensible results. With spin-0 particles, to get positive probabilities. With spin-$\frac{1}{2}$ particles, to get positive energies.

The Dirac theory in its one-particle form cannot describe properly the interaction between several particles. But so long as we are talking only about free particles, we can describe them with one-particle wave-functions.

2.8   Positron States

So which wave-function will describe a positron with momentum p and energy E? Clearly the wave function should be of the form

$$\phi(x,t) = v \exp\left(i\frac{\boldsymbol{p}\cdot\boldsymbol{x}}{\hbar} - i\frac{Et}{\hbar}\right) \tag{49}$$

as always in quantum mechanics. But the negative-energy electron whose *absence is* the positron has a wave-function

$$\psi(x,t) = u \exp\left(-i\frac{\boldsymbol{p}\cdot\boldsymbol{x}}{\hbar} + i\frac{Et}{\hbar}\right) \tag{50}$$

since it has a momentum $-p$ energy $-E$.

Thus we must take

$$\phi = C\psi^+, \qquad \text{i.e.} \qquad v = Cu^+ \tag{51}$$

where ψ^+ is ψ with complex conjugate elements but *not* transposed, and C is a suitable constant matrix;

$$\psi^+(x,t) = u^+ \exp\left(i\frac{\boldsymbol{p}\cdot\boldsymbol{x}}{\hbar} - i\frac{Et}{\hbar}\right)$$

We know that u is a solution of

$$Eu = \left(c\boldsymbol{\alpha}\cdot\boldsymbol{p} - mc^2\beta\right)u \tag{52}$$

We want the theory to make no distinction between electrons and positrons, and so v must also satisfy the Dirac equation

$$Ev = \left(c\boldsymbol{\alpha}\cdot\boldsymbol{p} + mc^2\beta\right)v$$
$$ECu^+ = \left(c\boldsymbol{\alpha}\cdot\boldsymbol{p} + mc^2\beta\right)Cu^+ \tag{53}$$

But from (52) we have for u^+ the equation

$$Eu^+ = \left(c\boldsymbol{\alpha}^+ \cdot \boldsymbol{p} - mc^2\beta^+\right)u^+ \tag{54}$$

In order that (53) and (54) be identical we should have

$$C\alpha^{k+} = \alpha^k C, \qquad\qquad C\beta^+ = -\beta C \tag{55}$$

Now in fact

$$\alpha^{1+} = \alpha^1 \qquad \alpha^{3+} = \alpha^3 \qquad \alpha^{2+} = -\alpha^2 \qquad \beta^+ = \beta$$

Therefore a suitable C will be

$$C = -i\beta\alpha^2 = \gamma_2 = \begin{pmatrix} 0 & \begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix} \\ \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & 0 \end{pmatrix} \tag{56}$$

The relation between ψ and ϕ is symmetrical because

$$C^2 = \mathbb{I} \quad \text{Hence} \quad \psi = C\phi^+ \tag{57}$$

The ϕ is called the *charge-conjugate* wave-function corresponding to the negative-energy electron ψ. Clearly

$$\phi^*\phi = \left(C\psi^+\right)^* \left(C\psi^+\right) = \psi^T C^* C\psi^+ = \psi^* \left(C^*C\right)^T \psi = \psi^*\psi \tag{58}$$

And

$$\phi^*\alpha^k\phi = \psi^T C^*\alpha^k C\psi^+ = \psi^* C\alpha^{kT} C\psi = \psi^*\alpha^k\psi \tag{59}$$

Thus the probability and flow densities are the same for a positron as for the conjugate negative electron.

For many purposes it is easier to represent positrons directly by the $\overline{\psi}$ wave-function, e.g. in computing cross-sections for pair creation and so forth as we shall do later. But if you actually want to *see* the positron, e.g. in describing the details of a positronium experiment, it is necessary to use the ϕ wave-function to represent e.g. the way the spin is pointing.

This is all we shall say about free electrons and positrons.

2.9   Electromagnetic Properties of the Electron

Given an external (c-number) electromagnetic field defined by the potentials

$$A_\mu \qquad \mu = 1, 2, 3, 4 \qquad A_4 = i\Phi$$

given functions of space and time. Then the motion of a particle in the field is found by substituting in the free-particle Lagrangian

$$E + e\,\Phi \quad \text{for } E$$

$$\boldsymbol{p} + \frac{e}{c}\boldsymbol{A} \quad \text{for } \boldsymbol{p} \tag{60}$$

where $(-e)$ is the electron charge. We write the momentum-energy 4-vector

$$p = (p_1,\ p_2,\ p_3,\ p_4 = iE/c) \tag{61}$$

Then we have to substitute simply

$$p_\mu + \frac{e}{c}A_\mu \quad \text{for } p_\mu \tag{62}$$

Now in the quantum theory

$$p_\mu \rightarrow -i\hbar\frac{\partial}{\partial x_\mu} \tag{63}$$

Therefore the Dirac equation with fields is

$$\sum_1^4 \gamma_\mu \left(\frac{\partial}{\partial x_\mu} + \frac{ie}{\hbar c}A_\mu \right)\psi + \frac{mc}{\hbar}\,\psi = 0 \tag{64}$$

$$\sum_1^4 \left(\frac{\partial}{\partial x_\mu} - \frac{ie}{\hbar c}A_\mu \right)\overline{\psi}\,\gamma_\mu - \frac{mc}{\hbar}\,\overline{\psi} = 0 \tag{65}$$

In the non-covariant notations this is

$$i\hbar\frac{\partial\psi}{\partial t} = \left[-e\,\Phi + \sum_1^3 \left(-i\hbar c\frac{\partial}{\partial x_k} + eA_k \right)\alpha^k + mc^2\beta \right]\psi \tag{66}$$

since by (57), we have $\overline{\psi}\gamma_\mu = \psi^*\beta\gamma_\mu = \left(C\phi^+\right)^T\beta\gamma_\mu = \phi^T C^T\beta\gamma_\mu$; the wave function $\phi = C\psi^+$ of a positron satisfies by (65)

$$\sum \left(\frac{\partial}{\partial x_\mu} - \frac{ie}{\hbar c}A_\mu \right)\gamma_\mu^T\beta C\phi - \frac{mc}{\hbar}\beta C\phi = 0 \tag{67}$$

Multiplying by $C\beta$ this gives

$$\sum \left(\frac{\partial}{\partial x_\mu} - \frac{ie}{\hbar c}A_\mu \right)\gamma_\mu\phi + \frac{mc}{\hbar}\,\phi = 0 \tag{68}$$

This is exactly the Dirac equation for a particle of positive charge $(+e)$. We have used

$$C\beta\gamma_\mu^T\beta C = -\gamma_\mu, \tag{69}$$

which follows from (15), (32), and (55).

2.10   The Hydrogen Atom

This is the one problem which it is possible to treat very accurately using
the one-electron Dirac theory. The problem is to find the eigenstates of the
equation

$$E\psi = H\psi$$

$$H = -i\hbar c\,\boldsymbol{\alpha} \cdot \nabla + mc^2\beta - \frac{e^2}{r} \tag{70}$$

As in the non-relativistic theory, we have as quantum numbers in addition
to E itself the quantities

$$j_z = -i\,[\boldsymbol{r} \times \nabla]_3 + \tfrac{1}{2}\sigma_3 \tag{71}$$

$$j(j+1) = J^2 = \left[-i\,(\boldsymbol{r} \times \nabla) + \tfrac{1}{2}\boldsymbol{\sigma}\right]^2 \tag{72}$$

where j_z and j are now *half-odd integers* by the ordinary theory of angular
momenta. These quantum numbers are not enough to fix the state, because
each value of j may correspond to two non-relativistic states with $\ell = j \pm \tfrac{1}{2}$.
Therefore we need an additional operator which commutes with H, which
will distinguish between states with $\boldsymbol{\sigma}$ parallel or antiparallel to $\boldsymbol{J}$. The
obvious choice is

$$Q = \boldsymbol{\sigma} \cdot \boldsymbol{J}$$

But $[H, \boldsymbol{\sigma}]$ is non-zero and rather complicated. So it is better to try

$$Q = \beta\boldsymbol{\sigma} \cdot \boldsymbol{J} \tag{73}$$

which is the same in the non-relativistic limit.

Then we have

$$[H, Q] = [H, \beta\boldsymbol{\sigma} \cdot \boldsymbol{J}] = [H, \beta\boldsymbol{\sigma}] \cdot \boldsymbol{J} + \beta\boldsymbol{\sigma} \cdot [H, \boldsymbol{J}]$$

But $[H, \boldsymbol{J}] = 0$; furthermore, since

$$\alpha^k \beta\sigma_\ell = \beta\sigma_\ell \alpha^k \quad k \neq \ell \quad \text{and} \quad \alpha^k \beta\sigma_k = -\beta\sigma_k \alpha^k$$

we get

$$[H, \beta\boldsymbol{\sigma}] = -i\hbar c\,\{(\boldsymbol{\alpha} \cdot \nabla)\,\beta\boldsymbol{\sigma} - \beta\boldsymbol{\sigma}\,(\boldsymbol{\alpha} \cdot \nabla)\} = -2i\hbar c \sum_{k=1}^{3} \alpha^k \sigma_k\,\beta\,\nabla_k$$

Therefore

$$[H, \beta\boldsymbol{\sigma}] \cdot \boldsymbol{J} = -2\hbar c \sum_{k=1}^{3} \alpha^k \sigma_k \beta \nabla_k (\boldsymbol{r} \times \nabla)_k - i\hbar c (\boldsymbol{\alpha} \cdot \nabla) \beta$$

$$= -i\hbar c (\boldsymbol{\alpha} \cdot \nabla) \beta = \left[H, \tfrac{1}{2}\beta\right]$$

because

$$\nabla \cdot \boldsymbol{r} \times \nabla = 0 \quad \text{and} \quad \alpha^k \sigma_k = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{for all } k\text{'s.}$$

Hence the quantity which commutes with H and is a constant of the motion is

$$K = \beta\boldsymbol{\sigma} \cdot \boldsymbol{J} - \tfrac{1}{2}\beta \tag{74}$$

There must be a relation between K and J. In fact

$$K^2 = \left(\frac{\boldsymbol{\sigma} \cdot \boldsymbol{L}}{\hbar} + 1\right)^2 = \frac{L^2}{\hbar^2} + \frac{\boldsymbol{\sigma} \cdot \boldsymbol{L}}{\hbar} + 1$$

$$J^2 = \left(\frac{\boldsymbol{L}}{\hbar} + \frac{1}{2}\boldsymbol{\sigma}\right)^2 = \frac{L^2}{\hbar^2} + \frac{\boldsymbol{\sigma} \cdot \boldsymbol{L}}{\hbar} + \frac{3}{4}$$

Hence

$$K^2 = J^2 + \tfrac{1}{4} = \left(j + \frac{1}{2}\right)^2 \tag{75}$$

Therefore K has *integer* eigenvalues not zero,

$$K = k = \pm\left(j + \frac{1}{2}\right) \tag{76}$$

$$j = |k| - \frac{1}{2}, \quad k = \pm 1, \ \pm 2, \ \pm 3, \ \dots \tag{77}$$

Using the eigenvalue for K, we can simplify the Hamiltonian, which we could not do as in the non-relativistic theory with the eigenvalue of L^2 alone. First

$$\boldsymbol{\sigma} \cdot \boldsymbol{r}\, \boldsymbol{\sigma} \cdot (\boldsymbol{r} \times \nabla) = i\boldsymbol{\sigma} \cdot (\boldsymbol{r} \times (\boldsymbol{r} \times \nabla)) = i(\boldsymbol{\sigma} \cdot \boldsymbol{r})(\boldsymbol{r} \cdot \nabla) - ir^2\boldsymbol{\sigma} \cdot \nabla \tag{78}$$

Let now

$$\epsilon = -i\alpha^1\alpha^2\alpha^3 \qquad \sigma_k = \epsilon\alpha^k \tag{79}$$

Then multiplying (78) by ϵ^{-1} we get:

$$-r^2\, i\boldsymbol{\alpha}\cdot\nabla = \boldsymbol{\alpha}\cdot\boldsymbol{r}\,\boldsymbol{\sigma}\cdot(\boldsymbol{r}\times\nabla) - i\,\boldsymbol{\alpha}\cdot\boldsymbol{r}\left(r\frac{\partial}{\partial r}\right)$$

Let $\alpha_r = \frac{1}{r}\boldsymbol{\alpha}\cdot\boldsymbol{r}$, then by (39) and (42)

$$-i\,\boldsymbol{\alpha}\cdot\nabla = \frac{1}{r}\alpha_r\left(i\boldsymbol{\sigma}\cdot\boldsymbol{J} - \frac{3}{2}i\right) - i\alpha_r\frac{\partial}{\partial r} = \frac{1}{r}\alpha_r\,(i\beta K - i) - i\alpha_r\frac{\partial}{\partial r}$$

Thus finally we can write (70) in the form

$$H = mc^2\beta - \frac{e^2}{r} + i\hbar c\,\alpha_r\left(\frac{\beta K}{r} - \frac{1}{r} - \frac{\partial}{\partial r}\right) \tag{80}$$

This gives the Dirac equation as an equation in the single variable r, having separated all angular variables.

For the solution of this equation, see – Dirac, *Quantum Mechanics*, Third Edition, Sec. 72, pp. 268–271.

2.11   Solution of Radial Equation

We may choose a two-component representation in which

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \alpha_r = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \qquad \psi = \begin{pmatrix} u \\ v \end{pmatrix} \tag{81}$$

Then

$$\left(E - mc^2\right)u = -\frac{e^2}{r}u + \hbar c\left(\frac{1+K}{r} + \frac{\partial}{\partial r}\right)v$$

$$\left(E + mc^2\right)v = -\frac{e^2}{r}v + \hbar c\left(-\frac{1-K}{r} - \frac{\partial}{\partial r}\right)u \tag{82}$$

Let now

$$a_1 = \frac{-E + mc^2}{\hbar c}; \qquad a_2 = \frac{E + mc^2}{\hbar c}; \qquad \alpha = \frac{e^2}{\hbar c}, \tag{83}$$

the fine structure constant. Then

$$\left(-a_1 + \frac{\alpha}{r}\right)u = \left(\frac{1+K}{r} + \frac{\partial}{\partial r}\right)v$$

$$\left(a_2 + \frac{\alpha}{r}\right)v = \left(\frac{-1+K}{r} - \frac{\partial}{\partial r}\right)u \tag{84}$$

Next put $a = \sqrt{a_1 a_2} = \sqrt{m^2 c^4 - E^2}/\hbar c$ which is the magnitude of the imaginary momentum of a free electron of energy E. Then $\psi \sim e^{-ar}$ at infinity. Hence we write

$$u = \frac{e^{-ar}}{r} f$$

$$v = \frac{e^{-ar}}{r} g \tag{85}$$

So

$$\left(\frac{\alpha}{r} - a_1\right) f = \left(\frac{\partial}{\partial r} - a + \frac{k}{r}\right) g$$

$$\left(\frac{\alpha}{r} + a_2\right) g = \left(-\frac{\partial}{\partial r} + a + \frac{k}{r}\right) f \tag{86}$$

Now we try solutions in series

$$f = \sum c_s r^s, \qquad g = \sum d_s r^s \tag{87}$$

This gives

$$\alpha\, c_s - a_1 c_{s-1} = -ad_{s-1} + (s+k)\, d_s$$

$$\alpha\, d_s + a_2 d_{s-1} = +ac_{s-1} + (-s+k)\, c_s \tag{88}$$

Putting

$$e_s = a_1 c_{s-1} - ad_{s-1} \quad \text{we have} \quad e_s = \alpha\, c_s - (s+k)\, d_s = \frac{a_1}{a} \left(\alpha\, d_s + (s-k)\, c_s\right)$$

$$c_s = \frac{a_1 \alpha + a\,(s+k)}{a_1 \alpha^2 + a_1(s^2 - k^2)} e_s \qquad d_s = \frac{a\alpha - a_1\,(s-k)}{a_1 \alpha^2 + a_1(s^2 - k^2)} e_s$$

$$e_{s+1} = \frac{\left(a_1^2 - a^2\right) \alpha + 2saa_1}{a_1 \alpha^2 + a_1\,(s^2 - k^2)} e_s$$

Suppose the series do not terminate. Then for large s

$$\frac{e_{s+1}}{e_s} \approx \frac{c_{s+1}}{c_s} \approx \frac{2a}{s} \quad \text{hence} \quad f \approx \exp(2ar)$$

This is permissible when a is imaginary. Thus there is a continuum of states with

$$E > mc^2 \tag{89}$$

For real a the series must terminate at both ends in order not to blow up at infinity. Suppose then[8] e_s is non-zero for

$$s = \epsilon + 1, \ \epsilon + 2, \ldots, \epsilon + n \qquad n \geq 1 \tag{90}$$

and otherwise zero. This gives

$$\alpha^2 + \epsilon^2 - k^2 = 0$$

$$\left(a_1^2 - a^2\right)\alpha + 2\left(\epsilon + n\right)aa_1 = 0$$

Now not both c_ϵ and d_ϵ are zero, thus the wave function $r^{-1+\epsilon}$ must be integrable at zero. This gives $\epsilon > -\frac{1}{2}$. But $\epsilon = \pm\sqrt{k^2 - \alpha^2}$. Now $k^2 \geq 1$, hence $\sqrt{k^2 - \alpha^2} > \frac{1}{2}$, and

$$\epsilon = +\sqrt{k^2 - \alpha^2} \tag{91}$$

Also[9]

$$(\epsilon + n)^2 = \left(\frac{a_1^2 - a^2}{2aa_1}\right)^2 \alpha^2 = \left(\frac{\left((mc^2 - E)^2 - (m^2c^4 - E^2)\right)^2}{4\left(m^2c^4 - E^2\right)(mc^2 - E)^2}\right)\alpha^2$$

$$= \frac{4E^2\alpha^2}{4\left(m^2c^4 - E^2\right)}$$

$$E^2 = \frac{m^2c^4}{\left(1 + \frac{\alpha^2}{(\epsilon+n)^2}\right)}$$

Hence in this case

$$E = \frac{mc^2}{\sqrt{1 + \frac{\alpha^2}{\left(n+\sqrt{k^2-\alpha^2}\right)^2}}} \tag{92}$$

Given this positive E, $\left(a_1^2 - a^2\right)$ is negative (see (83) and (84)) and so it is allowable to square $(\epsilon + n)$ to find these solutions, without introducing any difficulties. So for each

$$k = \pm 1, \pm 2, \pm 3, \ldots$$
$$n = 1, 2, 3, \ldots \tag{93}$$

solutions exist, with E given by (92).

The alternative possibility is that *all e_s are zero.* Suppose not both of c_ϵ and d_ϵ are zero. Then $\alpha^2 + \epsilon^2 - k^2 = 0$ as before and so $\epsilon = \sqrt{k^2 - \alpha^2}$. But now

$$a_1 c_\epsilon - a d_\epsilon = 0$$

$$\alpha\, c_\epsilon - (\epsilon + k)d_\epsilon = 0$$

Hence $a\alpha - a_1(\epsilon + k) = 0$ and k must be positive to make $\epsilon + k = \sqrt{k^2 - \alpha^2} + k > 0$. After this the solution goes as before. So solutions (92) exist for

$$n = 0, \qquad k = +1, +2, +3, \ldots \tag{94}$$

The principal quantum number N is

$$N = n + |k|$$

Expanding in powers of α

$$E = mc^2 \left[1 - \frac{1}{2} \frac{\alpha^2}{N^2} + \frac{\alpha^4}{N^3} \left(\frac{3}{8N} - \frac{1}{2|k|} \right) \right] \tag{95}$$

$$\underbrace{\phantom{1 - \frac{1}{2}\frac{\alpha^2}{N^2}}}_{\text{NR levels}} \underbrace{\phantom{\frac{\alpha^4}{N^3}\left(\frac{3}{8N} - \frac{1}{2|k|}\right)}}_{\text{fine structure}}$$

There is *exact* degeneracy between the two states of a given $|k|$. Nonrelativistic states are given by

$$j = \ell + \frac{1}{2} \rightarrow k = -(\ell + 1)$$

$$j = \ell - \frac{1}{2} \rightarrow k = +\ell$$

So

$$\left. \begin{array}{l} {}^2P_{1/2} \text{ is } j = \dfrac{1}{2}, \quad k = 1 \\[2mm] {}^2S_{1/2} \text{ is } j = \dfrac{1}{2}, \quad k = -1 \end{array} \right\} \rightarrow \text{degenerate}$$

$$\quad {}^2S_{3/2} \text{ is } j = \frac{1}{2}, \quad k = -2$$

2.12 Behaviour of an Electron in a Non-Relativistic Approximation

Multiplying the Dirac equation (64) by $\sum_\nu \gamma_\nu \left(\frac{\partial}{\partial x_\nu} + i\frac{e}{\hbar c} A_\nu \right) - \frac{mc}{\hbar}$ we have

$$\sum_\mu \sum_\nu \gamma_\mu \gamma_\nu \left(\frac{\partial}{\partial x_\mu} + i\frac{e}{\hbar c} A_\mu \right) \left(\frac{\partial}{\partial x_\nu} + i\frac{e}{\hbar c} A_\nu \right) \psi - \frac{m^2 c^2}{\hbar^2} \psi = 0 \tag{96}$$

Using $\gamma_\mu^2 = 1, \gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 0$ this gives

$$\sum_\mu \left\{ \left(\frac{\partial}{\partial x_\mu} + \frac{ie}{\hbar c} A_\mu \right)^2 \right\} \psi - \frac{m^2 c^2}{\hbar^2} \psi + \frac{ie}{2\hbar c} \sum_\mu \sum_\nu \sigma_{\mu\nu} F_{\mu\nu} \psi = 0 \qquad (97)$$

Here

$$\sigma_{\mu\nu} = \tfrac{1}{2} \left(\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu \right) \qquad F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu}$$

Thus $F_{12} = H_3$ magnetic field component

$$F_{14} = i \frac{\partial \Phi}{\partial x_1} + \frac{i}{c} \frac{\partial A_1}{\partial t} = -iE_1 \qquad \text{electric field}$$

$$\sigma_{12} = i\sigma_3 \qquad \text{spin component}$$

$$\sigma_{14} = i\alpha_1 \qquad \text{velocity component}$$

Thus (97) becomes

$$\sum_\mu \left\{ \left(\frac{\partial}{\partial x_\mu} + \frac{ie}{\hbar c} A_\mu \right)^2 \right\} \psi - \frac{m^2 c^2}{\hbar^2} \psi - \frac{e}{\hbar c} \left\{ \boldsymbol{\sigma} \cdot \boldsymbol{H} - i\boldsymbol{\alpha} \cdot \boldsymbol{E} \right\} \psi = 0 \qquad (98)$$

This is still exact.

Now in the non-relativistic approximation

$$i\hbar \frac{\partial}{\partial t} = mc^2 + O(1)$$

$$\left\{ \left(\frac{\partial}{\partial x_4} + \frac{ie}{\hbar c} A_4 \right)^2 \right\} - \frac{m^2 c^2}{\hbar^2} = \frac{1}{\hbar^2 c^2} \left\{ \left(-i\hbar \frac{\partial}{\partial t} - e\Phi \right)^2 - m^2 c^4 \right\}$$

$$= \frac{1}{\hbar^2 c^2} \left\{ \left(-i\hbar \frac{\partial}{\partial t} - e\Phi - mc^2 \right) \right.$$

$$\left. \times \left(-i\hbar \frac{\partial}{\partial t} - e\Phi + mc^2 \right) \right\}$$

$$= \frac{1}{\hbar^2 c^2} \{ -2mc^2 + O(1) \}$$

$$\times \left(-i\hbar \frac{\partial}{\partial t} - e\Phi + mc^2 \right)$$

Hence

$$\left(-i\hbar\frac{\partial}{\partial t} - e\Phi + mc^2\right)\psi - \frac{h^2}{2m}\sum_{k=1}^{3}\left\{\left(\frac{\partial}{\partial x_k} + \frac{ie}{\hbar c}A_k\right)^2\right\}\psi$$

$$+ \frac{e\hbar}{2mc}[\boldsymbol{\sigma}\cdot\boldsymbol{H} - i\boldsymbol{\alpha}\cdot\boldsymbol{E}]\psi + O\left(\frac{1}{mc^2}\right) = 0$$

The non-relativistic approximation means dropping the terms $O\left(1/mc^2\right)$. Thus the non-relativistic Schrödinger equation is

$$i\hbar\frac{\partial\psi}{\partial t} = \left\{mc^2 - e\Phi - \frac{h^2}{2m}\sum_{k=1}^{3}\left(\frac{\partial}{\partial x_k} + \frac{ie}{\hbar c}A_k\right)^2 + \frac{e\hbar}{2mc}\left(\boldsymbol{\sigma}\cdot\boldsymbol{H} - i\boldsymbol{\alpha}\cdot\boldsymbol{E}\right)\right\}\psi$$

$$(99)$$

The term $\boldsymbol{\alpha}\cdot\boldsymbol{E}$ is really relativistic, and should be dropped or treated more exactly. Then we have exactly the equation of motion of a non-relativistic particle with a spin magnetic moment equal to

$$M = -\frac{e\hbar}{2mc}\sigma \tag{100}$$

This is one of the greatest triumphs of Dirac, that he got this magnetic moment right out of his general assumptions without any arbitrariness.

It is confirmed by measurements to about one part in 1000. Note that the most recent experiments show a definite discrepancy, and agree with the value

$$M = -\frac{e\hbar}{2mc}\sigma\left\{1 + \frac{e^2}{2\pi\hbar c}\right\} \tag{101}$$

calculated by Schwinger using the complete many-particle theory.

Problem 2. Calculate energy values and wave functions of a Dirac particle moving in a homogeneous infinite magnetic field. This can be done exactly. See F. Sauter, *Zeitschrift für Physik* **69** (1931) 742.

Solution

Take the field $\boldsymbol{B}$ in the z direction.

$$A_1 = -\frac{1}{2}By, \qquad A_2 = \frac{1}{2}Bx$$

The second-order Dirac equation (98) gives for a stationary state of energy $\pm E$

$$\left(\frac{E^2}{\hbar^2 c^2} - \frac{m^2 c^2}{\hbar^2}\right)\psi + \left(\frac{\partial}{\partial x} - \frac{1}{2}\frac{ieB}{\hbar c}y\right)^2 \psi$$

$$+ \left(\frac{\partial}{\partial y} + \frac{1}{2}\frac{ieB}{\hbar c}x\right)^2 \psi + \frac{\partial^2}{\partial z^2}\psi - \frac{eB}{\hbar c}\sigma_z\psi = 0$$

Taking a representation with σ_z diagonal, this splits at once into two states with $\sigma_z = \pm 1$. Also

$$L_z = -i\hbar\left\{x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right\}$$

is a constant of the motion, say $L_z = \ell\hbar$ where ℓ is an integer. And $-i\hbar\frac{\partial}{\partial z} = p_z$. Let $\lambda = |eB\hbar c|$. Then

$$\left\{E^2 - m^2 c^4 - c^2 p_z^2 \pm (\ell_z \pm 1)\lambda\right\}\psi = \hbar^2 c^2 \left\{\frac{1}{4}\frac{\lambda^2 r^2}{\hbar^4 c^4} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\right\}\psi$$

This is an eigenvalue problem with eigenvalues of a two-dimensional harmonic oscillator.

Thus

$$E^2 = m^2 c^4 + c^2 p_z^2 + \lambda\left\{n \pm (\ell_z \pm 1)\right\}$$

where $\ell_z = 0, \pm 1, \pm 2, \ldots, \pm(n-1)$.

So the eigenvalues are

$$E = \sqrt{m^2 c^4 + c^2 p_z^2 + M|eB\hbar c|} \quad \text{with } M = 0, 1, 2, \ldots$$

The lowest state has energy exactly mc^2.

2.13 Summary of Matrices in the Dirac Theory in Our Notation

$$\alpha^k \alpha^\ell + \alpha^\ell \alpha^k = 2\delta_{k\ell}\mathbb{I} \qquad \alpha^k \beta + \beta\alpha^k = 0 \qquad \beta^2 = \mathbb{I} \qquad \sigma_k \sigma_\ell + \sigma_\ell \sigma_k = 2\delta_{k\ell}\mathbb{I}$$

$$\gamma_k = -i\beta\alpha^k \qquad \alpha^k = i\beta\gamma_k \qquad \gamma_4 = \beta \qquad \gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\delta_{\mu\nu}\mathbb{I} \qquad (\gamma_k)^* = \gamma_k$$

$$\alpha^k \gamma_\ell - \gamma_\ell \alpha^k = 2i\delta_{\ell k}\beta \qquad \gamma_5 = \gamma_1 \gamma_2 \gamma_3 \gamma_4 \qquad \gamma_\mu \gamma_5 + \gamma_5 \gamma_\mu = 0$$

$$\alpha^k \gamma_5 - \gamma_5 \alpha^k = 0 \qquad \gamma_5^2 = \mathbb{I}$$

We use the following representation:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \alpha^k = \begin{pmatrix} \mathbb{O} & \sigma_k \\ \sigma_k & \mathbb{O} \end{pmatrix}$$

i.e.

$$\alpha^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \alpha^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$

$$\alpha^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} \mathbb{I} & \mathbb{O} \\ \mathbb{O} & -\mathbb{I} \end{pmatrix}$$

i.e.

$$\beta = \gamma_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\gamma_5 = \begin{pmatrix} \mathbb{O} & -\mathbb{I} \\ -\mathbb{I} & \mathbb{O} \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad \gamma_k = \begin{pmatrix} \mathbb{O} & -i\sigma_k \\ i\sigma_k & \mathbb{O} \end{pmatrix}$$

i.e.

$$\gamma_1 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \quad \gamma_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

$$\gamma_3 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}$$

$$\sigma_k = \epsilon \alpha^k \qquad \alpha^k = \epsilon \sigma_k \qquad \eta = i\epsilon\beta \qquad \epsilon = -i\alpha^1\alpha^2\alpha^3 \epsilon^2 = \eta^2 = \mathbb{I}$$

$$\gamma_5 = -\epsilon \qquad \sigma_k = \eta\gamma_k \qquad \gamma_k = \eta\sigma_k \qquad \epsilon = -i\eta\beta \qquad \eta = -\alpha^1\alpha^2\alpha^3$$

$$\epsilon = \begin{pmatrix} \mathbb{O} & \mathbb{I} \\ \mathbb{I} & \mathbb{O} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\eta = \begin{pmatrix} \mathbb{O} & -i\mathbb{I} \\ i\mathbb{I} & \mathbb{O} \end{pmatrix} = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}$$

$$\alpha^k \sigma_\ell + \sigma_\ell \alpha^k = 2\delta_{\ell k}\epsilon \qquad \gamma_k \sigma_\ell + \sigma_\ell \gamma_k = 2\delta_{\ell k}\eta \qquad \beta\sigma_k - \sigma_k\beta = 0$$

$$\sigma_k \sigma_\ell = \alpha_k \alpha_\ell = \gamma_k \gamma_\ell = i\sigma_m \qquad k, \ell, m = (1,2,3) \text{ cyclicly permuted}$$

$$\alpha^k \epsilon - \epsilon\alpha^k = \gamma_\mu \epsilon + \epsilon\gamma_\mu = \sigma_k \epsilon - \epsilon\sigma_k = 0$$

$$\alpha^k \eta + \eta\alpha^k = \gamma_k \eta - \eta\gamma_k = \sigma_k \eta - \eta\sigma_k = \beta\eta + \eta\beta = 0$$

$$\left. \begin{aligned} \alpha_k \sigma_\ell &= i\alpha_m \\ \sigma_k \gamma_\ell &= i\gamma_m \\ \gamma_k \alpha_\ell &= \beta\sigma_m \end{aligned} \right\} \quad k, \ell, m = (1,2,3) \text{ cyclicly permuted}$$

Comparison with the Dirac notation: $\qquad \rho_1 = \epsilon \qquad \rho_2 = \eta \qquad \rho_3 = \beta$.
   Latin indices: 1, 2, 3. Greek indices: 1, 2, 3, 4.

2.14   Summary of Matrices in the Dirac Theory in the Feynman Notation

$$\alpha^k \alpha^\ell + \alpha^\ell \alpha^k = 2\delta_{k\ell}\mathbb{I} \qquad \alpha^k \beta + \beta\alpha^k = 0 \qquad g_{00} = +1 \qquad g_{kk} = -1$$

$$g_{\mu\nu} = 0, \ \mu \neq \nu \qquad \sigma_k \sigma_\ell + \sigma_\ell \sigma_k = 2\delta_{k\ell}\mathbb{I} \qquad \beta^2 = \mathbb{I} \qquad \gamma_k = \beta\alpha^k$$

$$\alpha^k = \beta\gamma_k \qquad \gamma_0 = \beta \qquad \gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2g_{\mu\nu}\mathbb{I} \qquad (\gamma_k)^* = -\gamma_k$$

$$\alpha^k \gamma_\ell - \gamma_\ell \alpha^k = -2\delta_{\ell k}\beta \qquad \gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3 \qquad \gamma_\mu \gamma_5 + \gamma_5 \gamma_\mu = 0$$

$$\alpha^k \gamma_5 - \gamma_5 \alpha^k = 0 \qquad \gamma_5^2 = -\mathbb{I}$$

Representation:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \alpha^k = \begin{pmatrix} \mathbb{O} & \sigma_k \\ \sigma_k & \mathbb{O} \end{pmatrix}$$

i.e.

$$\alpha^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad \alpha^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$

$$\alpha^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \qquad \beta = \begin{pmatrix} \mathbb{I} & \mathbb{O} \\ \mathbb{O} & -\mathbb{I} \end{pmatrix}$$

i.e.

$$\beta = \gamma_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \gamma_k = \begin{pmatrix} \mathbb{O} & \sigma_k \\ -\sigma_k & \mathbb{O} \end{pmatrix}$$

i.e.

$$\gamma_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \qquad \gamma_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$

$$\gamma_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\rho_1 = \begin{pmatrix} \mathbb{O} & \mathbb{I} \\ \mathbb{I} & \mathbb{O} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \gamma_5$$

$$\rho_2 = \begin{pmatrix} \mathbb{O} & -i\mathbb{I} \\ i\mathbb{I} & \mathbb{O} \end{pmatrix} = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}$$

$$\sigma_k = \rho_1 \alpha^k \qquad \alpha^k = \rho_1 \sigma_k \qquad \rho_2 = i\rho_1\beta \qquad \rho_1 = -i\alpha^1\alpha^2\alpha^3 \qquad \rho_1^2 = \rho_2^2 = \mathbb{I}$$

$$\sigma_k = -i\rho_2\gamma_k \qquad \gamma_k = i\rho_2\sigma_k \qquad \rho_1 = -i\rho_2\beta \qquad \rho_2 = -\alpha^1\alpha^2\alpha^3\beta$$

$$\alpha^k\sigma_\ell + \sigma_\ell\alpha^k = 2\delta_{\ell k}\rho_1 \qquad \gamma_k\sigma_\ell + \sigma_\ell\gamma_k = -2\delta_{\ell k}\rho_2 \qquad \beta\sigma_k - \sigma_k\beta = 0$$

$$\sigma_k\sigma_\ell = \alpha_k\alpha_\ell = -\gamma_k\gamma_\ell = i\sigma_m \qquad k,\ell,m = (1,2,3) \text{ cyclicly permuted}$$

$$\alpha^k \rho_1 - \rho_1 \alpha^k = \gamma_\mu \rho_1 + \rho_1 \gamma_\mu = \sigma_k \rho_1 - \rho_1 \sigma_k = 0$$

$$\alpha^k \rho_2 + \rho_2 \alpha^k = \gamma_k \rho_2 - \rho_2 \gamma_k = \sigma_k \rho_2 - \rho_2 \sigma_k = \beta \rho_2 + \rho_2 \beta = 0$$

$$\left.\begin{array}{l} \alpha_k \sigma_\ell = i\alpha_m \\ \sigma_k \gamma_\ell = i\gamma_m \\ \gamma_k \alpha_\ell = i\beta\sigma_m \end{array}\right\} \quad k, \ell, m = (1, 2, 3) \text{ cyclicly permuted}$$

Latin indices: 1, 2, 3. Greek indices: 0, 1, 2, 3.

Scattering Problems and Born Approximation

3.1   General Discussion

The problem of scattering of a Dirac particle by a potential can be treated exactly by finding the continuum solutions of the Dirac equation. This is a complicated business even for the simplest case of a Coulomb force. It was done by Mott, *Proc. Roy. Soc. A* **135** (1932) 429.

For most purposes in relativistic problems, and always when the scattering is produced by complicated effects involving radiation theory, one uses the *Born approximation*. That is, we treat the scattering only to first order in the interaction, or only to some definite order in which we are interested.

The formula for scattering from an initial state A to a final state B lying in a continuum of states is the transition probability per unit time,

$$w = \frac{2\pi}{\hbar} \rho_E \left| V_{BA} \right|^2 \tag{102}$$

This you ought to know. $\rho_E =$ density of final states per unit energy interval. V_{BA} is the matrix element of the potential V for the transition. Here V may be anything, and may be itself a second-order or higher order effect obtained by using higher-order perturbation theory.

The difficulties in real calculations usually come from the factors 2 and π and the correct *normalization of states*. Always I shall normalize the continuum states not in the usual way (one particle per unit volume) which is non-invariant, but instead

$$\text{One particle per volume } \frac{mc^2}{|E|} \tag{103}$$

where $|E|$ is the energy of the particles. Then under a Lorentz transformation the volume of a fixed region transforms like $1/|E|$ and so the definition stays invariant.

Thus a continuum state given by the spinor $\psi = u \exp\{(i\boldsymbol{p} \cdot \boldsymbol{x} - iEt)/\hbar\}$ is normalized so that

$$u^* u = \frac{|E|}{mc^2} \tag{104}$$

Now if we multiply the Dirac equation for a free particle, (44), by $\overline{u}$ on the left, we get $Eu^*\beta u = cu^*\beta \boldsymbol{\alpha} \cdot \boldsymbol{p}u + mc^2 u^* u$; its complex conjugate is $Eu^*\beta u = -cu^*\beta\boldsymbol{\alpha} \cdot \boldsymbol{p}u + mc^2 u^* u$ since $\beta\boldsymbol{\alpha}$ is anti-Hermitian; then by adding we get

$$E\,\overline{u}u = mc^2 u^* u \tag{105}$$

Therefore the normalization becomes

$$\left. \begin{array}{l} \overline{u}u = +1 \;\; \text{for electron states} \\ \phantom{\overline{u}u} = -1 \;\; \text{for positron states} \end{array} \right\} = \epsilon; \qquad \text{This is the definition of } \epsilon. \tag{106}$$

With this normalization the density of states in momentum space is one per volume h^3 of phase space, that is to say

$$\rho = \frac{1}{h^3} \frac{mc^2}{|E|} \, dp_1 dp_2 dp_3 \tag{107}$$

per volume $dp_1 dp_2 dp_3$ of momentum space, for *each* direction of spin and each sign of charge. Now we have again an invariant differential

$$\frac{dp_1 dp_2 dp_3}{|E|} \tag{108}$$

3.2   Projection Operators

Usually we are not interested in the spin either of an intermediate or of an initial or final state. Therefore we have to make sums over spin states which are of the form

$$S = \sum_2 (\overline{s}Ou)(\overline{u}Pr) \tag{109}$$

where O and P are some kind of operators, s and r some kind of spin states, and the sum is over the two spin states u of an electron of momentum $\boldsymbol{p}$ and energy E.

Let us write

$$\not p = \sum_\mu p_\mu \gamma_\mu \,, \qquad p_4 = iE/c \tag{110}$$

The Dirac equation satisfied by u is

$$(\not p - imc)\, u = 0 \tag{111}$$

The two spin states with momentum 4-vector $-p$ satisfy

$$(\not p + imc)\, u = 0 \tag{112}$$

As one can easily show from (48), these 4 states are all orthogonal in the sense that $(\overline{u}' u) = 0$ for each pair $u'u$. Therefore the identity operator may be written in the form

$$\mathbb{I} = \sum_4 (u\overline{u})\, \epsilon \tag{113}$$

summed over all 4 states with ϵ defined as earlier. Hence by (111) and (112) we can write (109) as

$$S = \sum_4 \left(\overline{s}O \frac{\not p + imc}{2imc} \epsilon u \right) (\overline{u}Pr) = (\overline{s}O\Lambda_+ Pr) \tag{114}$$

by virtue of (113); here the operator

$$\Lambda_+ = \frac{\not p + imc}{2imc} \tag{115}$$

is a projection operator for *electrons* of momentum p.

In the same way for a sum over the two positron states u with momentum p energy E

$$S = \sum_2 (\overline{s}Ou)\,(\overline{u}Pr) = (\overline{s}O\Lambda_- Pr) \tag{116}$$

with

$$\Lambda_- = \frac{\not p - imc}{2imc} \tag{117}$$

and we have

$$\Lambda_+ - \Lambda_- = \mathbb{I} \tag{118}$$

These projection operators are covariant. In Heitler the business is done in a different way which makes them non-covariant and more difficult to handle.

Note that here charge-conjugate wave functions are *not* used. The positrons of momentum p are represented by the electron wave functions u of momentum $-p$ energy $-E$.

3.3   Calculation of Traces

Suppose we have to calculate an expression such as

$$\frac{1}{2} \sum_I \sum_F (\overline{u}_F O u_I)(\overline{u}_I O u_F)$$

summed over electron states only. This gives

$$\frac{1}{2} \sum (\overline{u}_F O \Lambda_+ O \Lambda_+ u_F)\, \epsilon$$

summed over all four spin states u_F. To calculate this, let us consider the general expression

$$\sum_u \epsilon\, (\overline{u} Q u)$$

summed over all 4 spin states, where Q is any 4×4 matrix.

Let Q have the eigenvectors w_1, w_2, w_3, w_4 with eigenvalues $\lambda_1, \lambda_2, \lambda_3, \lambda_4$. Then

$$Q = \sum_{k=1}^{4} \lambda_k w_k w_k^*$$

and

$$\sum_u \epsilon\, (\overline{u} Q u) = \sum_u \epsilon \sum_{k=1}^{4} \lambda_k (\overline{u} w_k)(w^* u) = \sum_{k=1}^{4} \lambda_k w^* \left\{ \sum_u \epsilon\, (u \overline{u}) \right\} w_k$$

Hence by (113)

$$\sum_u \epsilon\, (\overline{u} Q u) = \sum \lambda$$

Now $\sum \lambda$ = sum of diagonal elements of Q = Trace of Q. Thus[10]

$$\sum_u \epsilon\, (\overline{u} O u) = \mathrm{Tr}\, Q$$

and this is always easy to calculate.

Problem 3. Given a steady potential V a function of position, and a beam of incident particles, electrons. Solve the Schrödinger equation in the Born approximation

(a) By stationary perturbation theory

(b) By time-dependent perturbation theory.

Show that the results agree, with a transition probability per unit time given by $w = (2\pi/\hbar)\rho_E|V_{BA}|^2$. Evaluate the cross section in the case $V = -Ze^2/r$, averaging spin over initial state and summing over final state.

(c) Repeat the calculation with particles obeying the Klein–Gordon equation, leaving out the V^2 term, by either method. Compare the angular distribution in the two cases.

Problem 4. A nucleus (O^{16}) has an even[11] $j = 0$ ground state and an even $j = 0$ excited state at 6 MeV. Calculate the total rate of emission of pairs, and the angular and momentum distributions.

Solution

Let ΔE be the excitation energy, ρ_N and $\boldsymbol{j}_N$ the charge and current density operators of the nucleus. Then for the transition we are interested in ρ_N and $\boldsymbol{j}_N$ are functions of position r with the time-variation of the single matrix element given by $\exp\{-i\Delta E/\hbar\}$. Also

$$\nabla \cdot \boldsymbol{j}_N = -\frac{\partial \rho_H}{\partial t} = i\frac{\Delta E}{\hbar}\rho_N \tag{119}$$

The electrostatic potential V of the nucleus has the matrix element given by

$$\nabla^2 V = -4\pi\rho_N \tag{120}$$

The states being spherically symmetric, ρ_N is a function of r only, and so the general solution of Poisson's equation simplifies to[12]

$$V(r) = -\frac{6\pi}{r} \int_0^r r_1^2 \rho_N(r_1)\, dr_1 \tag{121}$$

Outside the nucleus $V(r) = Ze^2/r$ is constant in time, and so the matrix element of $V(r)$ for this transition is zero. In fact from (119) and (120) we get by integration

$$V(r) = \frac{\hbar}{i\Delta E}(-4\pi)(-r)j_{No}(r) = \frac{4\pi r\hbar}{i\Delta E}j_{No}(r) \tag{122}$$

where j_{No} is the outward component of the current.

The interaction which creates pairs is then

$$I = \int \frac{4\pi r\hbar}{i\Delta E} j_{No}(r)\left(-e\psi^*\psi(r)\right) d\tau \tag{123}$$

As an approximation consider the de Broglie wavelengths of all pairs long compared with the nuclear size. Then

$$I = \psi^* \psi(0) \frac{4\pi \hbar e i}{\Delta E} \int r \, j_{No}(r) \, d\tau \tag{124}$$

The constant $\int r \, j_{No}(r) \, d\tau$ is not known exactly. Suppose as an order of magnitude estimate that the nucleus of charge Ze is uniformly spread over a sphere of radius r_o in the ground state, and also in the excited state. Since ρ_N is roughly uniform inside the nucleus we have by integrating (119):

$$j_N = \frac{i\Delta E}{3\hbar} \overline{r} \rho_N \,, \qquad \text{and thus}$$

$$I = \psi^* \psi(0) \left(\frac{-4\pi e}{3} \right) \int r^2 \rho_N(r) \, d\tau = \psi^* \psi(0) \left(\frac{-4\pi e}{3} \right) Q e^{-i\Delta Et/\hbar} \tag{125}$$

Q is roughly a measure of the charge-moment of inertia of the nucleus, and is equal to

$$\frac{3}{5} Z e r_o^2$$

Thus

$$I = -\frac{4\pi Z e^2 r_o^2}{5} \{\psi^* \psi(0)\} \, e^{-i\Delta Et/\hbar} \tag{126}$$

So the problem is just to compute the probabilities of pair-emission with this interaction. Note that real radiation is strictly forbidden in a 0–0 transition, and so these pairs are actually observed in the reaction

$$p + F_{19} \rightarrow O_{16}^* + \alpha \rightarrow O_{16} + e^+ + e^- + \alpha \tag{127}$$
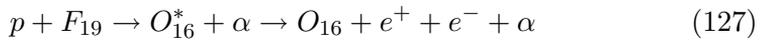
Is it correct to take for the interaction just

$$\int V(r) \, (-e\psi^*\psi) \, d\tau$$

taking the Coulomb potential of the nuclear charge and ignoring all electrodynamic effects? Yes. Because in general the interaction would be

$$\int \left\{ \varphi \, (-e\psi^*\psi) - \sum_k A_k \, (-e\psi^* \alpha_k \psi) \right\} d\tau \tag{128}$$

where φ, A_k are the scalar and vector potentials satisfying the Maxwell equations

$$\nabla^2 \varphi + \frac{1}{c} \nabla \cdot \frac{\partial \boldsymbol{A}}{\partial t} = -4\pi \rho_N$$

$$\nabla^2 \boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \nabla\left\{\nabla \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial \varphi}{\partial t}\right\} = -\frac{4\pi}{c}\boldsymbol{j}_N$$

The matrix element of the interaction (128) is unchanged by any gauge transformation of the $(\boldsymbol{A}, \varphi)$. Therefore we may take the gauge in which

$$\nabla \cdot \boldsymbol{A} = 0$$

Incidentally, since $\varphi = V(r)$, the second Maxwell equation reduces to

$$\nabla^2 \boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - = -\frac{4\pi}{c}\boldsymbol{j}_N$$

Now, since there is no free radiation present, also $\nabla \times \boldsymbol{A} = 0$, and hence $\boldsymbol{A} = 0$, in this gauge, and therefore we can indeed ignore all electrodynamic effects.

Let us calculate then the probability of pair emission with the interaction (126). A typical final state has an electron of momentum p_1 and a positron of momentum p_2, with energies E_1, E_2 and spins u_1, u_2 respectively. For the creation of this pair the matrix element of I is

$$I = -C\,\overline{u}_1 \beta u_2\,, \qquad C = \frac{4\pi Z e^2 r_o^2}{5} \tag{128a}$$

The density of final states is by (107)

$$\frac{1}{(2\pi\hbar)^6}\frac{m^2 c^4}{E_1 E_2}\,p_1^2 dp_1\,d\omega_1\,p_2^2 dp_2\,d\omega_2 \tag{129}$$

where $d\omega_1$ and $d\omega_2$ are the solid angles for p_1 and p_2. The creation probability per unit time is thus by (102)

$$w = \frac{2\pi}{\hbar}\frac{\rho_E}{dE}|I|^2 = \frac{2\pi}{\hbar}\frac{dp_1 dp_2}{d\left(E_1 + E_2\right)}C^2\frac{1}{(2\pi\hbar)^6}\frac{m^2 c^4\,p_1^2\,p_2^2\,d\omega_1\,d\omega_2}{E_1 E_2}\sum_{u_1,u_2}|\,\overline{u}_1 \beta u_2\,|^2 \tag{130}$$

Now fixing p_1,

$$\frac{dp_2}{d(E_1 + E_2)} = \frac{dp_2}{dE_2} = \frac{E_2}{c^2 p_2}$$

and

$$\sum_{u_1,u2}|\,\overline{u}_1 \beta u_2\,|^2 = \sum_{u_1,u2}(\overline{u}_1 \beta u_2)\,(\overline{u}_2 \beta u_1) = \mathrm{Tr}\left\{\beta\frac{\not{p}_2 - imc}{2imc}\,\beta\frac{\not{p}_1 + imc}{2imc}\right\}$$

$$= -1 + \frac{\boldsymbol{p}_1 \cdot \boldsymbol{p}_2}{m^2 c^2} + \frac{E_1 E_2}{m^2 c^4} = \frac{E_1 E_2 - m^2 c^4 + c^2 p_1 p_2 \cos\theta}{m^2 c^4}$$

where θ is the angle between the pair. Then writing in (130)

$$dE_1 = dp_1 \frac{c^2 p_1}{E_1}\,, \qquad d\omega_1 = 4\pi\,, \qquad d\omega_2 = 2\pi \sin\theta\, d\theta$$

we obtain the differential probability[13] in E_1 and θ

$$w_o = \frac{4Z^2 e^4 r_o^4}{25\pi c^4 \hbar^7}\, p_1 p_2\, dE_1 \left(E_1 E_2 - m^2 c^4 + c^2 p_1 p_2 \cos\theta\right) \sin\theta\, d\theta \qquad (131)$$

Since

$$\Delta E = 6\,\mathrm{MeV} = 12\, mc^2$$

we can to a good approximation treat all particles as extreme relativistic. Thus

$$w_o = \frac{4Z^2 e^4 r_o^4}{25\pi c^6 \hbar^7}\, E_1^2\, E_2^2\, dE_1\, (1 + \cos\theta) \sin\theta\, d\theta \qquad (132)$$

So the pairs have an angular distribution concentrated in the *same* hemisphere, and predominantly *equal* energies. Then, since

$$\int_0^\pi (1 + \cos\theta) \sin\theta\, d\theta = 2\,, \quad \text{and} \quad \int_0^{\Delta E} E_1^2 E_2^2\, dE_1$$

$$= \int_0^{\Delta E} E_1^2 (E_1 + \Delta E)^2 dE_1 = \frac{1}{15}(\Delta E)^5$$

the total creation probability per unit time is

$$w_T = \frac{4Z^2 e^4 r_o^4}{25\pi \hbar^7 c^6} \frac{1}{15}\,(\Delta E)^5 \qquad (133)$$

Numerically

$$\frac{Ze^2}{\hbar c} \approx \frac{1}{17} \quad \text{and} \quad \frac{\Delta E r_o}{\hbar c} \approx \frac{1}{10} \quad \text{since } r_o = 4 \times 10^{-13}\ \mathrm{cm}$$

Hence the lifetime will be

$$\tau = 15 \times 25\pi \times 10^5 \times 17^2 \times \frac{1}{4} \times \frac{r_o}{c} = 10^{10}\, \frac{r_o}{c} \approx 10^{-13}\ \mathrm{sec.} \qquad (134)$$

3.4 Scattering of Two Electrons in Born Approximation. The Møller Formula

We calculate now the transition scattering matrix element M between an initial state A consisting of 2 electrons with momenta p_1, p_2 and spin states u_1, u_2 and a final state B consisting of two electrons with momenta p'_1, p'_2 and spin states u'_1, u'_2. Thus M gives the probability amplitude for arriving in state B after a long time when the system is known to be in state A to begin with. Hence M itself should be an *invariant* relativistically.

We treat the interaction in Born approximation i.e. consider the particles to go directly from the free-particle state A to the free-particle state B by applying the interaction operator once to state A. For electrons at reasonably high or relativistic velocities this will be a very good approximation ($e^2/\hbar v \ll 1$). Also we treat the electromagnetic interaction classically, just as in the O^{16} problem, taking the field produced by particle 1 according to classical Maxwell equations to act directly on particle 2. This ignores the fact that the field consists of quanta. We shall see later, after we have developed the quantum field theories, that this introduces *no error* so long as we are in the Born approximation.

For the field produced by particle 1 in a transition from the state p_1, u_1 to p'_1, u'_1 we have the matrix elements $\varphi_{(1)}, \boldsymbol{A}_{(1)}$ say. We use now not the gauge in which $\nabla \cdot \boldsymbol{A} = 0$, but the covariant gauge in which[14]

$$\sum_\mu \frac{\partial A_\mu}{\partial x_\mu} = 0 \qquad A_4 = i\varphi \tag{135}$$

So using covariant notations we have in this gauge

$$\sum_\nu \frac{\partial^2}{\partial x_\nu^2} A_{\mu(1)} = +4\pi e s_{\mu(1)} \qquad \text{(charge is } - e) \tag{136}$$

$$s_{\mu(1)} = i \left(\overline{u}'_1 \gamma_\mu u_1\right) \exp\left\{\sum_\nu \frac{i}{\hbar} \left(p_{1\nu} - p'_{1\nu}\right) x_\nu\right\} \tag{137}$$

whence

$$A_{\mu(1)} = -4\pi i e \hbar^2 \left[\frac{(\overline{u}'_1 \gamma_\mu u_1) \exp\left\{\sum_\nu \frac{i}{\hbar} \left(p_{1\nu} - p'_{1\nu}\right) x_\nu\right\}}{\sum_\lambda \left(p_{1\lambda} - p'_{1\lambda}\right)^2}\right] \tag{138}$$

where

$$\sum_\nu (p_{1\nu} - p'_{1\nu})^2 = |\boldsymbol{p}_1 - \boldsymbol{p}'_1|^2 - \frac{1}{c^2}(E_1 - E'_1)^2 \tag{139}$$

The effect of the field (138) on particle 2 is given by the interaction term in the Dirac equation for particle 2

$$-e\varphi + e\boldsymbol{\alpha}\cdot\boldsymbol{A} = ie\beta\sum_\mu \gamma_\mu A_\mu \tag{140}$$

This gives for particle 2 for the transition from state p_2, u_2 to p'_2, u'_2 a transition matrix element

$$\int d\tau\, \psi'^*_2 \left(ie\beta\sum_\mu \gamma_\mu A_{\mu\,(1)}\right)\psi_2 \tag{141}$$

a 3-dimensional integral over space at the time t say. For the total transition matrix element M by first order perturbation method

$$M = -\frac{i}{\hbar}\int_{-\infty}^{\infty} dt \int d\tau\, \overline{\psi}'_2\left(ie\sum_\mu \gamma_\mu A_{\mu\,(1)}\right)\psi_2$$

$$= -\frac{i}{\hbar c}\int d^4x\, \overline{\psi}'_2\left(ie\sum_\mu \gamma_\mu A_{\mu\,(1)}\right)\psi_2 \tag{142}$$

where the 4-fold integral is $dx_1, dx_2, dx_3, dx_0, x_0 = ct$. Putting in the values of $A_{\mu\,(1)}, \psi'_2$ and ψ_2, we get

$$M = -\frac{4\pi e^2 \hbar i}{c}\sum_\mu (\overline{u}'_2 \gamma_\mu u_2)(\overline{u}'_1 \gamma_\mu u_1)\frac{1}{\sum_\nu (p_{1\nu} - p'_{1\nu})^2}\int d^4x$$

$$\times \exp\left\{\sum_\lambda \frac{i}{\hbar}(p_{1\lambda} - p'_{1\lambda} + p_{2\lambda} - p'_{2\lambda})x_\lambda\right\}$$

$$= -\frac{4\pi e^2 \hbar i}{c}\sum_{\mu,\nu}\frac{(\overline{u}'_2 \gamma_\mu u_2)(\overline{u}'_1 \gamma_\mu u_1)}{(p_{1\nu} - p'_{1\nu})^2}(2\pi\hbar)^4\delta^4(p_1 + p_2 - p'_1 - p'_2) \tag{143}$$

where $\delta^4(x) = \prod_{k=1}^4 \delta(x_k)$.

There is also the exchange process in which the particle p_1, u_1 goes to p'_2, u'_2 and *vice-versa*. This gives a contribution to M with a minus sign

since the wave function ought to be taken antisymmetric between the two particles. Hence the final result is

$$M = -\frac{4\pi e^2 \hbar i}{c}(2\pi\hbar)^4 \delta^4\left(p_1 + p_2 - p_1' - p_2'\right)$$

$$\times \sum_{\mu,\nu}\left\{\frac{(\overline{u}_2'\gamma_\mu u_2)\,(\overline{u}_1'\gamma_\mu u_1)}{(p_{1\nu} - p_{1\nu}')^2} - \frac{(\overline{u}_2'\gamma_\mu u_1)\,(\overline{u}_1'\gamma_\mu u_2)}{(p_{1\nu} - p_{2\nu}')^2}\right\} \qquad (144)$$

This covariant formula is elegant and easy to arrive at. The question now is, how does one go from such a formula to a cross-section?

Generally, suppose in such a 2-particle collision process the transition matrix is

$$M = K(2\pi\hbar)^4\delta^4\left(p_1 + p_2 - p_1' - p_2'\right) \qquad (145)$$

Then what will be the cross-section in terms of K? We do this calculation once here so that later we can stop when we have found formulae for M of the type (145) which come for example conveniently in this form out of radiation theory.

3.5   Relation of Cross-sections to Transition Amplitudes

Let w be the transition probability per unit volume and per unit time. This is related to the transition probability for a single final state, which is

$$w_s = c|K|^2(2\pi\hbar)^4\delta^4\left(p_1 + p_2 - p_1' - p_2'\right) \qquad (146)$$

since in $|M|^2$ one of the two $(2\pi\hbar)^4\delta^4(p_1 + p_2 - p_1' - p_2')\,/c$ factors represents merely the volume of space-time in which the interaction can occur. The number of final states is by (107)

$$\frac{1}{(2\pi\hbar)^6}\frac{mc^2}{|E_1'|}\frac{mc^2}{|E_2'|}\,dp_{11}'\,dp_{12}'\,dp_{13}'\,dp_{21}'\,dp_{22}'\,dp_{23}' \qquad (147)$$

Multiplying (146) by (147) gives the total transition probability

$$w = |K|^2\frac{1}{(2\pi\hbar)^2}\frac{m^2c^4}{E_1'E_2'}\,c\,\delta^4\left(p_1 + p_2 - p_1' - p_2'\right)\,dp_{11}'\,dp_{12}'\,dp_{13}'\,dp_{21}'\,dp_{22}'\,dp_{23}' \qquad (148)$$

As $\delta(ax) = \frac{1}{a}\delta(x)$, we have

$$\delta^4\left(p_1 + p_2 - p_1' - p_2'\right) = \delta^3\left(\boldsymbol{p}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1' - \boldsymbol{p}_2'\right)c\,\delta(E_1 + E_2 - E_1' - E_2')$$

and the integration over dp_2 gives then by the momentum conservation

$$w = |K|^2 \frac{c^2}{(2\pi\hbar)^2} \frac{m^2c^4}{E_1'E_2'} \delta(E_1 + E_2 - E_1' - E_2') \, dp_{11}' \, dp_{12}' \, dp_{13}' \qquad (148a)$$

Furthermore,

if $f(a) = 0$, we have $f(x) = f(a) + f'(a)(x - a) = f'(a)(x - a)$

and thus

$$\delta(f(x)) = \delta\{f'(a)(x - a)\} = \frac{\delta(x - a)}{f'(a)}$$

Applying this to (148a) with $f(x) = f(p_{13}') = E_1 + E_2 - E_1' - E_2'$ and $a = (p_{13}')_c$ = the value of p_{13}' giving momentum and energy conservation, we get

$$\delta(E_1 + E_2 - E_1' - E_2') = \frac{1}{\dfrac{d(E_1 + E_2 - E_1' - E_2')}{dp_{13}'}} \delta\{p_{13}' - (p_{13}')_c\}$$

Hence we finally obtain

$$w = |K|^2 \frac{m^2c^4}{E_1'E_2'} \frac{c^2}{(2\pi\hbar)^2} \frac{dp_{11}' \, dp_{12}' \, dp_{13}'}{d(E_1' + E_2')}$$

Choose a Lorentz-system in which p_1 and p_2 are both along the x_3-direction and take p_{11}' and p_{12}' as the variables over which the transition probability is taken. This is necessary for relativistic invariance. Then p_{11}' and p_{12}' being fixed and having from the momentum conservation $p_{13}' = -p_{23}'$ we get

$$\frac{d(E_1' + E_1')}{dp_{13}'} = \left| \frac{dE_1'}{dp_{13}'} - \frac{dE_2'}{dp_{23}'} \right| = c^2 \frac{|E_2' p_{13}' - E_1' p_{23}'|}{E_1' E_2'} \qquad (149)$$

Then the cross-section σ is defined in this system by

$$\sigma = \frac{wV_1V_2}{|\boldsymbol{v}_1 - \boldsymbol{v}_2|} \qquad (150)$$

where V_1 is the normalization volume for particle 1, and v_1 its velocity. In fact by (103)

$$V_1 = \frac{mc^2}{E_1} \qquad V_2 = \frac{mc^2}{E_2} \qquad (\boldsymbol{v}_1 - \boldsymbol{v}_2) = \frac{c^2\boldsymbol{p}_1}{E_1} - \frac{c^2\boldsymbol{p}_2}{E_2} \qquad (151)$$

Hence the cross-section becomes

$$\sigma = \frac{w\,\left(mc^2\right)^2}{c^2\,|\boldsymbol{p}_1 E_2 - \boldsymbol{p}_2 E_1|}$$

$$= |K|^2 \frac{\left(mc^2\right)^4}{c^2\,|E_2 p_{13} - E_1 p_{23}|\,|E_2' \, p_{13}' - E_1' \, p_{23}'|}\,\frac{1}{(2\pi\hbar)^2}\,dp_{11}'\,dp_{12}' \quad (152)$$

It is worth noting that the factor $\boldsymbol{p}_1 E_2 - \boldsymbol{p}_2 E_1$ is invariant under Lorentz transformations leaving the x_1 and x_2 components unchanged (e.g.boosts parallel to the x_3 axis.)[15] To prove this, we have to show that $p_{13} E_2 - p_{23} E_1 = \tilde{p}_{13} \tilde{E}_2 - \tilde{p}_{23} \tilde{E}_1$ (where ~ denotes the quantities after the Lorentz transformation) because we have chosen a Lorentz system in which the direction of the momentum vector is the x_3 axis. Then

$$\tilde{E} = E \cosh\theta - cp \sinh\theta$$

$$\tilde{p} = p \cosh\theta - \frac{E}{c}\sinh\theta$$

Since $E^2 = p^2 c^2 + m^2 c^4$, we can write

$$E = mc^2 \cosh\phi \qquad pc = mc^2 \sinh\phi\,, \quad \text{which makes}$$

$$\tilde{E} = mc^2 \cosh(\phi - \theta) \qquad \tilde{p}c = mc^2 \sinh(\phi - \theta) \quad \text{and thus}$$

$$\tilde{E}_2 \tilde{p}_{13} - \tilde{E}_1 \tilde{p}_{23} = m^2 c^3 \{\cosh(\phi_2 - \theta)\sinh(\phi_1 - \theta)$$
$$- \cosh(\phi_1 - \theta)\sinh(\phi_2 - \theta)\}$$
$$= m^2 c^3 \sinh(\phi_1 - \phi_2)$$

independently of θ. Hence we see that σ is invariant under Lorentz transformations parallel to the x_3 axis.

3.6   Results for Møller Scattering

One electron initially at rest, the other initially with energy $E = \gamma mc^2$;

$$\gamma = \frac{1}{\sqrt{1 - (v/c)^2}}$$

$$\text{scattering angle} = \theta \text{ in the lab system}$$
$$= \theta^* \text{ in the center-of-mass system}$$

Then the differential cross-section is (Mott and Massey, *Theory of Atomic Collisions*, 2nd ed., p. 368)

$$2\pi\sigma(\theta)\,d\theta = 4\pi \left(\frac{e^2}{mv^2}\right)^2 \left(\frac{\gamma+1}{\gamma^2}\right) dx$$

$$\times \left\{ \frac{4}{(1-x^2)^2} - \frac{3}{1-x^2} + \left(\frac{\gamma-1}{2\gamma}\right)^2 \left(1 + \frac{4}{1-x^2}\right) \right\} \quad (153)$$

with

$$x = \cos\theta^* = \frac{2 - (\gamma+3)\sin^2\theta}{2 + (\gamma-1)\sin^2\theta}$$

Without spin you get simply

$$4\pi \left(\frac{e^2}{mv^2}\right)^2 \left(\frac{\gamma+1}{\gamma^2}\right) dx \left\{ \frac{4}{(1-x^2)^2} - \frac{3}{1-x^2} \right\}$$

The effect of spin is a measurable *increase* of scattering over the Mott formula. The effect of exchange is roughly the $\dfrac{3}{1-x^2}$ term. Positron-electron scattering is very similar. Only the exchange effect is different because of annihilation possibility.

3.7   Note on the Treatment of Exchange Effects

The correctly normalized initial and final states in this problem are

$$\frac{1}{\sqrt{2}} \{\psi_1(1)\psi_2(2) - \psi_1(2)\psi_2(1)\}$$

$$\frac{1}{\sqrt{2}} \{\psi'_1(1)\psi'_2(2) - \psi'_1(2)\psi'_2(1)\}$$

(154)

where $\psi_2(1)$ means the particle 2 in the state 1, and so on. With these states the matrix element M is exactly as we have calculated it including the exchange term.

The number of possible final states is only one half of the number of states of two distinguishable particles. But this does not bring a factor $\frac{1}{2}$ into the differential cross-section, because the density of antisymmetrical states, in which one of the two particles has a momentum in a particular range dp_1, dp_2, dp_3 is exactly the same as the density of states of 2 distinguishable

particles in which the particle labelled 1 lies in the given range. Hence the general rule: the differential cross-section does *not* have a factor $\frac{1}{2}$, the total cross section *does* because each final state may only be counted once in integrating over the angles.

3.8   Relativistic Treatment of Several Particles

The Møller treatment of the interaction of two electrons succeeds because the field of particle 1 is calculated for all time without taking any account of the effect of particle 2 on particle 1. How can one do a better calculation taking such reactions into account? Clearly we must construct an equation of motion which follows the motions of both particles continuously in time and keeps them in step with each other. So we must have a Dirac equation for 2 electrons, taking exact account of their interaction by including in the equation the behaviour of the Maxwell field too.

This kind of 2-particle Dirac equation is no longer relativistically invariant, if we give each particle a separate position in space but all the same *time.* To avoid this Dirac constructed the *many-time* theory in which each electron has its own private time coordinate, and satisfies its private Dirac equation. This theory is all right in principle. But it becomes hopelessly complicated when pairs are created and you have equations with new time-coordinates suddenly appearing and disappearing. In fact the whole program of quantizing the electron theory as a theory of discrete particles each with its private time becomes nonsense when you are dealing with an infinite "sea" or an indefinite number of particles. So we have come to the end of what we can do with the relativistic quantum theory of particles.

Where did the theory go wrong? Obviously a lot of the troubles arose from the fact that a particle was always described by an operator r representing its position at the time t, t being a number and not an operator. This made the interpretation of the formalism *essentially* non-relativistic even when the equations were formally invariant. In equations like the Klein–Gordon and the Dirac equation, the space and time coordinates appear symmetrically. Thus we are led to the following new view-point.

Relativistic quantum theory is the study of quantities ψ which are functions of four coordinates x_1, x_2, x_3, x_0, all the coordinates being c-numbers and only the expressions containing ψ being operators describing the dynamical system.

The dynamical system is specified by the quantity ψ existing at all points of space-time, and so consists of a system of fields. Relativistic quantum theory is necessarily a field theory.

The process of reinterpreting a one-particle wave-function like the Dirac ψ as a quantized field operator is called *Second Quantization*.

Field Theory

Before we can begin on the program of constructing our quantum theory of fields, we must make some remarks about Classical Field Theory.

4.1 Classical Relativistic Field Theory

We take a field with components (vector, spinor etc.) labeled by a suffix α. Let

$$\phi_\mu^\alpha = \frac{\partial \phi^\alpha}{\partial x_\mu} \tag{155}$$

The theory is fully described by an invariant function of position called the Lagrangian Density,

$$\mathscr{L} = \mathscr{L}\left(\phi^\alpha(x), \phi_\mu^\alpha(x)\right), \tag{156}$$

a function of ϕ^α and its *first* derivatives at the point x. The behaviour of the field is fixed by the *Action Principle*. If Ω is any finite or infinite region of space-time, then

$$I(\Omega) = \frac{1}{c} \int_\Omega \mathscr{L} \, d^4x \tag{157}$$

is stationary for the physically possible fields ϕ^α. Thus the variation $\varphi^\alpha \rightarrow \phi^\alpha + \delta\phi^\alpha$ produces no change in I to first order in $\delta\phi^\alpha$, if $\delta\phi^\alpha$ is an arbitrary variation equal to zero on the boundary of Ω.

It is always assumed that $\mathscr{L}$ is at most quadratic in the ϕ_μ^α and is in various other respects a well-behaved function.

Let Σ be the boundary of Ω, and $d\sigma$ an element of 3–dimensional volume on Σ, n_μ the outward unit vector normal to $d\sigma$, and

$$
d\sigma_\mu = n_\mu d\sigma, \quad \sum_\mu n_\mu^2 = -1 \quad \mu = 1, 2, 3, 4 \quad x_0 = ct
$$
$$
d\sigma_\mu = (dx_2\, dx_3\, dx_0,\; dx_1\, dx_3\, dx_0,\; dx_1\, dx_2\, dx_0,\; -i\, dx_1\, dx_2\, dx_3)
$$
(158)

Then

$$
c\,\delta I(\Omega) = \int_\Omega \sum_\alpha \left(\frac{\partial \mathscr{L}}{\partial \phi^\alpha} \delta\phi^\alpha + \sum_\mu \frac{\partial \mathscr{L}}{\partial \phi^\alpha_\mu} \delta\phi^\alpha_\mu \right) d^4x
$$

$$
= \int_\Omega \sum_\alpha \left\{ \frac{\partial \mathscr{L}}{\partial \phi^\alpha} - \sum_\mu \frac{\partial}{\partial x_\mu} \left(\frac{\partial \mathscr{L}}{\partial \phi^\alpha_\mu} \right) \right\} \delta\phi^\alpha d^4x + \int_\Sigma \sum_{\alpha,\mu} n_\mu \frac{\partial \mathscr{L}}{\partial \phi^\alpha_\mu} \delta\phi^\alpha d\sigma
$$
(159)

So the principle of action gives the field equations

$$
\frac{\partial \mathscr{L}}{\partial \phi^\alpha} - \sum_\mu \frac{\partial}{\partial x_\mu} \left(\frac{\partial \mathscr{L}}{\partial \phi^\alpha_\mu} \right) = 0
$$
(160)

defining the motion of the fields.

The quantity

$$
\pi_\alpha = \frac{1}{c} \sum_\mu n_\mu \frac{\partial \mathscr{L}}{\partial \phi^\alpha_\mu}
$$
(161)

is the momentum conjugate to ϕ^α, defined at x and with respect to the surface Σ.

A more general type of variation is made by varying not only the ϕ^α but also the boundary of Ω, each point x_μ being moved to the position $(x_\mu + \delta x_\mu)$ where δx_μ is either constant or may vary over the surface. Writing $_N\phi^\alpha$ for the new ϕ^α and $_O\phi^\alpha$ for the old one, we have

$$
\delta\phi^\alpha(x) = {}_N\phi^\alpha(x + \delta x) - {}_O\phi^\alpha(x)
$$

$$
{}_O\phi^\alpha(x + \delta x) = {}_O\phi^\alpha(x) + \sum_\mu \delta x_\mu\, {}_O\phi^\alpha_\mu(x)
$$
(162)

$$
\Delta\phi^\alpha(x) = {}_N\phi^\alpha(x) - {}_O\phi^\alpha(x)
$$

Therefore under the joint variation

$$
c\,\delta I(\Omega) = \int_{\Omega_N} \mathscr{L}\left({}_N\phi^\alpha(x),\, {}_N\phi^\alpha_\mu(x) \right) d^4x - \int_{\Omega_O} \mathscr{L}\left({}_O\phi^\alpha(x),\, {}_O\phi^\alpha_\mu(x) \right) d^4x
$$

$$
= \left\{ \int_{\Omega_N} - \int_{\Omega_O} \right\} \mathscr{L}\left({}_N\phi^\alpha(x),\, {}_N\phi^\alpha_\mu(x) \right) d^4x
$$

$$- \int_{\Omega_O} \left\{ \mathscr{L} \left({}_N\phi^\alpha(x), {}_N\phi^\alpha_\mu(x) \right) - \mathscr{L} \left({}_O\phi^\alpha(x), {}_O\phi^\alpha_\mu(x) \right) \right\} d^4x$$

$$= \int_\Sigma \sum_{\alpha,\mu} n_\mu \delta x_\mu \, \mathscr{L} \left({}_N\phi^\alpha(x), {}_N\phi^\alpha_\mu(x) \right) \, d\sigma + c \int_\Sigma \sum_\alpha \pi_\alpha(x) \Delta\phi^\alpha(x) \, d\sigma$$

the latter being true by (159) if we assume (160).

Now since by (162)

$$\delta\phi^\alpha(x) = {}_N\phi^\alpha(x) + \sum_\mu \delta x_\mu \, {}_N\phi^\alpha_\mu(x) - {}_O\phi^\alpha(x) = \Delta\phi^\alpha + \sum_\mu \delta x_\mu \, {}_N\phi^\alpha_\mu(x)$$

hence we get finally

$$\delta I(\Omega) = \int_\Omega \sum_{\alpha,\mu} \left\{ \pi_\alpha \delta\phi^\alpha + \left(\frac{1}{c} n_\mu \mathscr{L} - \phi^\alpha_\mu \pi^\alpha \right) \delta x_\mu \right\} d\sigma \qquad (163)$$

with all the new quantities on the RHS.

In the case which is physically of importance, the actual motion is fixed uniquely by specifying the values of the ϕ^α everywhere on two space-time surfaces σ_2 and σ_1 which are the past and future boundaries of the volume Ω. A space-like surface is one on which every two points are outside each other's light-cones, so that the fields can be fixed independently at every point.

Special case of non-relativistic theory, both σ_1 and σ_2 are just space at the time t_1 and t_2, and δx_μ is ic times a displacement of the time by δt_1 and δt_2. Then we may write $n_\mu = (0,0,0,i)$, $\pi_\alpha = \partial\mathscr{L}/\partial\dot\phi^\alpha$, and then for the Hamiltonian

$$H = \int d\tau \left(\sum_\alpha \pi_\alpha \dot\phi^\alpha - \mathscr{L} \right) \qquad (164)$$

and thus

$$\delta I(\Omega) = \int d\tau \sum_\alpha \left\{ (\pi_\alpha \delta\phi^\alpha)(t_1) - (\pi_\alpha \delta\phi^\alpha)(t_2) \right\} - \left\{ H(t_1)\,\delta t_1 - H(t_2)\,\delta t_2 \right\}$$

$$\qquad (165)$$

The essential feature of this classical theory is that the Action Principle is stated only for variations vanishing on the boundary of Ω. From this one can *deduce* as in (163) and (165) the effect on $I(\Omega)$ of variations not vanishing on the boundary. This is possible because each state of motion is defined by fixing *as many* field quantities as can be fixed independently, (e.g. all the

fields on two space-like surfaces or all the fields and their time-derivatives on one surface) and then the whole past and future of the motion is determined by the field equations.

Field equations can be written in the Hamiltonian form

$$\dot{\phi}^\alpha = \frac{\partial H}{\partial \pi_\alpha}, \qquad \dot{\pi}_\alpha = -\frac{\partial H}{\partial \phi^\alpha}$$

Examples:

1. Klein-Gordon Field, real

$$\mathscr{L}_K = -\frac{1}{2}c^2 \left\{ \sum_\mu \left(\frac{\partial \psi}{\partial x_\mu} \right)^2 + \mu^2 \psi^2 \right\} \tag{166}$$

2. Klein–Gordon Field, complex

$$\mathscr{L}_K' = -c^2 \left\{ \sum_\mu \left(\frac{\partial \psi}{\partial x_\mu} \frac{\partial \psi^*}{\partial x_\mu} \right) + \mu^2 \psi \psi^* \right\} \tag{167}$$

where we consider ψ and ψ^* independent one-component fields.

3. Maxwell Field, four component A_μ, Fermi form,

$$\mathscr{L}_M = -\frac{1}{4} \sum_{\mu,\nu} \left(\frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} \right)^2 - \frac{1}{2} \sum_\mu \left(\frac{\partial A_\mu}{\partial x_\mu} \right)^2 \tag{168}$$

4. Dirac Field

$$\mathscr{L}_D = -\hbar c \, \overline{\psi} \left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} + \mu \right) \psi \qquad \mu = \frac{mc}{\hbar} \tag{169}$$

5. Dirac Field interacting with Maxwell Field

$$\mathscr{L}_Q = \mathscr{L}_D + \mathscr{L}_M - \sum_\lambda ie A_\lambda \overline{\psi} \gamma_\lambda \psi \tag{170}$$

here Q stands for quantum electrodynamics.

Problem 5. Work out these examples: find the field equations, the momentum conjugate to each component of the field, and the Hamiltonian function, (the momenta and Hamiltonian defined for the case of a flat space σ only). Verify that the Hamiltonian gives a correct canonical representation of the field equations as Hamiltonian equations of motion.

4.2   Quantum Relativistic Field Theory

The classical relativistic field theories were usually quantized by using the Hamiltonian form of the field equations and bringing in the commutation relations between coordinates and momenta taken from non-relativistic quantum mechanics. For this approach see Wentzel's book. It is a very bad method, it is complicated; and it is not at all obvious or even easy to prove that the theory so made is relativistic, because the whole Hamiltonian approach is non-covariant.

Just recently we learnt a much better way of doing it, which I shall now expound in these lectures. It is due to Feynman and Schwinger.[16]

References: R. P. Feynman,   *Rev. Mod. Phys.* **20** (1948) 367
Phys. Rev. **80** (1950) 440
J. Schwinger,   *Phys. Rev.* **82** (1951) 914

It is relativistic all the way, and it is much simpler than the old methods. It is based directly on the Action Principle form of the classical theory which I have just given you, not the Hamiltonian form.

In the quantum theory the ϕ^α are operators defined at each point of space-time as before. They satisfy the same field equations as before, and this is ensured if we assume that the Action Principle

$$\delta I(\Omega) = 0$$
$$I(\Omega) = \frac{1}{c} \int_\Omega \mathscr{L}\left(\phi^\alpha, \phi^\alpha_\mu\right) d^4x \tag{171}$$

holds for all variations $\delta\phi^\alpha$ of the operators vanishing on the boundaries of Ω.

In the quantum theory, because of complementarity relations, it is not possible to give numerical values to all field operators throughout a physical motion. In fact the state of motion is specified by giving numerical values to the ϕ^α on *one* space-like surface. The future of the state of motion cannot then be determined from the field equations, which are in general second-order differential equations. Therefore the action principle (171) which was enough for the classical theory is no longer enough. We must make some *additional* statement about the behaviour of δI for variations $\delta\phi^\alpha$ which are not zero on the boundaries of Ω.

A state of motion is specified by specifying a space-time surface σ and a set of numerical values ϕ'^α for the eigenvalue which the operators ϕ^α on σ

have in this state. The state is denoted by the Dirac ket vector $|\phi'^{\alpha}, \sigma\rangle$. This is a special kind of state in which the ϕ^{α} on σ have eigenvalues: the general state is a linear combination of $|\phi'^{\alpha}, \sigma\rangle$ with various values of ϕ'^{α}. The physically observable quantities are expressions such as the matrix element

$$\left\langle \phi_1'^{\alpha}, \sigma_1 \middle| \phi^{\beta}(x) \middle| \phi_2'^{\alpha}, \sigma_2 \right\rangle \tag{172}$$

of the field operator $\phi^{\beta}(x)$ between the two states specified by $\phi_1'^{\alpha}$ on σ_1 and by $\phi_2'^{\alpha}$ on σ_2. In particular, the transition probability amplitude between the two states is

$$\left\langle \phi_1'^{\alpha}, \sigma_1 \middle| \phi_2'^{\alpha}, \sigma_2 \right\rangle \tag{173}$$

The squared modulus of this gives the probability of finding the values $\phi_1'^{\alpha}$ for the fields on σ_1, in the motion which is defined by the fields being given the definite values $\phi_2'^{\alpha}$ on σ_2.

4.3   The Feynman Method of Quantization

The Feynman method of quantizing the theory consists in writing down an explicit formula for the transition amplitude (173). Namely

$$\left\langle \phi_1'^{\alpha}, \sigma_1 \middle| \phi_2'^{\alpha}, \sigma_2 \right\rangle = N \sum_H \exp\left\{ \frac{i}{\hbar} I_H(\Omega) \right\} \tag{174}$$

Here H represents a *History* of the fields between σ_2 and σ_1, i.e. any set of classical functions $\phi^{\alpha}(x)$ which are defined in the region Ω between σ_2 and σ_1 and which take the values $\phi_1'^{\alpha}$ on σ_1 and $\phi_2'^{\alpha}$ on σ_2. $I_H(\Omega)$ is the value of $I(\Omega)$ calculated with these particular functions. The sum $\sum_H$ is taken over all possible histories, a continuously infinite sum whose exact mathematical definition is not easy to formulate. N is a normalization factor independent of the particular states considered, chosen so as to make the sum of the squares of the amplitudes from a given state to all other states equal to 1. This formula is derived by Feynman from very general considerations, applying a Huygen's principle to the solution of wave-mechanics just as it is done in wave-optics. By this one formula the whole theory is quantized and the answer to any physical problem in principle given. The method applies not only to field theory but to ordinary non-relativistic quantum theory too. We do not try to derive or justify the Feynman formula here. We just show that it gives the same results as the usual quantum mechanics.

For a discussion of the difficulties in defining the sum $\sum_H$, and a method of doing it in simple cases, see C. Morette, *Phys. Rev.* **81** (1951) 848.

From formula (174) we derive at once the most general *Correspondence Principle* giving us back the classical theory in the limit as $\hbar \to 0$. For suppose $\hbar \to 0$ then the exponential factor in (174) becomes an extremely rapidly oscillating function of H for all histories H *except* that one for which $I(\Omega)$ is stationary. Therefore in the limit the sum $\sum_H$ reduces to the contribution from the classical motion leading from $\phi_2'^\alpha$ on σ_2 to $\phi_1'^\alpha$ on σ_1, all other contributions interfering destructively. The classical motion is defined by the condition that[17] its $\delta I(\Omega) = 0$ for all small variations of the ϕ^α between σ_2 and σ_1. This passage to the classical theory is precisely analogous to the passage from wave-optics to geometrical optics when the wave-length of light is allowed to tend to zero. The WKB approximation is gotten by taking $\hbar$ small but not quite zero.

To establish a connection between the Feynman method and the ordinary method of quantization, Feynman has to define what he means by an operator in his formulation. This he does as follows: Let x be any space-time point inside Ω. Let $\mathcal{O}(x)$ be any field operator defined at x, for example $\phi^\beta(x)$ or $\phi_\mu^\beta(x)$. Then $\mathcal{O}(x)$ is given a meaning by defining its matrix element between the states $|\phi_2'^\alpha, \sigma_2\rangle$ and $|\phi_1'^\alpha, \sigma_1\rangle$, where σ_2 and σ_1 are any two surfaces to the past and future of x. This matrix element is

$$\langle \phi_1'^\alpha, \sigma_1 \big| \mathcal{O}(x) \big| \phi_2'^\alpha, \sigma_2 \rangle = N \sum_H \mathcal{O}_H(x) \exp\left\{ \frac{i}{\hbar} I_H(\Omega) \right\} \qquad (175)$$

The number $\mathcal{O}_H$ is just the value which the expression $\mathcal{O}$ takes when the ϕ^α are given the values which they have in the history H. It is easily verified that the definitions (174) and (175) are physically reasonable and give the right formal properties of transition amplitudes and operator matrix elements.

The Feynman method has one fatal drawback: we cannot use it until we have some way of calculating or at least using the sums over histories, and so far nobody has suggested a practical way of doing this. But Schwinger has shown how to derive from the Feynman method an Action Principle formulation of the theory which avoids this difficulty.

4.4   The Schwinger Action Principle

Let the sets of eigenvalues $\phi_1'^\alpha$ and $\phi_2'^\alpha$ in (174) be held fixed. Let the numbers $\phi_H^\alpha(x)$ be varied so that $\phi_H^\alpha(x)$ is replaced by $\phi_H^\alpha(x) + \delta\phi^\alpha(x)$ where

$\delta\phi^\alpha(x)$ is an arbitrary infinitesimal c-number quantity. Let the surfaces σ_1 and σ_2 be varied so that the point x_μ moves to $x_\mu + \delta x_\mu$. And let the function $\mathscr{L}$ also be varied so that it is replaced by $\mathscr{L} + \delta\mathscr{L}$ where $\delta\mathscr{L}$ is any expression involving the ϕ^α and ϕ^α_μ. Under this triple variation (174) gives

$$\delta\left\langle\phi'^\alpha_1, \sigma_1 \,\middle|\, \phi'^\alpha_2, \sigma_2\right\rangle = N\sum_H \left\{\frac{i}{\hbar}\delta I_H(\Omega)\exp\left(\frac{i}{\hbar}I_H(\Omega)\right)\right\} \tag{176}$$

Using (175) this may be written

$$\delta\left\langle\phi'^\alpha_1, \sigma_1 \,\middle|\, \phi'^\alpha_2, \sigma_2\right\rangle = \frac{i}{\hbar}\left\langle\phi'^\alpha_1, \sigma_1 \,\middle|\, \delta I(\Omega) \,\middle|\, \phi'^\alpha_2, \sigma_2\right\rangle. \tag{177}$$

Here $\delta I(\Omega)$ is the operator obtained by making the three variations on the operator $I(\Omega)$. Formally $\delta I(\Omega)$ is the same as the variation obtained in the classical theory,

$$\delta I(\Omega) = \frac{1}{c}\int_\Omega \left\{\delta\mathscr{L} + \sum_{\alpha,\mu}\left(\frac{\partial\mathscr{L}}{\partial\phi^\alpha} - \frac{\partial}{\partial x_\mu}\frac{\partial\mathscr{L}}{\partial\phi^\alpha_\mu}\right)\delta\phi^\alpha\right\}d^4x$$

$$+ \left\{\int_{\sigma_1} - \int_{\sigma_2}\right\}\sum_{\alpha,\mu}\left\{\pi_\alpha\delta\phi^\alpha + \left(\frac{1}{c}n_\mu\mathscr{L} - \phi^\alpha_\mu\pi_\alpha\right)\delta x_\mu\right\}d\sigma \tag{178}$$

Only now everything on the RHS of (178) is an operator.

Now what is the meaning of this triple variation as applied to the left side of (174)? Since the $\phi^\alpha_H(x)$ are only variables of summation, the change from $\phi^\alpha_H(x)$ to $\phi^\alpha_H(x) + \delta\phi^\alpha(x)$ only affects the left side in changing the boundary values which $\phi^\alpha_H(x)$ must take on σ_1 and σ_2. Thus instead of $\phi^\alpha_H(x) = \phi^{\alpha\prime\prime}_1(x)$ on σ_1 we now have the new summation variable

$$\phi^\alpha_H(x) + \delta\phi^\alpha = \phi^{\alpha\prime\prime}_1(x) + \delta\phi^\alpha \quad \text{on } \sigma_1$$

Therefore the change in ϕ^α_H is equivalent simply to changing

$$\phi^{\alpha\prime\prime}_1 \quad \text{to} \quad \phi^{\alpha\prime\prime}_1 + \delta\phi^\alpha \quad \text{on } \sigma_1$$
$$\phi^{\alpha\prime}_2 \quad \text{to} \quad \phi^{\alpha\prime}_2 + \delta\phi^\alpha \quad \text{on } \sigma_2.$$

The change in $\mathscr{L}$ and in the position of σ produces a change in the left side of (174), by virtue of the change in the operators ϕ^α on σ_1 and σ_2 resulting from the variations $\delta\mathscr{L}$ and δx_μ in consequence of the field equations.

Hence the net result of the triple variation on the left side of (174) is to give the change in the matrix element $\left\langle\phi^{\alpha\prime\prime}_1, \sigma_1 \,\middle|\, \phi^{\alpha\prime}_2, \sigma_2\right\rangle$ if the $\phi^{\alpha\prime\prime}_1$ and $\phi^{\alpha\prime}_2$ are

left fixed, the operators $\phi^\alpha(x)$ on σ_1 and σ_2 being modified in consequence of the variations $\delta\mathscr{L}$ and δx_μ according to the field equations, and in addition $\phi^\alpha(x)$ on σ_1 and σ_2 being changed to $\phi^\alpha(x) - \delta\phi^\alpha(x)$.

Schwinger takes equation (177) as the fundamental principle for setting up the quantum theory. Thus he gets rid of the unpleasant $\sum_H$. Out of this action principle come very simply all the main features of a quantum field theory below.

4.4.1 The Field Equations

If we take the special case of a variation $\delta\phi^\alpha$ which vanishes on the boundary of Ω and $\delta\mathscr{L} = \delta x_\mu = 0$, then $\langle \phi_1^{\alpha\prime}, \sigma_1 | \phi_2^{\alpha\prime}, \sigma_2 \rangle$ depends only on the operators ϕ^α on σ_1 and σ_2 and is unaffected by the variation. Therefore for all such variations

$$\delta I(\Omega) = 0 \tag{171}$$

$$\frac{\partial \mathscr{L}}{\partial \phi^\alpha} - \sum_\mu \frac{\partial}{\partial x_\mu} \frac{\partial \mathscr{L}}{\partial \phi^\alpha_\mu} = 0 \tag{179}$$

That is to say, the classical action principle and the classical field equations are valid for the quantum field operators.

We see that (177) is exactly the kind of generalization we want of the old variation principle (171). It includes the information, necessary for a quantum theory, concerning the effect on $I(\Omega)$ of variations not vanishing at the boundary of Ω.

4.4.2 The Schrödinger Equation for the State-function

Specialize σ_1 and σ_2 to be the whole space at the times t_1 and t_2. Then

$$\langle \phi_1^{\prime\alpha}, \sigma_1 | \phi_2^{\prime\alpha}, \sigma_2 \rangle = \langle \phi_1^{\prime\alpha}, t_1 | \phi_2^{\prime\alpha}, t_2 \rangle = \Psi\left(\phi_1^{\prime\alpha}, t_1\right)$$

is a Schrödinger wave-function giving the probability amplitude for finding the system in the state $\phi_1^{\prime\alpha}$ at the time t_1, given the initial conditions $\phi_2^{\prime\alpha}$ at t_2. The development of $\Psi\left(\phi_1^{\prime\alpha}, t_1\right)$ with time t_1 is thus a description of the development of the state of the system with time in the Schrödinger representation.

Take in (177) a variation in which $\delta\phi^\alpha = \delta\mathscr{L} = 0$, the surface σ_1 being just moved through the displacement δt in the time direction. Then using (165) and (164)

$$\delta\Psi\left(\phi_1^{\prime\alpha}, t_1\right) = -\frac{i}{\hbar} \langle \phi_1^{\prime\alpha}, t_1 | H\left(t_1\right) | \phi_2^{\prime\alpha}, t_2 \rangle \delta t_1$$

or

$$i\hbar \frac{d}{dt} \left\langle \phi_1'^{\,\alpha}, t_1 \,\middle|\, \phi_2'^{\,\alpha}, t_2 \right\rangle = \left\langle \phi_1'^{\,\alpha}, t_1 \,\middle|\, H\left(t_1\right) \,\middle|\, \phi_2'^{\,\alpha}, t_2 \right\rangle \tag{180}$$

This is the ordinary Schrödinger equation in Dirac's notation. It shows that the Schwinger action principle contains enough information for predicting the future behaviour of a system given initially in a known quantum state.

4.4.3   *Operator Form of the Schwinger Principle*

Feynman defined operators by giving the formula (175) for their matrix elements between states specified on two different surfaces. The initial state had to be specified in the past, the final state in the future, the operator referring to some particular time which is taken as present.

The usual and generally more useful way of defining operators is to specify their matrix elements between states defined on the same surface. Thus we are interested in a matrix element

$$\left\langle \phi'^{\,\alpha}, \sigma \,\middle|\, \mathcal{O} \,\middle|\, \phi''^{\,\alpha}, \sigma \right\rangle \tag{181}$$

where $\phi'^{\,\alpha}$ and $\phi''^{\,\alpha}$ are given sets of eigenvalues and σ is a surface which may be past, present or future in relation to the field-points to which $\mathcal{O}$ refers.

Suppose that a reference surface σ_o is chosen in the remote past. Let the ϕ^α, σ and $\mathcal{L}$ be varied in such a way that everything on σ_o remains fixed. For such a variation, (178) gives if we assume that (179) holds

$$\delta I(\Omega) = \frac{1}{c} \int_\Omega \delta\mathcal{L}\, d^4x + \int_\sigma \sum_{\alpha,\mu} \left\{ \pi_\alpha \delta\phi^\alpha + \left(\frac{1}{c} n_\mu \mathcal{L} - \phi_\mu^\alpha \pi_\alpha \right) \delta x_\mu \right\} d\sigma \tag{182}$$

where Ω is the region bounded by σ_o and σ. Let us now first calculate the variation of (181) arising from the change in the meaning of the states $|\phi'^{\,\alpha}, \sigma\rangle$ and $|\phi''^{\,\alpha}, \sigma\rangle$. The operator $\mathcal{O}$ itself is at this point fixed and not affected by the variations in ϕ^α, σ and $\mathcal{L}$. Then

$$\left\langle \phi'^{\,\alpha}, \sigma \,\middle|\, \mathcal{O} \,\middle|\, \phi''^{\,\alpha}, \sigma \right\rangle$$
$$= \sum_{\phi_o'} \sum_{\phi_o''} \left\langle \phi'^{\,\alpha}, \sigma \,\middle|\, \phi_o'^{\,\alpha}, \sigma_o \right\rangle \left\langle \phi_o'^{\,\alpha}, \sigma_o \,\middle|\, \mathcal{O} \,\middle|\, \phi_o''^{\,\alpha}, \sigma_o \right\rangle \left\langle \phi_o''^{\,\alpha}, \sigma_o \,\middle|\, \phi''^{\,\alpha}, \sigma \right\rangle \tag{183}$$

therefore, denoting

$$\left\langle \phi'^{\,\alpha}, \sigma \,\middle|\, \mathcal{O} \,\middle|\, \phi''^{\,\alpha}, \sigma \right\rangle = \left\langle \sigma' \,\middle|\, \mathcal{O} \,\middle|\, \sigma'' \right\rangle \quad \text{etc., we have}$$

$$\delta \langle \sigma' | \mathcal{O} | \sigma'' \rangle = \sum_{'} \sum_{''} \left(\delta \langle \sigma' | \sigma'_o \rangle \right) \langle \sigma'_o | \mathcal{O} | \sigma''_o \rangle \langle \sigma''_o | \sigma'' \rangle$$

$$+ \sum_{'} \sum_{''} \langle \sigma' | \sigma'_o \rangle \langle \sigma'_o | \mathcal{O} | \sigma''_o \rangle \left(\delta \langle \sigma''_o | \sigma'' \rangle \right)$$

because $|\phi_o'^{\alpha}\rangle$ and $|\phi_o''^{\alpha}\rangle$ are not changed by the variation, and neither is $\mathcal{O}$. Therefore, using (177) we have

$$\delta \langle \sigma' | \mathcal{O} | \sigma'' \rangle = \sum_{'} \sum_{''} \frac{i}{\hbar} \langle \sigma' | \delta I_{\sigma - \sigma_o} \mathcal{O} | \sigma'' \rangle + \sum_{'} \sum_{''} \frac{i}{\hbar} \langle \sigma' | \mathcal{O} \, \delta I_{\sigma_o - \sigma} | \sigma'' \rangle$$

where the subscript $\sigma - \sigma_o$ refers to the surface integrals in (178). Since $\delta I_{\sigma - \sigma_o} = -\delta I_{\sigma_o - \sigma}$, we get finally

$$\delta \langle \phi'^{\alpha}, \sigma | \mathcal{O} | \phi''^{\alpha}, \sigma \rangle = \frac{i}{\hbar} \langle \phi'^{\alpha}, \sigma | [\, \delta I(\Omega), \mathcal{O} \,] | \phi''^{\alpha}, \sigma \rangle \qquad (184)$$

where $[\, P, R \,] = PR - RP$. This applies for the case when $\mathcal{O}$ is fixed and the states vary.

Now we want to calculate the variation of $\langle \phi'^{\alpha}, \sigma | \mathcal{O} | \phi''^{\alpha}, \sigma \rangle$ for the case when the states are fixed, and $\mathcal{O} = \mathcal{O}(\phi^{\alpha}(\sigma))$ changes. This, however, will be the same as for the previous case, except with the opposite sign, because the variation of the matrix element[18]

$$\langle \phi'^{\alpha}, \sigma | \mathcal{O} | \phi''^{\alpha}, \sigma \rangle \qquad (185)$$

if *both* the states and $\mathcal{O}$ change simultaneously is zero. Therefore, if we use a representation in which matrix elements of $\mathcal{O}$ are defined between states *not* subject to variation we get[19]

$$i\hbar \, \delta \mathcal{O}(\sigma) = [\, \delta I(\Omega), \mathcal{O}(\sigma) \,] \qquad (186)$$

This is the Schwinger action principle in operator form. It is related to (177) exactly as the Heisenberg representation is to the Schrödinger representation in elementary quantum mechanics.

4.4.4   *The Canonical Commutation Laws*

Taking for σ the space at time t, for $\mathcal{O}(\sigma)$ the operator $\phi^{\alpha}(r, t)$ at the space-point r, and $\delta x_{\mu} = \delta \mathcal{L} = 0$ we have by (182) and (186) for an arbitrary variation $\delta \phi^{\alpha}$

$$-i\hbar \, \delta \phi^{\alpha}(r, t) = \sum_{\beta} \int [\, \pi_{\beta}(r', t) \, \delta \phi^{\beta}(r', t), \, \phi^{\alpha}(r, t) \,] \, d^3 r' \qquad (187)$$

because $d\sigma = -n_\mu \, d\sigma_\mu = -i(-i \, dx'_1 dx'_2 dx'_3) = -d^3 r'$ by (158); the unit vector in the *increasing* time direction is i, and this is the outward direction since we choose σ_o in the past. Hence for every r, r'

$$[\,\phi^\alpha(r,t),\, \pi_\beta(r',t)\,] = i\hbar \, \delta_{\alpha\beta} \, \delta^3(\boldsymbol{r} - \boldsymbol{r}') \tag{188}$$

Also since the $\phi^\alpha(r)$ on σ are assumed independent variables,

$$[\,\phi^\alpha(r,t),\, \phi^\beta(r',t)\,] = 0 \tag{189}$$

So this method gives automatically the correct canonical commutation laws for the fields. There is no need to prove that the commutation rules are consistent with the field equations, as was necessary in the older methods.

4.4.5   *The Heisenberg Equation of Motion for the Operators*

Suppose that σ is a flat surface at time t, and that a variation is made by moving the surface through the small time δt as in B above. But now let $\mathcal{O}(t) = \mathcal{O}(\sigma)$ be an operator built up out of the field-operators ϕ^α on σ. Then by (165) and (186) the change in $\mathcal{O}(t)$ produced by the variation is given by

$$i\hbar \, \delta\mathcal{O}(t) = [\, -H(t) \, \delta t,\, \mathcal{O}(t)\,]$$

That is to say, $\mathcal{O}(t)$ satisfies the Heisenberg equation of motion

$$i\hbar \, \frac{d\mathcal{O}(t)}{dt} = [\, \mathcal{O}(t),\, H(t)\,] \tag{190}$$

where $H(t)$ is the total Hamiltonian operator.

4.4.6   *General Covariant Commutation Laws*

From (186) we derive at once the general covariant form of the commutation laws discovered by Peierls in 1950 [13]. This covariant form is not easy to reach in the Hamiltonian formalism.

Let two field points z and y be given, and two operators $\mathcal{R}(z)$ and $\mathcal{Q}(y)$ depending on the field quantities ϕ^α at z and y. Let a reference surface σ_o be fixed, past of both z and y. Suppose the quantity,

$$\delta_\mathcal{R}(\mathscr{L}) = \epsilon \, \delta^4(x - z)\mathcal{R}(z) \tag{191}$$

is added to the Lagrangian density $\mathscr{L}(x)$, where ϵ is an infinitesimal c-number. This will make at most a certain infinitesimal change $\epsilon \, \delta_\mathcal{R} \phi^\alpha(x)$ in

the solutions $\phi^\alpha(x)$ of the field equations. Supposing the new $\phi^\alpha(x)$ to be identical with the old one on σ_o then $\delta_{\mathcal{R}}\phi^\alpha(x)$ is different from zero only in the future light-cone of z.

Similarly adding

$$\delta_{\mathcal{Q}}(\mathscr{L}) = \epsilon\,\delta^4(x-y)\mathcal{Q}(y) \tag{192}$$

to $\mathscr{L}(x)$ produces at most a change $\epsilon\,\delta_{\mathcal{Q}}\phi^\alpha(x)$ in the $\phi^\alpha(x)$. Let $\epsilon\,\delta_{\mathcal{R}}\mathcal{Q}(y)$ be the change in $\mathcal{Q}(y)$ produced by the addition (191), while $\epsilon\,\delta_{\mathcal{Q}}\mathcal{R}(z)$ be the change in $\mathcal{R}(z)$ produced by (192). Suppose y lies on a surface σ lying in the future of z. Then we take $\mathcal{Q}(y)$ for $\mathcal{O}(\sigma)$ in (186), and $\delta\mathscr{L}$ given by (191). The $\delta I(\Omega)$ given by (182) reduces then simply to

$$\delta I(\Omega) = \frac{1}{c}\,\epsilon\,\mathcal{R}(z)$$

It is assumed that there is no intrinsic change $\delta\phi^\alpha$ of ϕ^α or δx_μ of σ apart from the change whose effect is already included in the $\delta\mathscr{L}$ term. Thus (186) gives

$$\begin{aligned}
[\,\mathcal{R}(z),\,\mathcal{Q}(y)\,] &= i\hbar c\,\delta_{\mathcal{R}}\mathcal{Q}(y) \quad (y_0 > z_0) \\
[\,\mathcal{R}(z),\,\mathcal{Q}(y)\,] &= -i\hbar c\,\delta_{\mathcal{Q}}\mathcal{R}(z) \quad (z_0 > y_0)
\end{aligned} \tag{193}$$

When y and z are separated by a space-like interval, the commutator is zero, because the disturbance $\mathcal{R}(z)$ propagates with a velocity at most c and therefore can affect things only in the future lightcone of z; this means $\delta_{\mathcal{R}}\mathcal{Q}(y) = 0$ in this case.

Peierls' formula, valid for any pair of field operators, is

$$[\,\mathcal{R}(z),\,\mathcal{Q}(y)\,] = i\hbar c\,\{\delta_{\mathcal{R}}\mathcal{Q}(y) - \delta_{\mathcal{Q}}\mathcal{R}(z)\} \tag{194}$$

This is a useful formula for calculating commutators in a covariant way.

4.4.7 *Anticommuting Fields*

There is one type of field theory which can be constructed easily by Schwinger's action principle, but which does not come out of Feynman's picture. Suppose a classical field theory in which a group of field operators ψ^α always occurs in the Lagrangian in bilinear combinations like $\overline{\psi}^\beta\psi^\alpha$ with the group of field operators $\overline{\psi}$. Examples, the Dirac $\mathscr{L}_D$ and the quantum electrodynamics $\mathscr{L}_Q$.

Then instead of taking every ϕ^α on a given surface σ to commute as in (189), we may take every pair of ψ^α to anticommute, thus[20]

$$\{\psi^\alpha(r,t),\psi^\beta(r',t)\} = 0 \qquad \{\mathcal{P},\mathcal{R}\} = \mathcal{P}\mathcal{R} + \mathcal{R}\mathcal{P} \tag{195}$$

The bilinear combination will still commute, like the ϕ^α's did before. The ψ^α commute as before with any field quantities on σ other than the ψ and $\overline{\psi}$. Schwinger then assumes (177) to hold precisely as before, except that in calculating $\delta I(\Omega)$ according to (178), the variation $\delta\psi^\alpha$ *anticommutes* with all operators ψ^α and $\overline{\psi}^\beta$. In these theories it turns out that the momentum π_α conjugate to ψ^α is just a linear combination of $\overline{\psi}$, because the Lagrangian is only linear in the derivatives of ψ. With the anticommuting fields the field equations (179) are deduced as before, also the Schrödinger equation (180), the commutation rules being given by (186) and (187). But now in order to make (187) valid, since $\delta\psi^\beta$ anticommutes with the ψ and π operators,[21] the canonical commutation law must be written

$$\{\,\psi^\alpha(r,t),\,\pi_\beta(r',t)\,\} = -i\hbar\,\delta_{\alpha\beta}\,\delta^3(\boldsymbol{r} - \boldsymbol{r}') \tag{196}$$

The general commutation rule (194) is still valid provided that $\mathcal{Q}$ and $\mathcal{R}$ are also expressions bilinear in the $\overline{\psi}$ and ψ.

The interpretation of the operators, and the justification for the Schwinger principle, in the case of anticommuting fields, is not clear. But it is clear that the Schwinger principle in this case gives a consistent and simple formulation of a relativistic quantum field theory. And we may as well take advantage of the method, even if we do not quite understand its conceptual basis. The resulting theory is mathematically unambiguous, and gives results in agreement with experiment; that should be good enough.

Examples of Quantized Field Theories

5.1 The Maxwell Field

Lagrangian

$$\mathscr{L}_M = -\frac{1}{4}\sum_{\mu,\nu}\left(\frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu}\right)^2 - \frac{1}{2}\sum_\mu\left(\frac{\partial A_\mu}{\partial x_\mu}\right)^2 \tag{168}$$

Field equations

$$\sum_\mu \frac{\partial^2}{\partial x_\mu^2}A_\lambda = \Box^2 A_\lambda = 0 \tag{197}$$

Commutation rules for the A_λ: To find these we uses the Peierls method. Take two points y and z with $z_0 > y_0$. Let $\mathcal{Q}(y) = A_\lambda(y)$, $\mathcal{R}(z) = A_\mu(z)$. Note: in this section x, y, z, k etc. are meant to have components 1, 2, 3 and 0, while in x_μ, y_μ, z_μ, k_μ, etc. we mean $\mu = 1, 2, 3$ and 4. When $\delta_{\mathcal{Q}}(\mathscr{L}) = \epsilon\, \delta^4(x-y)A_\lambda(y)$ is added to $\mathscr{L}_M$, the field equation for A_μ becomes

$$\Box^2 A_\mu + \delta_{\lambda\mu}\,\epsilon\,\delta^4(x-y) = 0 \tag{198}$$

This equation is satisfied by $A_\mu + \delta_{\mathcal{Q}}A_\mu(z)$ (by definition), and hence also by $\delta_{\mathcal{Q}}A_\mu(z)$ because of (197). Therefore $\delta_{\mathcal{Q}}A_\mu(z)$ is defined by the conditions

$$\Box^2(\delta_{\mathcal{Q}}A_\mu(z)) = -\delta_{\lambda\mu}\,\delta^4(z-y)$$
$$\delta_{\mathcal{Q}}A_\mu(z) = 0 \quad \text{for } z_0 < y_0\,. \tag{199}$$

That is to say, $\delta_Q A_\mu(z)$ is a c-number and is the retarded potential created by a point source acting instantaneously at the space-time point y.

$$\delta_Q A_\mu(z) = \delta_{\lambda\mu} D_R(z - y)$$
$$\square^2 D_R(z - y) = -\delta^4(z - y)$$
(200)

If x is any 4-vector, using

$$\delta(x^2 - a^2) = \frac{1}{2a} \left\{ \delta(x - a) + \delta(x + a) \right\}, \quad a > 0$$

we get

$$D_R(x) = \frac{1}{2\pi} \Theta(x) \, \delta(x^2) = \frac{1}{4\pi|\boldsymbol{r}|} \, \delta(x_0 - |\boldsymbol{r}|)$$
(201)

Here

$$|\boldsymbol{r}| = \sqrt{x_1^2 + x_2^2 + x_3^2}; \qquad x_o = ct \,;$$

$$x^2 = r^2 - x_o^2 \qquad \Theta(x) = \begin{cases} +1 & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases}$$

In the same way

$$\delta_{\mathcal{R}} A_\lambda(y) = \delta_{\lambda\mu} D_A(z - y) = \delta_{\lambda\mu} D_R(y - z)$$
(202)

where D_A is the advanced potential of the same source,

$$D_A(x) = \frac{1}{4\pi|\boldsymbol{r}|} \, \delta(x_0 + |\boldsymbol{r}|)$$

Hence we have the the commutation rule (194)

$$[\, A_\mu(z), \, A_\lambda(y) \,] = i\hbar c \, \delta_{\lambda\mu} \, [D_A(z - y) - D_R(z - y)]$$

$$= i\hbar c \, \delta_{\lambda\mu} D(z - y) \quad \text{(definition of } D) \tag{203}$$

This invariant D-function satisfies by (200)

$$\square^2 D(x) = -\delta^4(x) - (-\delta^4(x)) = 0$$
(204)

as it must. Also

$$D(x) = \frac{1}{4\pi|\boldsymbol{r}|} \, [\delta(x_0 + |\boldsymbol{r}|) - \delta(x_0 - |\boldsymbol{r}|)]$$

$$= -\frac{1}{2\pi} \, \epsilon(x) \, \delta(x^2) \qquad \epsilon(x) = \text{sign}(x_0) \tag{205}$$
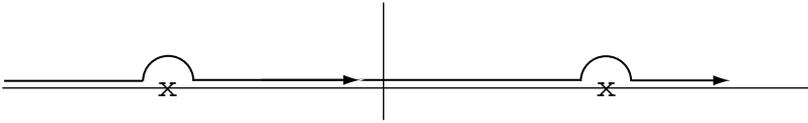
5.1.1 Momentum Representations

We have

$$\delta^4(x) = \frac{1}{(2\pi)^4} \int \exp(ik \cdot x) \, d^4k \tag{206}$$

where the integral is fourfold, over $dk_1 \, dk_2 \, dk_3 \, dk_4$. Therefore

$$D_R(x) = \frac{1}{(2\pi)^4} \int_+ \exp(ik \cdot x) \frac{1}{k^2} \, d^4k \tag{207}$$

where $k^2 = |\boldsymbol{k}|^2 - k_0^2$. The integration with respect to k_1, k_2, k_3 is an ordinary real integral. That with respect to k_0 is a contour integral going along the real axis and *above* the two poles at $k_0 = \pm|\boldsymbol{k}|$.
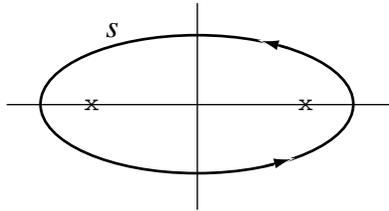

For detailed calculations, see the Appendix below. This gives the correct behaviour of D_R being zero for $x_0 < 0$. Similarly

$$D_A(x) = \frac{1}{(2\pi)^4} \int_- \exp(ik \cdot x) \frac{1}{k^2} \, d^4k \tag{208}$$

with a contour going *below* both the poles. Therefore

$$D(x) = \frac{1}{(2\pi)^4} \int_s \exp(ik \cdot x) \frac{1}{k^2} \, d^4k \tag{209}$$

with a contour s as shown.
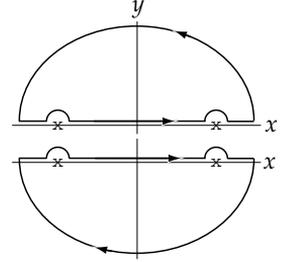

Evaluating the residues we have

$$D(x) = -\frac{i}{(2\pi)^3} \int \exp(ik \cdot x) \, \delta(k^2) \, \epsilon(k) \, d^4k \tag{210}$$

this being now an ordinary real integral.

Appendix

Let us verify, e.g. (207). For $x_0 < 0$, we have to take the top path shown, otherwise the integrand blows up; this gives obviously 0.

For $x_0 > 0$, we have to take the bottom path; then

$$D_R(x) = \frac{1}{(2\pi)^4} \int_+ \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{x}}e^{-ik_0x_0}}{\boldsymbol{k}^2 - k_0^2}\, d^3k\, dk_0$$

where $\boldsymbol{k}$ and $\boldsymbol{x}$ are 3 dimensional vectors.

Now because of the clockwise direction

$$\int_+ \frac{e^{-ik_0x_0}}{\boldsymbol{k}^2 - k_0^2}\, dk_0 = -\int_+ \frac{e^{-ik_0x_0}}{(k_0 - |\boldsymbol{k}|)(k_0 + |\boldsymbol{k}|)}\, dk_0$$

$$= 2\pi i(\text{Residue at } k_0 = |\boldsymbol{k}| + \text{Residue at } k_0 = -|\boldsymbol{k}|)$$

$$= 2\pi i\left(\frac{e^{-i|\boldsymbol{k}|x_0}}{2|\boldsymbol{k}|} - \frac{e^{+i|\boldsymbol{k}|x_0}}{2|\boldsymbol{k}|}\right)$$

Hence

$$D_R(x) = \frac{i}{(2\pi)^3} \int \frac{1}{2|\boldsymbol{k}|} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\left\{e^{i|\boldsymbol{k}|x_0} - e^{-i|\boldsymbol{k}|x_0}\right\} d^3k$$

$$= \frac{i}{(2\pi)^3}\, 2\pi \iint \frac{1}{2|\boldsymbol{k}|} e^{i|\boldsymbol{k}||\boldsymbol{x}|\cos\theta}\left\{e^{i|\boldsymbol{k}|x_0} - e^{-i|\boldsymbol{k}|x_0}\right\} |\boldsymbol{k}|^2\, d|\boldsymbol{k}|\, \sin\theta\, d\theta$$

$$= -\frac{i}{(2\pi)^2} \int_0^\infty \frac{1}{2|\boldsymbol{k}|}\frac{1}{i|\boldsymbol{k}||\boldsymbol{x}|}\left\{e^{i|\boldsymbol{k}|x_0} - e^{-i|\boldsymbol{k}|x_0}\right\}|\boldsymbol{k}|^2$$

$$\times \left(e^{i|\boldsymbol{k}||\boldsymbol{x}|\alpha}\Big|_{\alpha=-1}^{\alpha=+1}\right) d|\boldsymbol{k}|$$

$$= -\frac{1}{4\pi^2}\frac{1}{2|\boldsymbol{x}|} \int_0^\infty \left\{e^{i|\boldsymbol{k}|(x_0+|\boldsymbol{x}|)} - e^{i|\boldsymbol{k}|(x_0-|\boldsymbol{x}|)} - e^{-i|\boldsymbol{k}|(x_0-|\boldsymbol{x}|)}\right.$$

$$\left. + e^{-i|\boldsymbol{k}|(x_0+||\boldsymbol{x})}\right\} d|\boldsymbol{k}|$$

$$= -\frac{1}{4\pi|\boldsymbol{x}|}\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left\{e^{i|\boldsymbol{k}|(x_0+|\boldsymbol{x}|)} - e^{i|\boldsymbol{k}|(x_0-|\boldsymbol{x}|)}\right\}d|\boldsymbol{k}|$$

$$= -\frac{1}{4\pi|\boldsymbol{x}|}\,\delta(x_0 - |\boldsymbol{x}|) \quad \text{for } x_0 > 0.$$

5.1.2 *Fourier Analysis of Operators*

Let us analyze the potential A_μ into Fourier components

$$A_\mu(x) = B\int d^3k\,|\boldsymbol{k}|^{-1/2}\left\{a_{k\mu}\exp(ik\cdot x) + \tilde{a}_{k\mu}\exp(-ik\cdot x)\right\} \qquad (211)$$

where the factor $|\boldsymbol{k}|^{-1/2}$ appears only as a matter of convenience; the actual Fourier coefficients are then $|\boldsymbol{k}|^{-1/2}\,a_{k\mu}$ and $|\boldsymbol{k}|^{-1/2}\,\tilde{a}_{k\mu}$. The integration is over all 4-vectors (k) with $k_0 = +|\boldsymbol{k}|$. B is a normalization factor to be determined later. The $a_{k\mu}$ and $\tilde{a}_{k\mu}$ are operators independent of x.

Since A_1, A_2, A_3 and A_0 are Hermitian,

$$\tilde{a}_{k\mu} = a_{k\mu}^* = \text{ Hermitian conjugate of } a_{k\mu}\,, \quad \mu = 1,2,3,0 \text{ and therefore}$$
$$\tilde{a}_{k4} = -a_{k4}^* = -\text{Hermitian conjugate of } a_{k4}$$
$$(212)$$

Computing the commutator $[A_\mu(z),\,A_\lambda(y)]$ from (211) and comparing the result with (203) in the momentum representation (210), we have first, since the result is a function of $(z - y)$ only[22]

$$[\,a_{k\mu},\,a_{k'\lambda}\,] = 0$$
$$[\,\tilde{a}_{k\mu},\,\tilde{a}_{k'\lambda}\,] = 0 \qquad (213)$$
$$[\,a_{k\mu},\,\tilde{a}_{k'\lambda}\,] = \delta^3(\boldsymbol{k} - \boldsymbol{k}')\,\delta_{\mu\lambda}$$

And the two results for the commutator agree then precisely if we take

$$B = \sqrt{\frac{\hbar c}{16\pi^3}} \qquad (214)$$

5.1.3 *Emission and Absorption Operators*

The operators $A_\mu(x)$ obey the Heisenberg equations of motion for operators (190)

$$i\hbar\,\frac{\partial A_\mu}{\partial t} = [\,A_\mu,\,H\,] \qquad (190a)$$

Therefore the operator $a_{k\mu} \exp(ik \cdot x)$ has matrix elements between an initial state of energy E_1 and a final state of energy E_2, only if

$$i\hbar(-ick_0) = E_1 - E_2 = \hbar c|\boldsymbol{k}| \tag{215}$$

because by (211) and (190a) we have

$$\psi_1 \, i\hbar(-ick_0)a_{k\mu} \, \psi_2 = \psi_1 \, \hbar c|\boldsymbol{k}| \, a_{k\mu} \, \psi_2$$
$$= \psi_1 \, [\, a_{\mu k}, \, H \,] \, \psi_2 = (E_1 - E_2)\psi_1 \, a_{\mu k} \, \psi_2 \tag{216}$$

Now $\hbar c|\boldsymbol{k}|$ is a constant energy, characteristic of the frequency $\omega = ck$ characteristic of the particular Fourier components of the field. The operator $a_{k\mu}$ can only operate so as to *reduce* the energy of a system by a lump of energy of this size. In the same way, $\tilde{a}_{k\mu}$ will only operate when

$$E_1 - E_2 = -\hbar c|\boldsymbol{k}|$$

to *increase* the energy by the same amount.

This is the fundamental property of the quantized field operators, that they change the energy of a system not continuously but in jumps. This shows that our formalism includes correctly the experimentally known quantum behaviour of radiation.

We call $a_{k\mu}$ the *absorption* operator for the field oscillator with propagation vector k and polarization direction μ. Likewise $\tilde{a}_{k\mu}$ the *emission* operator.

We have thus 4 directions of polarization for a photon of given momentum. There are not all observed in electromagnetic radiation. Free radiation can only consist of transverse waves, and has only 2 possible polarizations. This is because the physically allowable states Ψ are restricted by a supplementary condition

$$\sum_{\mu} \frac{\partial A_{\mu}^{(+)}}{\partial x_{\mu}} \Psi = 0 \tag{217}$$

where $A_{\mu}^{(+)}$ is the positive frequency part of A_{μ}, i.e. the part containing the absorption operators. In the classical theory we have

$$\sum_{\mu} \frac{\partial A_{\mu}}{\partial x_{\mu}} = 0$$

the condition imposed in order to simplify the Maxwell equations to the simple form $\Box^2 A_\mu = 0$. In the quantum theory it was usual to take

$$\sum_\mu \frac{\partial A_\mu}{\partial x_\mu} \Psi = 0$$

but this means that photons of a certain kind cannot be *emitted* in a physical state which is physically difficult to understand and brings mathematical inconsistencies into the theory. So we assume only (216) which only says these photons are *not present* and cannot be absorbed from a physical state, which makes good sense. Also in the classical limit $\sum_\mu \partial A_\mu / \partial x_\mu$ is a real quantity, and so $\sum_\mu \partial A_\mu / \partial x_\mu = 0$ follows correctly from (216) alone.

The method of using (216) as supplementary condition is due to Gupta and Bleuler[23];

References: S. N. Gupta, *Proc. Roy. Soc. A* **63** (1950) 681.
K. Bleuler, *Helv. Phys. Acta* **23** (1950) 567.

The older treatment is unnecessary and difficult, so we will not bother about it.

By (211), (216) is equivalent to assuming

$$\sum_\mu \left(k_\mu a_{k\mu} \right) \Psi = 0 \tag{216a}$$

for each momentum vector k of a photon.

As a result of this work of Gupta and Bleuler, the supplementary conditions do not come into the practical use of the theory at all. We use the theory and get correct results, forgetting about the supplementary conditions.

5.1.4   *Gauge-Invariance of the Theory*

The theory is gauge-invariant. That is to say, adding a gradient[24] $\Lambda_\mu = \partial \Lambda / \partial x_\mu$ to the potentials does not change the fields by anything which is physically observable. Therefore all states which differ only by such an addition to the potentials are physically identical.

If Ψ is any state then

$$\Psi' = \left(1 + \lambda \sum_\mu k_\mu \tilde{a}_{k\mu} \right) \Psi$$

is a state obtained from Ψ by emitting a pseudo-photon with potentials proportional to $\partial\Lambda/\partial x_\mu$. Hence Ψ' should be indistinguishable from Ψ.

Now if Ψ_2 is any state whatever satisfying the supplementary condition (216a) the matrix element

$$
\begin{aligned}
\left(\Psi'^*, \Psi_1\right) &= \left(\Psi^* \left(1 + \lambda \sum_\mu k_\mu^* \, \tilde{a}_{k\mu}^*\right), \Psi_2\right) \\
&= \left(\Psi^*, \left(1 + \lambda \sum_\mu k_\mu \, a_{k\mu}^*\right) \Psi_2\right) = (\Psi^*, \Psi_2)
\end{aligned}
\tag{216b}
$$

Hence the matrix elements of Ψ' and Ψ to any physical state Ψ_2 whatever are equal, and so the results of the theory are all independent of whether the state Ψ is represented by the vector Ψ or by Ψ'. This is enough to show that the theory is properly gauge-invariant, in spite of the fact that states are specified by the potentials which are not themselves gauge-invariant.

5.1.5   *The Vacuum State*

The vacuum state is by definition the state of lowest energy, so that all absorption operators operating on it give zero:

$$
a_{k\mu}\Psi_o = 0
\tag{217a}
$$

and therefore by (212)

$$
(a_{k\mu}\Psi_o)^* = (\Psi_o^* a_{k\mu}^*) = \pm(\Psi_o^* \tilde{a}_{k\mu}) = 0
\tag{217b}
$$

Given any operator $\mathcal{Q}$, we are interested in the "vacuum expectation value" of $\mathcal{Q}$ defined as

$$
\langle Q \rangle_o = (\Psi_o^*, \mathcal{Q}\Psi_o)
\tag{218}
$$

Then we have at once

$$
\langle a_{k\mu} a_{k'\lambda} \rangle_o = 0 \quad \text{by (217a)}
\tag{219a}
$$
$$
\langle \tilde{a}_{k\mu} \tilde{a}_{k'\lambda} \rangle_o = 0 \quad \text{by (217b)}
\tag{219b}
$$
$$
\langle \tilde{a}_{k\mu} a_{k'\lambda} \rangle_o = 0 \quad \text{by (217a,b)}
\tag{219c}
$$

And by the commutation laws (213) and (219c) we have

$$
\langle a_{k\mu} \tilde{a}_{k'\lambda} \rangle_o = \langle\, [\, a_{k\mu}, \, \tilde{a}_{k'\lambda} \,] \,\rangle_o = \delta^3(\boldsymbol{k} - \boldsymbol{k}')\, \delta_{\mu\lambda}
\tag{220}
$$

The vacuum expectation value $\langle A_\mu(z)\,A_\lambda(y)\rangle_o$ is thus just the part of the commutator $[\,A_\mu(z),\,A_\lambda(y)]$ which contains positive frequencies $\exp\{ik\cdot(z-y)\}$ with $k_0 > 0$, as one can see using (211), (219) and (220). Thus[25]

$$\langle A_\mu(z)\,A_\lambda(y)\rangle_o = i\hbar c\,\delta_{\mu\lambda}\,D^+(z-y) \tag{221}$$

$$D^+(x) = -\frac{i}{(2\pi)^3}\int_{k_0>0}\exp(ik\cdot x)\,\delta(k^2)\,\Theta(k)\,d^4k \tag{222}$$

We write

$$D(x) = D^+(x) + D^-(x) \tag{223}$$

$$D^+ = \frac{1}{2}\left(D - iD^{(1)}\right) \qquad D^- = \frac{1}{2}\left(D + iD^{(1)}\right) \tag{224}$$

The even function $D^{(1)}$ is then defined by

$$\langle A_\mu(z)\,A_\lambda(y) + A_\lambda(y)\,A_\mu(z)\rangle_o = \hbar c\,\delta_{\mu\lambda}\,D^{(1)}(z-y) \tag{225}$$

$$D^{(1)}(x) = \frac{1}{(2\pi)^3}\int \exp(ik\cdot x)\,\delta(k^2)\,d^4k \tag{226}$$

It is then not hard to prove (see the Appendix below) that

$$D^{(1)}(x) = \frac{1}{2\pi^2 x^2} \tag{227}$$

The functions D and $D^{(1)}$ are the two independent solutions of $\square^2 D = 0$, one odd and the other even. Then we define the function

$$\overline{D}(x) = -\frac{1}{2}\epsilon(x)\,D(x) = \frac{1}{2}\left(D_R(x) + D_A(x)\right) = \frac{1}{4\pi}\,\delta(x^2)$$
$$= \frac{1}{(2\pi)^4}\int \exp(ik\cdot x)\,\frac{1}{k^2}\,d^4k \tag{228}$$

the last being a real principal value integral: This is the even solution of the point-source equation

$$\square^2\overline{D}(x) = -\delta^4(x) \tag{229}$$

Appendix

$$D^{(1)}(x) = \frac{1}{(2\pi)^3} \int e^{ik\cdot x}\,\delta(k^2)\,d^4k$$

$$= -\frac{1}{(2\pi)^2} \int_{-1}^{+1} d\mu \int_{-\infty}^{\infty} dk_0 \int_0^{\infty} d|\boldsymbol{k}|e^{i|\boldsymbol{k}||\boldsymbol{x}|\mu}e^{-ik_0x_0}$$

$$\times \frac{\{\delta(k_0 - |\boldsymbol{k}|) + \delta(k_0 + |\boldsymbol{k}|)\}|\boldsymbol{k}|^2}{2|\boldsymbol{k}|}$$

$$= \frac{1}{(2\pi)^2} \int_0^{\infty} d|\boldsymbol{k}|\frac{1}{i|\boldsymbol{k}||\boldsymbol{x}|}\left\{e^{i|\boldsymbol{k}||\boldsymbol{x}|} - e^{-i|\boldsymbol{k}||\boldsymbol{x}|}\right\}$$

$$\times \frac{1}{2|\boldsymbol{k}|}\left\{e^{-i|\boldsymbol{k}|x_0} + e^{i|\boldsymbol{k}|x_0}\right\}|\boldsymbol{k}|^2$$

$$= \frac{1}{2\pi^2}\frac{1}{|\boldsymbol{x}|} \int_0^{\infty} \sin(|\boldsymbol{k}||\boldsymbol{x}|)\cos(|\boldsymbol{k}|x_0)\,d|\boldsymbol{k}|$$

$$= \frac{1}{2\pi^2}\frac{1}{2|\boldsymbol{x}|} \int_0^{\infty} d|\boldsymbol{k}|\left\{\sin((|\boldsymbol{x}| + x_0)|\boldsymbol{k}|) + \sin((|\boldsymbol{x}| - x_0)|\boldsymbol{k}|)\right\}$$

Taking the integral in the Abelian sense

$$\lim_{\epsilon\to 0} \int_0^{\infty} e^{-\epsilon x}\sin ax\,dx = \lim_{\epsilon\to 0} \frac{a}{\epsilon^2 + a^2} = \frac{1}{a}$$

hence in our case[26]

$$D^{(1)}(x) = \frac{1}{2\pi^2}\frac{1}{2|\boldsymbol{x}|}\left\{\frac{1}{|\boldsymbol{x}| + x_0} + \frac{1}{|\boldsymbol{x}| - x_0}\right\} = \frac{1}{2\pi^2 x^2}.$$

5.1.6 The Gupta-Bleuler Method

There is one difficulty in the preceding theory. We assume according to (220)

$$\langle a_{k\mu}a_{k'\lambda}^*\rangle_o = \pm\delta^3(\boldsymbol{k} - \boldsymbol{k}')\,\delta_{\mu\lambda} \tag{220a}$$

Here the plus sign holds for $\mu = 1, 2, 3$; the minus sign for $\mu = 4$. Now if the operators $a_{k\mu}$, $a_{k\mu}^*$ are represented in the usual way by matrices, as it is done in the elementary theory of the harmonic oscillator (see Wentzel p. 33, Eq. (6.16)), the vacuum expectation values of a product $(a_{k\mu}\,a_{k\mu}^*)$ will

always be *positive*, i.e. the plus sign should hold in (220a) for $\mu = 4$ also. In fact $(a_{k\mu} \, a^*_{k\mu})$ will have a positive expectation value in any state whatever, if the photon oscillators are treated as ordinary elementary oscillators.

Therefore we must distinguish between the scalar product $(\Psi_1^*, \, \Psi_2)$ as we have defined it by our covariant theory, and the scalar product $(\Psi_1^*, \, \Psi_2)_E$ which one would calculate using the explicit matrix representation of the operators. The product $(\Psi_1^*, \, \Psi_2)_E$ has no physical significance, because the matrix representations of the a_{k4} refer to states with photons polarized purely in the time dimension, which cannot occur physically. However, it is convenient also to be able to use the matrix representations in practice.

To use the matrix representations, we have only to define an operator η by the condition

$$\eta\Phi = (-1)\Phi \tag{220b}$$

where Φ is any state in which there is a definite number N of photons polarized in the 4-direction. Then the physical scalar product is given in terms of the explicit matrix representations by

$$(\Psi_1^*, \, \Psi_2) = (\Psi_1^*, \, \eta\Psi_2)_E \tag{220c}$$

The definition (220c), introduced by Gupta, makes the matrix representations consistent with all the requirements of the covariant theory, in particular also it gives (220) correctly. The physical scalar product is thus an *indefinite metric* regarded from the point of view of the matrix representations. However, we have seen in (216b) that for any *physical* states the scalar product $(\Psi_1^*, \, \Psi_2)$ is equal to $(\Psi_{1T}^*, \, \Psi_{1T})$ where Ψ_{1T} is a state involving transverse photons only, and hence is positive. Thus for physical states the metric is definite and this is all that we require of it.

5.1.7 *Example: Spontaneous Emission of Radiation*

This is a purely quantum-mechanical effect. A classical treatment, considering the reaction of the atom to a classical applied Maxwell field, gives a correct account of the absorption of radiation and of stimulated emission, but fails to give the spontaneous emission.

Let an atom have 2 states, the ground state 1 and an excited state 2 with energy $\hbar c q$. For the transition $2 \to 1$ let the charge-current density of the atom have the unintegrated matrix elements

$$j_{\mu A}(x) = j_{\mu A}(r, t) \quad \text{at the point } x = (r, t)$$

The interaction with the Maxwell field has matrix element[27]

$$I = -\frac{1}{c} \int \sum_\mu j_{\mu A}(r,t) \langle A_\mu(r,t)\rangle_{\text{emit}} \, d^3 \boldsymbol{r} \tag{230}$$

for making a transition with emission of a photon. The total emission probability per unit time is obtained using time dependent perturbation theory:[28]

$$w = \frac{1}{T} \sum |a_1(T)|^2$$

$$= \frac{1}{T} \sum_{\substack{\text{photon} \\ \text{states}}} \left| \frac{1}{i\hbar} \int \left\{ -\frac{1}{c} \int \sum_\mu j_{\mu A}(r,t) \langle A_\mu(r,t)\rangle_{\text{emit}} \, d^3 \boldsymbol{r} \right\} dt \right|^2$$

$$= \frac{1}{Tc^4\hbar^2} \iint \sum_{\lambda,\mu} j^*_{\lambda A}(x') \, j_{\mu A}(x) \langle A^*_\lambda(x') A_\mu(x)\rangle_o \, d^4x \, d^4x' \tag{231}$$

the integral being over all space for a long time T, the sum over the physical photon states only. It is *not* correct to take in (231) the photon states to be the 4 states with polarization in the directions $\mu = 1,2,3,4$, because these are not physical states.

Using a sum-rule to sum over the states

$$w = \frac{1}{Tc^4\hbar^2} \iint \sum_{\lambda,\mu} j^*_{\lambda A}(x') \, j_{\mu A}(x) \langle A^*_\lambda(x') A_\mu(x)\rangle_o \, d^4x \, d^4x'$$

Write $\tilde{j}_{\lambda A}(x')$ for the matrix element[29] of $j_{\lambda A}(x')$ in the reverse transition $1 \to 2$. Then

$$j^*_{\lambda A}(x') = \eta_\lambda \, \tilde{j}_{\lambda A}(x'), \qquad A^*_\lambda(x') = \eta_\lambda A_\lambda(x')$$

with $\eta_\lambda = +1$, $\lambda = 1,2,3$; $\eta_4 = -1$.

Hence

$$w = \frac{1}{Tc^4\hbar^2} \iint \sum_{\lambda,\,\mu=1}^{4} j^*_{\lambda A}(x') \, j_{\mu A}(x) \langle A^*_\lambda(x') A_\mu(x)\rangle_o \, d^4x \, d^4x'$$

$$= \frac{\hbar c}{(2\pi)^3} \int d^4k \, \delta(k^2) \Theta(k_0) \sum_{\mu=1}^{4} j_{\mu A}(k) \tilde{j}_{\mu A}(k) \frac{1}{Tc^4\hbar^2}$$

$$\times \int_0^{cT} \int_0^{cT} dx_0 \, dx_0' \exp\left\{i(x - x_0)(q - k_o)\right\}$$

$$= \frac{1}{(2\pi)^2 \hbar c^2} \sum_{\mu=1}^{4} \int d^3\boldsymbol{k} \, \delta(|\boldsymbol{k}|^2 - q^2) j_{\mu A}(\boldsymbol{k}) \tilde{j}_{\mu A}(\boldsymbol{k}) \tag{232}$$

since

$$\frac{1}{cT} \int_0^{cT} \int_0^{cT} dx_0 \, dx_0' \exp\left\{i(x - x_0)(q - k_o)\right\}$$

$$= cT \frac{\sin^2 \frac{cT}{2}(q - k_0)}{\left(\frac{cT}{2}(q - k_0)\right)^2} \to \pi\delta(q - k_0) \quad \text{if } cT \to \infty$$

where[30]

$$j_{\mu A}(\boldsymbol{k}) = \int j_{\mu A}(r) \, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \, d^3\boldsymbol{r} \qquad \tilde{j}_{\mu A}(\boldsymbol{k}) = \int \tilde{j}_{\mu A}(r) \, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \, d^3\boldsymbol{r}$$

By the charge conservation law

$$\sum_\mu k_\mu j_{\mu A}(k) = 0 \qquad\qquad \sum_\mu k_\mu \tilde{j}_{\mu A}(k) = 0$$

and so

$$\sum_\mu j_{\mu A}(k) \tilde{j}_{\mu A}(k) = \frac{1}{q^2} q j_{4A} \, q \tilde{j}_{4A} + \sum_i^3 |j_{iA}(k)|^2$$

$$= -\frac{1}{q^2} \sum_{i,\ell=1}^{3} \left\{ k_i j_{iA}(k) \, k_\ell \tilde{j}_{\ell A}(k) \right\} + \sum_i^3 |j_{iA}(k)|^2$$

$$= -\frac{1}{q^2} |\boldsymbol{k}|^2 |\boldsymbol{j}_A|^2 \cos^2\theta + |\boldsymbol{j}_A|^2$$

$$= |\boldsymbol{j}_A|^2 (1 - \cos^2\theta) = |j_{1A}|^2 + |j_{2A}|^2 \tag{232a}$$

where 1 and 2 are the two directions of transverse polarization. This shows how the third and fourth polarization directions do not appear in real emission problems. The same result would be obtained if we used the indefinite metric explicitly, i.e. take the sum in (231) over the 4 polarization states

$\mu = 1, 2, 3, 4$, with the $\mu = 4$ given a minus sign arising from the η in (220c). But it is always simpler to work directly with the covariant formalism, than to bother with the non-physical photon states and then have to use η to get the right answers.

Finally, the emission probability in direction of polarization 1 and in direction of propagation given by the solid angle $d\Omega$ is, using (232), (232a), and $\delta(|\boldsymbol{k}|^2 - q^2) = \frac{1}{2q}\delta(|\boldsymbol{k}| - q)$ for $q > 0$

$$w = \frac{q\, d\Omega}{8\pi^2 \hbar c^2}\, |j_{1A}(x)|^2 \tag{233}$$

For dipole radiation by a one-electron atom with coordinates (x, y, z)

$$j_1 = e\dot{x} = iecqx \qquad \boldsymbol{k} \cdot \boldsymbol{r} \ll 1$$

and

$$w = \frac{e^2 q^3\, d\Omega}{8\pi^2 \hbar}\, |\langle x \rangle_{12}|^2 \tag{234}$$

This checks with Bethe's *Handbuch* article.[31]

The example shows how covariant methods will work, even for problems of this elementary sort for which they are not particularly suited. The covariant method avoids the necessity of having to think about the normalization of the photon states, the factors of 2 and π etc. being given automatically when one uses (221).

5.1.8   *The Hamiltonian Operator*

From the equation

$$i\hbar \frac{\partial A_\mu}{\partial t} = [\, A_\mu, H\,]$$

we find

$$[\, a_{k\mu}, H\,] = \hbar c |\boldsymbol{k}| a_{k\mu}$$
$$[\, \tilde{a}_{k\mu}, H\,] = -\hbar c |\boldsymbol{k}| \tilde{a}_{k\mu}$$

Using the commutation rules (213) we can find an operator H which satisfies all these conditions simultaneously. Namely

$$H = \int d^3\boldsymbol{k}\, \hbar c |\boldsymbol{k}| \sum_1^4 \tilde{a}_{k\lambda} a_{k\lambda} \tag{234a}$$

This operator is in fact unique apart from an arbitrary additive constant. To fix the constant we require[32] $\langle H \rangle_o = 0$ which leads to the result (234a) precisely, as one can see at once from (219). Hence (234a) is the Hamiltonian of this theory, which is very simple in this momentum representation.

To derive H from the Lagrangian is also possible but much more tedious. From (234a) we see that

$$N_{k\lambda} = \tilde{a}_{k\lambda} a_{k\lambda} \qquad \text{(not summed)}$$

is an operator just representing the number of quanta in the frequency k and polarization λ. It follows at once from the commutation rules (213), from the singular δ-function factor which comes from the continuous spectrum, $N_{k\lambda}$ being in fact the number of quanta per unit frequency range, that $\int N_{k\lambda} d^3 \mathbf{k}$ integrated over *any* region of momentum-space has the integer eigenvalues $0, 1, 2, \ldots$. This is so, because the state with n_i particles with momentum k_i is $\Psi = \prod_{i=1}^{\ell} (\tilde{a}_{k_i\lambda})^{n_i} \Psi_o$. Then taking $\int_\Omega N_{k\lambda} d^3 \mathbf{k}$ over Ω including the momenta $k_1, k_2, \ldots, k_j$ we get

$$\int_\Omega N_{k\lambda} d^3 \mathbf{k} \, \Psi = \int_\Omega \tilde{a}_{k\lambda} a_{k\lambda} \prod_{i=1}^{\ell} (\tilde{a}_{k_i\lambda})^{n_i} \Psi_o \, d^3 \mathbf{k}$$

$$= \int_\Omega \tilde{a}_{k\lambda} \left\{ \sum_{i=1}^{\ell} n_i \prod_{i=1}^{\ell} \tilde{a}_{k_i\lambda}^{(n_i-1)} [\, a_{k\lambda}, \, \tilde{a}_{k_i\lambda} \,] \right.$$

$$\left. + \prod_{i=1}^{\ell} (a_{k_i\lambda})^{n_i} a_{k\lambda} \right\} \Psi_o \, d^3 \mathbf{k} \qquad \text{by (213) and (217a)}$$

$$= \int_\Omega \tilde{a}_{k\lambda} \sum_{i=1}^{\ell} n_i \prod_{i=1}^{\ell} \tilde{a}_{k_i\,\lambda}^{(n_i-1)} \delta^3(\mathbf{k} - \mathbf{k}_i) \Psi_o \, d^3 \mathbf{k}$$

$$= \sum_{i=1}^{j} n_i \prod_{i=1}^{\ell} (\tilde{a}_{k_i\lambda})^{n_i} \Psi_o = \sum_{i=1}^{j} n_i \Psi$$

5.1.9 *Fluctuations of the Fields*

Since the electromagnetic fields $\mathbf{E}$ and $\mathbf{H}$ are quantum-mechanical variables, they do not have well-determined values in any state in which energy and momentum are well-defined, for example the vacuum state. A state of the fields can be specified *either* by fixing the values of $\mathbf{E}$ and $\mathbf{H}$ *or* by specifying

the numbers of quanta present with various momenta and energies. The two descriptions are complementary and are both possible only in the classical limit of large numbers of quanta and very strong fields.

An educational discussion of these questions, considering in detail the example of a cavity resonator with one mode of oscillation, has been given by L. P. Smith, *Phys. Rev.* **69** (1946) 195. This is worth reading. Essential is the fact that you cannot fix the time-dependence of the field (phase) with a fixed number of quanta (energy).

We consider a more general problem. What is the mean-square fluctuation in the vacuum state of a field-quantity? We define

$$
\begin{aligned}
E_1(VT) &= \frac{1}{VT} \int_{VT} E_1(x)\, d\tau\, dt \\[2mm]
H_1(VT) &= \frac{1}{VT} \int_{VT} H_1(x)\, d\tau\, dt
\end{aligned}
\tag{235}
$$

averaged over some finite space-volume V and also over a time T. Let $V(\mathbf{k}) = \int_V e^{-i\mathbf{k}\cdot\mathbf{r}}\, d\tau$. Then since $\mathbf{H} = \nabla \times \mathbf{A}$, we have

$$
\Big\langle \{H_1(VT)\}^2 \Big\rangle_o
$$

$$
= \frac{1}{V^2 T^2} \iint d\tau\, d\tau'\, dt\, dt' \left\langle \left(\frac{\partial}{\partial x_2} A_3 - \frac{\partial}{\partial x_3} A_2 \right) \left(\frac{\partial}{\partial x_2'} A_3' - \frac{\partial}{\partial x_3'} A_2' \right) \right\rangle_o
$$

$$
= \frac{1}{V^2 T^2} \frac{\hbar c}{2} \iint d\tau\, d\tau'\, dt\, dt' \left(\frac{\partial}{\partial x_2} \frac{\partial}{\partial x_2'} + \frac{\partial}{\partial x_3} \frac{\partial}{\partial x_3'} \right)
$$

$$
\times\, D^{(1)}(x - x') \quad \text{using (225)}
$$

$$
= \frac{1}{V^2 T^2} \frac{\hbar c}{16\pi^3} \iint d\tau\, d\tau'\, dt\, dt' \left(\frac{\partial}{\partial x_2} \frac{\partial}{\partial x_2'} + \frac{\partial}{\partial x_3} \frac{\partial}{\partial x_3'} \right)
$$

$$
\times \int d^4k\, e^{ik\cdot(x-x')} \delta(k^2) \quad \text{using (226)}
$$

$$
= \frac{\hbar c}{16\pi^3 V^2 T^2} \int_0^T \int_0^T \int \frac{d^3\mathbf{k}}{|\mathbf{k}|} (k_2^2 + k_3^2)\, |V(k)|^2\, e^{i|\mathbf{k}|x_0}\, e^{-i|\mathbf{k}|x_0'}\, dt\, dt'
$$

$$
= \frac{\hbar c}{16\pi^3 V^2 T^2} \int \frac{d^3\mathbf{k}}{|\mathbf{k}|} (k_2^2 + k_3^2)\, |V(k)|^2\, \frac{4\sin^2(\tfrac{1}{2}c|\mathbf{k}|T)}{c^2|\mathbf{k}|^2}
$$

$$\left\langle \{E_1(VT)\}^2 \right\rangle_o$$

$$= \frac{1}{V^2T^2}\frac{\hbar c}{2}\iint d\tau\, d\tau'\, dt\, dt' \left(-\frac{\partial}{\partial x_4}\frac{\partial}{\partial x_4'} - \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_1'}\right) D^{(1)}(x-x')$$

$$= \frac{\hbar c}{16\pi^3 V^2 T^2}\int \frac{d^3\mathbf{k}}{|\mathbf{k}|}(|\mathbf{k}|^2 - k_1^2)\,|V(k)|^2\,\frac{4\sin^2(\frac{1}{2}c|\mathbf{k}T)}{c^2|\mathbf{k}|^2}$$

$$= \left\langle \{H_1(VT)\}^2 \right\rangle_o \tag{236}$$

Taking for V any finite volume and T a finite time, this mean-square fluctuation is finite. Example: a sphere of radius R gives

$$V(K) = \frac{4\pi}{|\mathbf{k}|^3}\left(\sin R|\mathbf{k}| - R|\mathbf{k}|\cos R|\mathbf{k}|\right) \tag{237}$$

But if either R or T tends to zero, the fluctuations tend to ∞ and in the limit actually diverge.[33] That is to say, only measurements of field-quantities averaged both in space and in time can have any physical reality.

5.1.10 Fluctuation of Position of an Electron in a Quantized Electromagnetic Field. The Lamb Shift

Consider an electron represented by an extended spherical charge of radius R, lying in a stationary state in the potential $\phi(r)$ of a hydrogen atom. It has a certain wave-function $\psi(r)$. We consider everything non-relativistically except for the quantized radiation field with which the electron interacts. The effect of this fluctuating field is to produce a rapid fluctuation in position of the electron. In fact we have for rapid fluctuations

$$m\ddot{\mathbf{r}} = -e\mathbf{E}$$

Thus a fluctuating component of $\mathbf{E}$ with frequency $c|\mathbf{K}|$ produces the same fluctuation in $\mathbf{r}$ with amplitude multiplied by a factor $\frac{e}{m}\frac{1}{c^2|\mathbf{K}|^2}$. The slow fluctuations of $\mathbf{E}$ cannot be followed by the electron if the frequency is less than the atomic frequency cK_H. Hence we find from (236), making $T \to 0$

$$\langle r_1^2\rangle_o = \frac{e^2}{m^2}\frac{\hbar c}{16\pi^3 V^2}\int_{K_H}^{\infty}\frac{d^3\mathbf{K}}{|\mathbf{K}|}(K_2^2 + K_3^2)\,|V(K)|^2\frac{1}{c^4|\mathbf{K}|^4}$$

because $\lim\limits_{x\to 0}\frac{\sin^2 x}{x^2} = 1$. The integral now converges at ∞, because of the finite size of the electron. Since R is very small we may approximate (237)

by

$$V(K) = \begin{cases} \dfrac{4}{3}\pi R^3 = V & \text{for } |\boldsymbol{K}|R < 1 \\ 0 & \text{for } |\boldsymbol{K}|R > 1 \end{cases}$$

Then, since $(K_2^2 + K_3^2) = |\boldsymbol{K}|^2(1 - \cos^2\theta)$ and $\int_0^\pi \sin^2\theta \sin\theta\, d\theta = \frac{4}{3}$, we have[34]

$$\langle r_1^2 \rangle_o = \frac{e^2\hbar}{6m^2c^3\pi^2} \int_{K_H}^{1/R} \frac{d|\boldsymbol{K}|}{|\boldsymbol{K}|} = \frac{e^2\hbar}{6m^2c^3\pi^2} \log\left(\frac{1}{RK_H}\right) \tag{238}$$

This fluctuation in position produces a change in the effective potential acting on the electron. Thus

$$\langle V(r + \delta r) \rangle = V(r) + \langle \delta\boldsymbol{r} \cdot \nabla V(r) \rangle_o + \frac{1}{2} \langle (\delta\boldsymbol{r})^2 \rangle_o \frac{\partial^2 V}{\partial r^2} + \dots$$

$$= V(r) + \frac{1}{2} \langle r_1^2 \rangle_o \nabla^2 V$$

because $\langle \delta\boldsymbol{r} \cdot \nabla V(r) \rangle_o = 0$, being odd. Now in a hydrogen atom, $\nabla^2 V = e^2\delta^3(\boldsymbol{r})$ (Heaviside units!) Hence the change in the energy of the electron due to the fluctuations is[35] ($a_o = $ Bohr radius)

$$\Delta E = \int \psi^* \, \delta V \, \psi \, d\tau = \frac{1}{2} \langle r_1^2 \rangle_o e^2 |\psi(0)|^2$$

$$= \begin{cases} \dfrac{e^4\hbar}{12\pi^2 m^2 c^3} \log\left(\dfrac{1}{RK_H}\right) \dfrac{1}{\pi n^3 a_o^3} & \text{for s-states} \\ 0 & \text{for all others} \end{cases} \tag{239}$$

because for the hydrogen atom[36] ($\rho = r\sqrt{-8m_r E_n}/\hbar$)

$$\psi_{n\ell m}(r, \theta, \varphi) = -\frac{1}{\sqrt{2\pi}} e^{im\varphi} \left\{ \frac{(2\ell + 1)(\ell - |m|)!}{2(\ell + |m|)!} \right\}^{1/2}$$

$$\times P_\ell^{|m|}(\cos\theta) \left[\left(\frac{2}{na_o}\right)^3 \frac{(n - \ell - 1)!}{2n[(n + \ell)!]^3} \right]^{1/2} e^{-\rho/2} \rho^\ell L_{n+\ell}^{2\ell+1}(\rho)$$

and

$$\psi_{n00}(0, \theta, \varphi) = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{2}}{2} \left\{ -\left[\left(\frac{2}{na_o}\right)^3 \frac{(n - 1)!}{2n(n!)^3} \right]^{1/2} \right\} \left(-\frac{(n!)^2}{(n - 1)!} \right)$$

$$= \frac{1}{\pi^{1/2} a_o^{3/2} n^{3/2}}$$

There will also be a (much bigger) addition to kinetic energy, arising from the fluctuations. We ignore this on the grounds that it will be the same for all atomic states and so will not give any relativistic displacement. Of course this is not a good argument.

Hence we find the first approximation to the Lamb shift; the $2s$ state is shifted relative to the $2p$ states by

$$\Delta E = + \frac{e^4 \hbar}{96\pi^3 m^2 c^3 a_o^3} \log \frac{1}{RK_H}$$

Now[37]

$$a_o = \frac{4\pi \hbar^2}{me^2} = \frac{1}{\alpha} \frac{\hbar}{mc} \quad \text{(Bohr radius)}$$

$$\text{Ry} = \frac{e^4 m}{32\pi^2 \hbar^4} \quad \text{(Rydberg energy unit)}$$

$$K_H = \frac{\text{Ry}}{4\hbar c}$$

We take $R = (\hbar/mc)$, the electron Compton wave-length since it is at this frequency that the non-relativistic treatment becomes completely wrong. Then

$$RK_H = \frac{\text{Ry}}{4mc^2} = \frac{1}{8}\alpha^2$$

$$\Delta E = + \frac{\alpha^3}{3\pi} \log(8 \times 137^2) \, \text{Ry}$$

(240)

Actually $\frac{\alpha^3}{3\pi}$ Ry $= 136$ Mc in frequency units. This gives an effect of the right sign and order of magnitude. This method is due to Welton [14].

The size of the log is wrong because the low-frequency cut-off was badly done. We find $\Delta E \sim 1600$ Mc instead of the correct value of 1060 Mc. But physically the origin of the shift is correctly described in this way.

5.2 Theory of Line Shift and Line Width

To make a better treatment of the effect of the radiation interaction on energy-levels, we must try to solve more exactly the equation of motion for the system atom plus radiation field. The effect of the field is shown not only in a shift of energy levels, but also in a finite *width* of the levels due to real radiation. Roughly, if the state has a life-time T for decay by radiation,

the width Γ of the level, or the mean variation in energy of the emitted photons, is given by the uncertainty principle $\Gamma \approx \hbar/T$. The line-shift and line-width are effects of the same kind and cannot be correctly treated except in combination.

So we make a theory now which treats the atom non-relativistically, but takes account properly of the radiation interaction. This means, we repeat the calculation of spontaneous emission by an atom, but now including the reaction of the radiation on the atom instead of taking the atom as given by a fixed charge-current oscillator.

For this sort of calculation it is always convenient to work in a special representation called the *Interaction Representation*.

5.2.1   *The Interaction Representation*

In the Schrödinger representation the wave-function Ψ satisfies the equation of motion

$$i\hbar\frac{\partial}{\partial t}\Psi = H\Psi \tag{241}$$

where H is the Hamiltonian. In the case of an atom interacting with the radiation field, we have

$$H = H_A + H_M + H_I^S \tag{242}$$

where H_A is the Hamiltonian operator for the atom, and H_M that for the Maxwell field without interaction. H_M is given by (234a), and in quantum electrodynamics according to (170)

$$H_I^S = -\frac{1}{c}\int\sum_\mu j_\mu^S(r)A_\mu^S(r)d^3\boldsymbol{r} \tag{243}$$

because $j_\mu^S(r) = ie\,\overline{\psi}\gamma_\mu\psi$, and in this case $\mathscr{H} = \sum_\mu \pi_\mu A_\mu - \mathscr{L} = -\mathscr{L}$ because $\pi_\mu = \partial\mathscr{L}/\partial\dot{A}_\mu$, and $\mathscr{L}_I$ does not contain $\dot{A}_\mu$. All the operators in (242), (243) are time-independent[38], Schrödinger representation operators, and so are the given the label S.

Now we choose a new wave function $\Phi(t)$ which is given in terms of Ψ by

$$\Psi(t) = \exp\left\{-\frac{i}{\hbar}\left(H_A + H_M\right)t\right\}\Phi(t) \tag{244}$$

This $\Phi(t)$ will be a constant for any state representing the atom and the Maxwell field without interaction. Thus the time-variation of $\Phi(t)$ in an

actual state describes *just* the effect of the interaction in perturbing the atomic states. From (241) and (244), the time-variation of Φ is given by the Schrödinger equation

$$i\hbar \frac{\partial \Phi}{\partial t} = H_I(t)\Phi \tag{245}$$

with

$$H_I(t) = \exp\left\{\frac{i}{\hbar}(H_A + H_M)t\right\} H_I^S \exp\left\{-\frac{i}{\hbar}(H_A + H_M)t\right\} \tag{246}$$

Thus

$$H_I = -\frac{1}{c}\int \sum_\mu j_\mu(r,t)A_\mu(r,t)d^3\boldsymbol{r} \tag{247}$$

with

$$j_\mu(r,t) = \exp\left\{\frac{i}{\hbar}H_A t\right\} j_\mu^S(\boldsymbol{r}) \exp\left\{-\frac{i}{\hbar}H_A t\right\} \tag{248}$$

$$A_\mu(r,t) = \exp\left\{\frac{i}{\hbar}H_M t\right\} A_\mu^S(r) \exp\left\{-\frac{i}{\hbar}H_M t\right\} \tag{249}$$

These operators $j_\mu(r,t)$ and $A_\mu(r,t)$ have precisely the time-dependence of the field-operators in the Heisenberg representation, for the two systems, atom and radiation field, taken separately without interaction. Thus in the Interaction Representation (IR) the time-dependence of the Schrödinger wave-function is split into two parts, the operators taking the time-dependence of the non-interaction systems, the wave-function taking a time-dependence showing the effects of the interaction only. The operators $A_\mu(r,t)$ satisfy the wave-equation $\Box^2 A_\mu = 0$ and the covariant commutation laws (203), because we see from (249) that

$$\frac{\partial A_\mu(r,t)}{\partial t} = i\hbar\,[\,H_M,\,A_\mu(r,t)\,]$$

i.e. the time variation of $A_\mu(r,t)$ is the same as that of $A_\mu(x)$ in the Heisenberg representation without interaction (see (190a)), which in turn leads to the field equations (197). Matrix elements of IR operators given by (246), (248) or (249) between IR wave-functions given by (244) are of course the same as the matrix elements which would be obtained in any other representation.

5.2.2   *The Application of the Interaction Representation to the Theory of Line-Shift and Line-Width*

Consider the solution of Eq. (245) in which the atom is given initially in a stationary unperturbed state O with energy E_o, the Maxwell field being in the vacuum state, with no photons present. Let Φ_o be the interaction representation wave-function representing the atom in state O and the Maxwell field in the vacuum state, without interaction. Φ_o is independent of time.

The initial condition $\Phi(t) = \Phi_o$ at time $t = t_o$ is a physically unreal one. It would mean putting the atom into existence at time 0 without any radiation field being excited at this time. This we cannot do physically. In fact the initial condition for an atom in an excited state will depend on how it got into the excited state. This cannot be formulated in a simple way: one needs a complicated model in order to describe the initial excitation of the atom.

We are interested in calculating the variation with time of $(\Phi_o^*\Phi(t))$, the probability amplitude for finding the atom still in the unperturbed state Φ_o at time t. By (245) we have

$$\frac{d}{dt}\left(\Phi_o^*\Phi(t)\right) = -\frac{i}{\hbar}\left(\Phi_o^* H_I(t)\Phi(t)\right) \tag{250}$$

Suppose we take the physically unreal initial condition

$$\Phi(t) = \Phi_o \qquad \text{at } t = t_o$$

Then (250) will give

$$\frac{d}{dt}\left(\Phi_o^*\Phi(t)\right)_{t=t_o} = 0 \tag{251}$$

from the H_I given by (247) has zero expectation value in the Maxwell field vacuum, since $A_\mu(r,t)$ also has a zero expectation value in vacuum as one can see from (211) and (217). Thus $(\Phi_o^*\Phi(t))$ is momentarily stationary at $t = t_o$. This is however not interesting since the conditions at $t = t_o$ are entirely unphysical.

The physically meaningful quantity is the value of (250) at a time t a long time after t_o. Then the atom will have "settled down" to a quasi-stationary state of radiative decay, and we may expect that the value we find for (250) is independent of the particular initial condition chosen and will be correct for an atom which has been excited by any reasonable method to the state Φ_o.

We make the calculation so as to include effects of the radiation H_I up to second order. This means we include effects of emitting and absorbing one photon only. In fact we know physically that effects from two or more photons are very small, so the approximation is a good one.

Suppose that $(t - t_o)$ is long compared with all the atomic frequencies. Then a solution of (245), valid to first order in H_I, is

$$\Phi_1(t) = \left[1 - \frac{i}{\hbar} \int_{-\infty}^{t} H_I(t')\, dt'\right] a(t)\, \Phi_o$$

$$+ \text{ terms involving other atomic states } \Phi_n \text{ with two or more}$$

$$\text{photons present.} \tag{252}$$

Here $a(t) = (\Phi_o^* \Phi(t))$ is a slowly varying amplitude, constant to first order in H_I, representing the slow decay of the atom. Note that our treatment is not only a perturbation theory correct to second order in H_I, but it also must account *exactly* for the big effects produced by the radiative decay over long periods of time. Therefore we do not put $a(t) = 1$ in (252) although this would be correct to first order in H_I.

If we put $a(t) = 1$ in (252) we should have just the solution of the radiation emission problem, neglecting all effects of radiation reaction on the atom, which we obtained before from Eg. (230).

The value of $\frac{d}{dt} a(t) = \frac{d}{dt}(\Phi_o^* \Phi(t))$ is obtained correct to second order in H_I, and including the radiative reaction effects, by substituting (252) into (250).

Hence we have

$$\frac{1}{a(t)} \frac{d}{dt} a(t) = -\frac{1}{\hbar} \int_{-\infty}^{t} dt' \left\{ \Phi_o^* H_I(t) H_I(t') \Phi_o \right\} \tag{253}$$

Using (247), (221) this gives[39]

$$\frac{1}{a} \frac{da}{dt} = -\frac{i}{\hbar c} \int_{-\infty}^{t} dt' \iint d^3 r\, d^3 r'\, D^+(r - r', t - t') \sum_\mu \left\langle j_\mu(r,t) j_\mu(r',t') \right\rangle_{oo}$$

$$= -\frac{1}{(2\pi)^3 \hbar c} \int \frac{d^3 \boldsymbol{k}}{2|\boldsymbol{k}|} \int_{-\infty}^{t} dt' \iint d^3 \boldsymbol{r}\, d^3 \boldsymbol{r'}$$

$$\times \exp\left\{ i\boldsymbol{k} \cdot (\boldsymbol{r} - \boldsymbol{r'}) - ic|\boldsymbol{k}|(t - t') \right\} \sum_\mu \left\langle j_\mu(r,t) j_\mu(r',t') \right\rangle_{oo} \tag{254}$$

Let the atomic states be labelled by n, the state n having energy E_n. Let

$$j_\mu^k(n, m) \tag{255}$$

be the matrix element of the operator

$$\int j_\mu^S(r) e^{-i\boldsymbol{k} \cdot \boldsymbol{r}} \, d^3\boldsymbol{r} \tag{256}$$

in the transition $m \rightarrow n$. Then using a matrix product to evaluate $\langle j_\mu(r, t) j_\mu(r', t') \rangle_{oo}$,

$$\frac{1}{a}\frac{da}{dt} = -\frac{1}{16\pi^3 \hbar c} \int \frac{d^3\boldsymbol{k}}{|\boldsymbol{k}|} \int_{-\infty}^{t} dt'$$

$$\times \sum_n \exp\left\{ \frac{i}{\hbar}(t - t')(E_o - E_n - \hbar c|\boldsymbol{k}|) \right\} \sum_\mu |j_\mu^k(n, 0)|^2 \tag{257}$$

where we made use of (248).

As before, the sum extends only over the two transverse polarizations μ, the other two cancelling each other exactly. Now we have to calculate

$$\int_{-\infty}^{0} e^{iax} \, dx = \pi\delta(a) + \frac{1}{ia} = 2\pi\delta_+(a) \tag{258}$$

this being the definition of the δ_+ function. Thus

$$\frac{1}{a}\frac{da}{dt} = -\frac{1}{8\pi^2 c} \int \frac{d^3\boldsymbol{k}}{|\boldsymbol{k}|} \sum_{n,\mu} |j_\mu^k(n, 0)|^2 \, \delta_+ (E_n - E_0 + \hbar c|\boldsymbol{k}|) \tag{259}$$

We write

$$\frac{1}{a}\frac{da}{dt} = -\frac{1}{2}\Gamma - \frac{i}{\hbar}\Delta E \tag{260}$$

Then ΔE and Γ are real constants given by

$$\Delta E = -\frac{\hbar}{16\pi^3 c} \int \frac{d^3\boldsymbol{k}}{|\boldsymbol{k}|} \sum_{n,\mu} \frac{|j_\mu^k(n, 0)|^2}{E_n - E_0 + \hbar c|\boldsymbol{k}|} \tag{261}$$

$$\Gamma = \frac{1}{8\pi^2 c} \int \frac{d^3\boldsymbol{k}}{|\boldsymbol{k}|} \sum_{n,\mu} |j_\mu^k(n, 0)|^2 \, \delta (E_n - E_0 + \hbar c|\boldsymbol{k}|) \tag{262}$$

These are independent of t. Therefore the amplitude of the state Φ_o in the wave-function $\Phi(t)$ is given for all $t \gg t_o$ by

$$a(t) = (\Phi_o^* \Phi(t)) = \exp\left\{-\frac{i}{\hbar}\Delta E\,(t - t_o) - \frac{1}{2}\Gamma\,(t - t_o)\right\} \qquad (263)$$

The state Φ_o, as a result of the perturbation by the radiation field, has its energy shifted by ΔE, and decays exponentially like

$$|a(t)|^2 = e^{-\Gamma\,(t - t_o)} \qquad (263a)$$

Comparing (232) and (262), we see that Γ is exactly the total probability per unit time of radiation from the state o to all other states n, calculated neglecting radiation reaction. This gives the physical interpretation for the decay law (263a). When the denominators in (261) have zeroes, the integration over $|\boldsymbol{k}|$ is to be taken as a Cauchy principal value. The energy shift ΔE is exactly what would be obtained from elementary second-order perturbation theory, if the difficulties arising from vanishing denominators were just ignored.

We calculate now the spectrum of radiation emitted in the transition from the level o to the level n, including effects of the level shifts ΔE_o and ΔE_n and the widths Γ_o and Γ_n. Let b_{nk} be the amplitude at time t, of the state in which the atom is in state n and the photon is present with propagation vector k. The equation of motion for b_{nk}, including effects of radiation out of the state n, is

$$\frac{db_{nk}}{dt} = \left\{-\frac{1}{2}\Gamma_n - \frac{i}{\hbar}\Delta E_n\right\} b_{nk}$$

$$- Q \exp\left\{\frac{i}{\hbar}\left(E_n - E_o + \hbar c|\boldsymbol{k}|\right) t\right\} a(t) \qquad (264)$$

with $a(t)$ given by (263).[40] Here the last term represents the effects of transitions $o \to n$, and Q is the space part of a matrix element of H_I^S which is independent of t and varies only slowly with k so that we can regard Q as a constant for all values of k within the line-width. The exponential is the time part of the matrix element, the exponent being proportional to the energy difference between the atom in state n plus the photon, and the atom in the zero state. The solution of (264) is, taking for convenience $t_o = 0$,

$$b_{nk} = A\left\{\exp(-\beta t) - \exp(-\gamma t)\right\} \qquad (265)$$

using the initial condition $b_{nk} = 0$ at $t = 0$. Here

$$
\begin{aligned}
\beta &= \frac{1}{2}\Gamma_o + \frac{i}{\hbar}\left(E_o + \Delta E_o - E_n - \hbar c|\boldsymbol{k}|\right) \\
\gamma &= \frac{1}{2}\Gamma_n + \frac{i}{\hbar}\Delta E_n
\end{aligned}
\tag{266}
$$

and $A = Q/(\beta - \gamma)$.

The probability that the atom leaves the state n by a second radiative transition at time t, so that a quantum k remains, is

$$
\Gamma_n |b_{nk}(t)|^2
$$

The quantum k remains from the first transition $o \to n$. After the atom makes the second transition to a continuum of possible states, the final states will no longer be coherent, and so the quanta left behind at different times t will not interfere with each other. The total probability for the emission of a quantum of frequency k in the first transition is thus

$$
P(k) = \Gamma_n |Q|^2 \frac{1}{|\beta - \gamma|^2} \int_0^\infty |e^{-\beta t} - e^{-\gamma t}|^2 \, dt
\tag{267}
$$

Now

$$
\frac{1}{|\beta - \gamma|^2} \int_0^\infty |e^{-\beta t} - e^{-\gamma t}|^2 \, dt
$$

$$
= \frac{1}{(\beta - \gamma)(\beta^* - \gamma^*)} \left\{ \frac{1}{\beta + \beta^*} + \frac{1}{\gamma + \gamma^*} - \frac{1}{\beta + \gamma^*} - \frac{1}{\beta^* + \gamma} \right\}
$$

$$
= \frac{1}{\beta - \gamma} \left\{ \frac{1}{(\gamma + \gamma^*)(\beta^* + \gamma)} - \frac{1}{(\beta + \beta^*)(\beta + \gamma^*)} \right\}
$$

$$
= \frac{\beta + \beta^* + \gamma + \gamma^*}{(\beta + \beta^*)(\gamma + \gamma^*)(\beta + \gamma^*)(\gamma + \beta^*)} = \frac{1}{2} \frac{\operatorname{Re}(\beta + \gamma)}{\operatorname{Re}(\beta)\operatorname{Re}(\gamma)|\beta + \gamma^*|^2}
\tag{268}
$$

Hence

$$
P(k) = |Q|^2 \frac{\Gamma_o + \Gamma_n}{\Gamma_o} \frac{\hbar^2}{(E_o + \Delta E_o - E_n - \Delta E_n - \hbar c|\boldsymbol{k}|)^2 + \frac{1}{4}\hbar^2(\Gamma_o + \Gamma_n)^2}
\tag{269}
$$

This formula for P(k) gives the natural shape of a spectral line. The maximum intensity occurs at

$$
\hbar c|\boldsymbol{k}| = (E_o + \Delta E_o) - (E_n + \Delta E_n)
\tag{270}
$$

i.e. at the difference between the energies of the two levels including the radiative level shifts. The width at half-maximum is

$$\hbar(\Gamma_o + \Gamma_n) \tag{271}$$

just the sum of the two widths of the levels given by their reciprocal lifetimes.

These formulae (270) and (271) are important in interpreting the modern radio-frequency spectroscopic experiments with their very accurate measurements of line shapes and positions.

5.2.3   *Calculation of Line-Shift, Non-Relativistic Theory*

In all atomic systems, the line widths are finite and easily calculated from the known transition amplitudes. For this, non-relativistic theory is accurate enough for all purposes. The line-shift (261) is much more difficult, and non-relativistic theory is not accurate enough to handle it properly. Still we shall calculate (261) using the non-relativistic theory, to see what it gives. It turns out to give quite a lot that is interesting.

First, in a non-relativistic calculation we use the dipole approximation which we also used to derive (234). Supposing a one-electron atom, the electron having mass m and charge $-e$, we put

$$j_1^k(n0) = -\frac{e}{m}(p_1)_{n0} = -\frac{e}{m}\left\{\int \psi_n^* \left(-i\hbar\frac{\partial}{\partial x}\right)\psi_0\, d^3\boldsymbol{r}\right\} \tag{272}$$

The line shift (261) becomes

$$\Delta E = -\frac{e^2\hbar}{16\pi^3 m^2 c}\int \frac{d^3\boldsymbol{k}}{|\boldsymbol{k}|}\sum_n \frac{|(p_1)_{n0}|^2 + |(p_2)_{n0}|^2}{E_n - E_o + \hbar c|\boldsymbol{k}|}$$

and integrating over the direction of $\boldsymbol{k}$, (compare with (238)),

$$\Delta E = -\frac{e^2\hbar}{6\pi^2 m^2 c}\int_0^\infty |\boldsymbol{k}|\, d|\boldsymbol{k}|\sum_n \frac{|\boldsymbol{p}_{n0}|^2}{E_n - E_o + \hbar c|\boldsymbol{k}|} \tag{273}$$

The integral over $|\boldsymbol{k}|$ is now obviously divergent, even before summing over n. Therefore the line-shift is *infinite*. When a complete relativistic theory with positrons is used, the divergence becomes only logarithmic instead of linear, but it still definitely diverges. This was for many years a disaster which destroyed all faith in the theory, and no way of avoiding the difficulty was found until 1947.

5.2.4   *The Idea of Mass Renormalization*

The line shift (273) is also infinite for a *free* electron of momentum $\boldsymbol{p}$. In this case $\boldsymbol{p}$ is a diagonal operator and the sum over n reduces to the term $n = 0$. Therefore

$$\Delta E_F = -\frac{1}{6\pi^2}\frac{e^2}{m^2c^2}\left(\int_0^\infty d|\boldsymbol{k}|\right)\boldsymbol{p}^2 \tag{274}$$

The effect of the radiation interaction is just to give a free electron an additional energy proportional to its kinetic energy $(\boldsymbol{p}^2/2m)$. If the integral in (274) is cut off at an upper limit $K \sim (mc/\hbar)$ in order to allow for the fact that the theory anyway is wrong in the relativistic region; then

$$\Delta E_F \approx -\frac{1}{6\pi^2}\frac{e^2}{\hbar c}\frac{\boldsymbol{p}^2}{m}$$

is a small correction to the kinetic energy, which would be produced by an increase in the rest-mass of the electron from m to $(m + \delta m)$,

$$\delta m = \frac{1}{3\pi^2}\frac{e^2}{c^2}\int_0^\infty d|\boldsymbol{k}| \tag{275}$$

We must now take into account that the observed rest-mass of *any* electron, bound or free, is not m but $m + \delta m$. Therefore in (273) a part

$$-\frac{1}{6\pi^2}\frac{e^2}{c^2}\left(\int_0^\infty d|\boldsymbol{k}|\right)\langle\boldsymbol{p}^2\rangle_{oo} \tag{276}$$

expresses only the effect of the mass-change δm on the kinetic energy of the bound electron; this part is already included in the kinetic energy of the electron, when the observed mass $(m + \delta m)$ is taken for the mass in the formula $(\boldsymbol{p}^2/2m)$. Therefore the part (276) has to be subtracted from (273), to give the observable line-shift. The subtraction just cancels out the error that was made in identifying the mass m of a "bare" electron without electromagnetic interaction with the observed electronic mass.

The idea of this mass-renormalization is that, although the "bare" mass m appears in the original description of the atom without radiation field, all the final results of the theory should depend only on the physically observable $m + \delta m$. The idea is originally due to Kramers [16], developed by Bethe (*Phys. Rev.* **72** (1947) 339.)

Subtracting (276) from (273) gives the physically observable line-shift

$$\Delta E = \frac{e^2}{6\pi^2 m^2 c^2} \int_0^\infty d|\boldsymbol{k}| \sum_n \frac{(E_n - E_o)\,|\boldsymbol{p}_{no}|^2}{E_n - E_o + \hbar c|\boldsymbol{k}|} \tag{277}$$

The divergence at high $|\boldsymbol{k}|$ is now only logarithmic. Taking an upper limit cut-off for the integral at the point

$$\hbar c|\boldsymbol{k}| = K$$

where K is an energy on the order of magnitude of mc^2, we have

$$\Delta E = \frac{e^2}{6\pi^2 m^2 c^3 \hbar} \sum_n (E_n - E_o)\,|\boldsymbol{p}_{no}|^2 \log \frac{K}{E_n - E_o} \tag{278}$$

remembering that the integration over $|\boldsymbol{k}|$ in (277) is to be taken as a Cauchy principal value when $(E_n - E_0)$ is negative.

From this formula (278) the line-shift for hydrogen states can be calculated numerically, as was done by Bethe, Brown and Stehn (*Phys. Rev.* **77** (1950) 370.)

Since the log in (278) will be quite large (~ 7) for states n which are in the non-relativistic range, it is convenient to write

$$\sum_n (E_n - E_o)\,|\boldsymbol{p}_{no}|^2 \log|E_n - E_o| = \left\{ \sum_n (E_n - E_o)\,|\boldsymbol{p}_{no}|^2 \right\} \log (E - E_o)_{\mathrm{av}} \tag{279}$$

this being the definition of $(E - E_o)_{\mathrm{av}}$. Then $(E - E_o)_{\mathrm{av}}$ is a non-relativistic energy. Exact calculation gives for the 2s state in hydrogen

$$(E - E_o)_{\mathrm{av}} = 16.6\,\mathrm{Ry} \tag{280}$$

Thus the important transitions are to states which although non-relativistic are continuum states with very high excitation. This is surprising.

Note that in (278) the terms are all positive if E_o is the ground state. For higher states there will be both positive and negative contributions. In particular, we shall see that for a Coulomb potential the positive and negative terms cancel almost exactly, for all except s states[41]. The cancellation is more or less accidental and seems to have no deeper meaning.

Now using a sum rule

$$\sum_n (E_n - E_o)\,|\boldsymbol{p}_{no}|^2 = \langle \boldsymbol{p} \cdot [H, \boldsymbol{p}] \rangle_{oo} \tag{281}$$

where H is the Hamiltonian for the atom[42]

$$H = \frac{1}{2m}\boldsymbol{p}^2 + V \qquad\qquad V = -\frac{1}{4\pi}\frac{e^2}{r}, \qquad\qquad [H, \boldsymbol{p}] = i\hbar(\nabla V) \quad (282)$$

$$\langle \boldsymbol{p} \cdot [H, \boldsymbol{p}] \rangle_{oo} = \hbar^2 \int \psi_o^* \nabla \cdot (\psi_o \nabla V)\, d\tau$$

$$= \frac{\hbar^2}{2} \left\{ \int \psi_o^* \nabla \cdot (\psi_o \nabla V)\, d\tau + \int \psi_o \nabla \cdot (\psi_o^* \nabla V)\, d\tau \right\}$$

$$= \frac{\hbar^2}{2} \left\{ 2 \int \psi_o^* \psi_o \nabla^2 V\, d\tau + \int \nabla (\psi_o^* \psi_o) \cdot \nabla V\, d\tau \right\}$$

$$= \frac{\hbar^2}{2} \int \psi_o^* \psi_o \nabla^2 V\, d\tau = \frac{1}{2} e^2 \hbar^2 |\psi_o(0)|^2 \qquad\qquad (283)$$

where we used Green's vector theorem, the fact that $\langle \boldsymbol{p} \cdot [H, \boldsymbol{p}] \rangle_{oo}$ is real, and the result that $\nabla^2 V = e^2 \delta^3(\boldsymbol{r})$ in Heaviside units.
Hence[43]

$$\Delta E = \frac{e^4 \hbar}{12\pi^2 m^2 c^3} |\psi_o(0)|^2 \log \frac{K}{(E - E_o)_{\mathrm{av}}}$$

$$= \frac{e^4 \hbar}{12\pi^2 m^2 c^3} \log \frac{K}{(E - E_o)_{\mathrm{av}}} \times \begin{cases} 1/(\pi n^3 a_o^3) & \text{for } s \text{ states} \\ 0 & \text{for others} \end{cases} \qquad (284)$$

Compare this with (239). It differs only in having the $\log(K/(E - E_o)_{\mathrm{av}})$ replacing $\log(1/RK_H)$. The low frequency photons have now been treated properly instead of being estimated. Only the high-frequency end is still inaccurate because of the vagueness of the cut-off K. Taking $K = mc^2$, (284) gives for the Lamb shift $2s - 2p$ the value *1040 Megacycles*. Remarkably close to the experimental value of 1062.

The success of this calculation of the line shift shows that the correct treatment of the interaction between an electron and the Maxwell field, with the help of the idea of mass renormalization, will give sensible results in agreement with experiment. This calculation could be done non-relativistically because the line shift is mainly a low-frequency and non-relativistic effect.

There are other effects of the radiation interaction, especially the anomalous increase in the observed magnetic moment of the electron by a factor

$\left(1 + \frac{\alpha}{2\pi}\right)$ over the value given by the Dirac theory, which are essentially relativistic in character. For studying these effects, and for calculating the Lamb shift accurately without an arbitrary cut-off, we need to use a complete relativistic quantum electrodynamics, in which both electrons and the Maxwell field are handled relativistically.

Therefore we must go back to where we left off the theory of the Dirac electron on page 31, and start to construct a relativistic field theory of electrons and positrons, similar to the quantized Maxwell field theory.

5.3 Field Theory of the Dirac Electron, Without Interaction

We apply to the Dirac equation the method of field quantization for anti-commuting fields. The reason why we must do this, and not use commuting fields, we will see later. Write

$$\mu = (mc/\hbar), \qquad m = \text{electron mass}$$

Lagrangian

$$\mathscr{L}_o = -\hbar c \, \overline{\psi} \left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} + \mu \right) \psi \tag{285}$$

Note the factor $\hbar$ here. This means that the theory has no classical limit in the sense of the Correspondence Principle. In the classical limit, only charges and currents composed of many particles have any meaning; the ψ field disappears entirely from view. The $\hbar$ has to be put into (285) to make the dimensions right, since $(\overline{\psi}\psi)$ has dimensions (1/Volume) just as in the 1-particle Dirac theory of which this is an extension. Field equations

$$\sum_\lambda \gamma_\lambda \frac{\partial \psi}{\partial x_\lambda} + \mu \psi = 0$$

$$\sum_\lambda \frac{\partial \overline{\psi}}{\partial x_\lambda} \gamma_\lambda - \mu \overline{\psi} = 0 \tag{286}$$

The charge-conjugate field ϕ can be defined by

$$\phi = C\psi^+$$

according to (51), and it also satisfies

$$\left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} + \mu \right) \phi = 0 \tag{287}$$

5.3.1   *Covariant Commutation Rules*

We proceed as for the Maxwell field. Take two points z and y with $z_0 > y_0$. Let

$$\mathcal{Q}(y) = \overline{\psi}(y)\, u$$
$$\mathcal{R}(z) = \overline{v}\,\psi(z) \quad \text{or} \quad \overline{\psi}(z)\, v \tag{288}$$

Here u and v are spinor operators, not depending on y or z and anticommuting with all the ψ and $\overline{\psi}$ operators in our equations, as we assumed at the beginning of this section. For example take $u = \psi(w)$ where w is a point far away outside the light-cones of both y and z. We make a change in the Lagrangian by

$$\delta_\mathcal{Q}\mathscr{L} = \epsilon\delta^4(x - y)\,\overline{\psi}(y)\, u \tag{289}$$

The factor u must be put in to make $\delta_\mathcal{Q}\mathscr{L}$ a bilinear expression, which is necessary for the Peierls method to be applicable. In fact only bilinear expressions have a physically observable meaning, and it is never meaningful under any circumstances to add together a term linear and a term bilinear in the field operators.

The changed field equations for ψ and $\overline{\psi}$ are

For $\overline{\psi}:$ No change

$$\text{For } \psi: \quad \left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} + \mu\right)\psi - \tfrac{\epsilon}{\hbar c}\,\delta^4(x - y)\, u = 0 \tag{290}$$

Thus $\delta_\mathcal{Q}\overline{\psi}(z) = 0$ and $\epsilon\delta_\mathcal{Q}\psi(z)$ satisfies (290). (Compare with (198).)

Hence $\delta_\mathcal{Q}\psi(z)$ is defined by the conditions

$$\left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} + \mu\right)\delta_\mathcal{Q}\psi(z) = \frac{1}{\hbar c}\,\delta^4(z - y)\, u \tag{291}$$

$$\delta_\mathcal{Q}\psi(z) = 0 \quad \text{for } z_0 < y_0$$

From (291), $\delta_\mathcal{Q}\psi(z)$ is a c-number spinor. We write

$$\delta_\mathcal{Q}\psi(z) = -\frac{1}{\hbar c}\, S_R(z - y)\, u \tag{292}$$

Then $S_R(x)$ is a c-number Dirac matrix function of x, satisfying

$$\left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} + \mu\right) S_R(x) = -\delta^4(x) \tag{293}$$

$$S_R(x) = 0 \quad \text{for } x_0 < 0$$

and the 4×4 unit matrix $\mathbb{I}$ is understood on the right of (293).

If we have $\mathcal{R} = \overline{\psi}(z)\,v$ then $\delta\overline{\psi} = 0$, as before. If $\mathcal{R} = \overline{v}\,\psi(z)$ then

$$\delta_{\mathcal{R}}\overline{\psi}(y) = -\frac{1}{\hbar c}\,\overline{v}\,S_A(z - y)$$

$$\delta_{\mathcal{R}}\psi(y) = 0$$

(294)

where $S_A(x)$ is the Dirac matrix satisfying

$$\left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} + \mu\right) S_A(x) = -\delta^4(x)$$

$$S_A(x) = 0 \quad \text{for } x_0 > 0$$

(295)

Finally, if we had chosen $\mathcal{Q} = \overline{u}\,\psi(y)$ we should have had $\delta_{\mathcal{Q}}\psi(z) = 0$ in the same way.

Hence using the Peierls commutation law (194) with (292) and (294) we find

$$[\overline{v}\,\psi(z)\,,\,\overline{u}\,\psi(y)] = [\overline{\psi}(z)\,v\,,\,\overline{\psi}(y)\,u] = 0$$

$$[\overline{v}\,\psi(z)\,,\,\overline{\psi}(y)\,u] = -i\overline{v}\,[S_A(z - y) - S_R(z - y)]\,u$$

(296)

These hold for every choice of u and v if we now choose u and v anti-commuting with all the ψ and $\overline{\psi}$ operators. Therefore writing

$$S(x) = S_A(x) - S_R(x)$$

(297)

we can write down the commutation rules for the operator components:

$$\{\psi_\alpha(z),\,\psi_\beta(y)\} = \{\overline{\psi}_\alpha(z),\,\overline{\psi}_\beta(y)\} = 0$$

(298)

$$\{\psi_\alpha(z),\,\overline{\psi}_\beta(y)\} = -iS_{\alpha\beta}(z - y)$$

(299)

The invariant $S-$function satisfies by (293) and (295)

$$\left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} + \mu\right) S(x) = 0$$

(300)

There are no simple formulae for the S-functions like (261) and (265) for the D-functions, in coordinate space. However in momentum representations the S-functions are equally simple.

5.3.2 *Momentum Representations*

Write

$$S_R(x) = \left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} - \mu \right) \Delta_R(x), \quad S_A(x) = \left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} - \mu \right) \Delta_A(x)$$

$$S(x) = \left(\sum_\lambda \gamma_\lambda \frac{\partial}{\partial x_\lambda} - \mu \right) \Delta(x) \tag{301}$$

Then we have

$$\left(\Box^2 - \mu^2 \right) \Delta_R(x) = \left(\Box^2 - \mu^2 \right) \Delta_A(x) = -\delta^4(x)$$
$$\left(\Box^2 - \mu^2 \right) \Delta(x) = 0 \qquad \Delta(x) = \left(\Delta_A - \Delta_R \right)(x) \tag{302}$$

with the boundary conditions as before. And the Δ-functions are exact analogues of the D-functions, the D-functions being the special case $\mu = 0$. Instead of (207) we have by making the formal substitution $k^2 \to k^2 + \mu^2$

$$\Delta_R(x) = \frac{1}{(2\pi)^4} \int_+ e^{ik \cdot x} \frac{1}{k^2 + \mu^2} \, d^4k \tag{303}$$

where the contour in the k_0-plane goes above the two poles at $k_0 = \pm\sqrt{|\boldsymbol{k}|^2 + \mu^2}$. Similarly for (208). And instead of (210)

$$\Delta(x) = -\frac{i}{(2\pi)^3} \int e^{ik \cdot x} \, \delta(k^2 + \mu^2) \, \epsilon(k) \, d^4k \tag{304}$$

Hence using (301) and the notation (110)

$$S(x) = \frac{1}{(2\pi)^3} \int e^{ik \cdot x} \left(\not{k} + i\mu \right) \delta(k^2 + \mu^2) \, \epsilon(k) \, d^4k \tag{305}$$

Note the projection operator Λ_+ appearing here, as defined by (115) with the momentum $p = \hbar k$. Thus the S-function distinguishes automatically between the electron states $k_0 = +\sqrt{|\boldsymbol{k}|^2 + \mu^2}$ and the positron states $k_0 = -\sqrt{|\boldsymbol{k}|^2 + \mu^2}$

5.3.3 *Fourier Analysis of Operators*

We analyze ψ_α into Fourier components, written in a quite general form:

$$\psi_\alpha(x) = Q \int d^3\boldsymbol{k} \left(\frac{\mu^2}{|\boldsymbol{k}|^2 + \mu^2} \right)^{1/4} \left\{ \sum_{u+} u_\alpha e^{ik \cdot x} \, b_{ku} + \sum_{u-} u_\alpha e^{-ik \cdot x} \, b_{ku} \right\}$$
$$\tag{306}$$

where, as in (211), the factor $\left(\mu^2/(|\boldsymbol{k}|^2 + \mu^2)\right)^{1/4}$ only makes the notation simpler. The integration is over all 4-vectors k with $k_0 = +\sqrt{|\boldsymbol{k}|^2 + \mu^2}$. For each k, the sum $\sum_{u+}$ is over the two spin-states u satisfying, by (111), the equation

$$(\not{k} - i\mu) \, u = 0 \tag{307}$$

and the sum $\sum_{u-}$ is over the two spin-states u satisfying by (112)

$$(\not{k} + i\mu) \, u = 0 \tag{308}$$

the normalization being given by (106) and (113). The b_{ku} are operators independent of x and α, whose properties are to be determined.

Taking the adjoint to (306) we have

$$\overline{\psi}_\alpha(x) = Q \int d^3\boldsymbol{k} \left(\frac{\mu^2}{|\boldsymbol{k}|^2 + \mu^2}\right)^{1/4} \left\{ \sum_{u+} b_{ku}^* \overline{u}_\alpha e^{-ik\cdot x} + \sum_{u-} b_{ku}^* \overline{u}_\alpha e^{ik\cdot x} \right\} \tag{309}$$

Here the b_{ku}^* are ordinary Hermitian conjugates of b_{ku}.

Computing the anti-commutators (298), (299) from (306) and (309), and comparing the results with the momentum integral (305), using (115) and the properties of Λ_+ we have

$$\{\, b_{ku}, \, b_{k'v} \,\} = \{\, b_{ku}^*, \, b_{k'v}^* \,\} = 0 \tag{310}$$

$$\{\, b_{ku}, \, b_{k'v}^* \,\} = \delta^3(\boldsymbol{k} - \boldsymbol{k}') \, \delta_{uv} \tag{311}$$

and we find that the constant Q in (306) and (309) is given by[44]

$$Q = (2\pi)^{-3/2} \tag{312}$$

5.3.4 *Emission and Absorption Operators*

Let

$$E_k = \hbar c \sqrt{|\boldsymbol{k}|^2 + \mu^2} \tag{313}$$

be the energy of an electron or a positron with momentum $\hbar k$. We apply the same argument which led to (215) for the Maxwell field. From this it follows that

$$b_{ku} \text{ for electron states } u$$
$$b_{ku}^* \text{ for positron states } u$$

have matrix elements only for transitions from an initial state of energy E_1 to a final state of energy E_2, where

$$E_1 - E_2 = E_k \qquad (314)$$

And

$$b_{ku} \text{ for positron states } u$$
$$b_{ku}^* \text{ for electron states } u$$

have matrix elements which are non-zero only when

$$E_2 - E_1 = E_k \qquad (315)$$

Thus we see as before that the field has the properties we require of a quantized field. It can carry energy only in discrete lumps of magnitude E_k for each frequency k. And the energy can be carried by two kinds of excitation, which we have called electrons and positrons anticipating the later results of the theory. We see already that these two excitation states have particle properties, and that there are two kinds of particles.

The absorption operators are

$$b_{ku} \text{ for electrons}$$
$$b_{ku}^* \text{ for positrons}$$

the emission operators are

$$b_{ku} \text{ for positrons}$$
$$b_{ku}^* \text{ for electrons}$$

5.3.5   *Charge-Symmetrical Representation*

We use the charge-conjugate field ϕ defined by (51), in order to put the whole theory into a form where there is complete symmetry between electrons and positrons. This symmetry is known as the charge-symmetry of the theory.

Let k be given, and a spinor u satisfying (308), representing a positron state. We represent the positron state alternatively by the charge conjugate spinor

$$v = Cu^+ \qquad (316)$$

which will satisfy (307) as u does for electron states. We denote by

$$b_{kv}^C = b_{ku}^*$$

the absorption operator for the positron state v. Then instead of (306), (309) we may write the pair of equations

$\psi_\alpha(x)$

$$= Q \int d^3\mathbf{k} \left(\frac{\mu^2}{|\mathbf{k}|^2 + \mu^2} \right)^{1/4} \left\{ \sum_{u+} u_\alpha e^{i\mathbf{k}\cdot\mathbf{x}} b_{ku} + \sum_{v+} \{Cv^+\}_\alpha e^{-i\mathbf{k}\cdot\mathbf{x}} b_{kv}^{*C} \right\}$$

(317)

$\phi_\alpha(x)$

$$= Q \int d^3\mathbf{k} \left(\frac{\mu^2}{|\mathbf{k}|^2 + \mu^2} \right)^{1/4} \left\{ \sum_{v+} v_\alpha e^{i\mathbf{k}\cdot\mathbf{x}} b_{kv}^{C} + \sum_{u+} \{Cu^+\}_\alpha e^{-i\mathbf{k}\cdot\mathbf{x}} b_{ku}^{*} \right\}$$

(318)

The ψ and ϕ fields are thus entirely symmetrical between positrons and electrons; ϕ could be taken as the starting-point and ψ derived from it, just as easily as vice versa.

The commutation rules (311) become

$$\{ b_{ku}, b_{k'u'}^{*} \} = \delta^3(\mathbf{k} - \mathbf{k}') \, \delta_{uu'}$$
$$\{ b_{kv}^{C}, b_{k'v'}^{*C} \} = \delta^3(\mathbf{k} - \mathbf{k}') \, \delta_{vv'}$$
$$\{ b_{kv'}^{C}, b_{ku}^{*} \} = 0, \qquad \text{etc.}$$

(318a)

These are also symmetrical between electron and positron.

5.3.6   *The Hamiltonian*

The Hamiltonian H has commutation rules with the emission and absorption operators as for the Maxwell field. These rules come straight from the Heisenberg equations of motion for ψ and ϕ. For any electron state u or positron v

$$[\, b_{ku}, H \,] = E_k b_{ku} \qquad\qquad [\, b_{kv}^{C}, H \,] = E_k b_{kv}^{C}$$
$$[\, b_{ku}^{*}, H \,] = -E_k b_{ku}^{*} \qquad [\, b_{kv}^{*C}, H \,] = -E_k b_{kv}^{*C}$$

(319)

Hence the Hamiltonian of the theory is

$$H = \int d^3\mathbf{k} \, E_k \left\{ \sum_{u+} b_{ku}^{*} b_{ku} + \sum_{v+} b_{kv}^{*C} b_{kv}^{C} \right\}$$

(320)

as one can verify at once by substituting into (319).

The additive constant is again chosen so that $\langle H \rangle_o$, the expectation value of H in the vacuum state, is zero. This eliminates a possible arbitrary additive constant from H.

In (317), (318) and (320) there is complete symmetry between electrons and positrons. The theory could just have well been built up from the positron as the fundamental particle, instead of from the electron.

But for practical calculations we shall generally not use (317), (318), (320). It is generally easier to work with the unsymmetrical form of the theory, with the fields ψ and $\overline{\psi}$.

5.3.7   *Failure of Theory with Commuting Fields*

Suppose we had taken the theory up to this point, only assuming ψ and $\overline{\psi}$ to be ordinary commuting fields. Then u and v would be taken to be quantities commuting with all ψ and $\overline{\psi}$, in the relations (296). Thus (298) and (299) would still be true only with commutators everywhere instead of anticommutators. Likewise (310) and (311). However in this symmetric representation, instead of (318a) we should have

$$
\begin{aligned}
[\, b_{ku}, \, b^*_{k'u'} \,] &= \delta^3(\boldsymbol{k} - \boldsymbol{k}')\, \delta_{uu'} \\
[\, b^C_{kv}, \, b^{*C}_{k'v'} \,] &= -\delta^3(\boldsymbol{k} - \boldsymbol{k}')\, \delta_{vv'}
\end{aligned}
\tag{321}
$$

And then the Hamiltonian instead of being given by (320) would have to be

$$
H = \int d^3\boldsymbol{k}\, E_k \left\{ \sum_{u^+} b^*_{ku} b_{ku} - \sum_{v^+} b^{*C}_{kv} b^C_{kv} \right\}
\tag{322}
$$

So positrons would actually be particles of negative energy, like the negative energy electrons of the one-electron theory. This is physically inadmissible.

Thus the use of anticommuting fields is the only thing which gives us a proper positive energy for the positrons. This is reasonable, because the intuitive Dirac hole theory can only work by virtue of the Pauli exclusion principle, and the exclusion principle is a feature of anticommuting fields.

5.3.8   *The Exclusion Principle*

Take any creation operator b^*_{ku}. As a special case of (310) we have identically

$$
b^*_{ku} b^*_{ku} = 0
\tag{323}
$$

Given any state Ψ, the result of creating *two* electrons with frequency k and spin u in that state is $b^*_{ku} b^*_{ku} \Psi = 0$. Thus there are no states in which two electrons have the same momentum and spin. So we have the Pauli exclusion principle valid for both electrons and positrons. Also, an electron and a positron do not exclude each other.

It is a very great success of the general field theory, that it has given us the Pauli principle automatically, not by special hypothesis as in the old particle theory of electrons.[45]

The most general state of the fields is described by specifying for each electron and positron state the number of particles occupying it. This number in each case can take just the two values 0 and 1.

5.3.9   *The Vacuum State*

The vacuum state Ψ_o is defined by

$$b_{ku}\Psi_o = 0 \text{ and thus} \qquad \Psi_o^* b^*_{ku} = 0 \text{ for electron states } u$$
$$b^*_{ku}\Psi_o = 0 \text{ and thus} \qquad \Psi_o^* b_{ku} = 0 \text{ for positron states } u \qquad (323a)$$

So the vacuum expectation value of products of emission and absorption operators are given by (311). We find using (323a)

$$\langle b_{ku} b_{k'v} \rangle_o = \langle b^*_{ku} b^*_{k'v} \rangle_o = 0$$
$$\langle b_{ku} b^*_{k'v} \rangle_o = \Theta_u \, \delta^3(\boldsymbol{k} - \boldsymbol{k}') \, \delta_{uv} \qquad (324)$$
$$\langle b^*_{ku} b_{k'v} \rangle_o = (1 - \Theta_u) \, \delta^3(\boldsymbol{k} - \boldsymbol{k}') \, \delta_{uv}$$

where

$$\Theta_u = \begin{cases} = 1 & \text{for electron states } u, \\ = 0 & \text{for positron states.} \end{cases}$$

Hence by (306) and (309) the expectation value $\langle \psi_\alpha(z) \overline{\psi}_\beta(y) \rangle_o$ is just the part of the anti-commutator $\{\psi_\alpha(z), \overline{\psi}_\beta(y)\}$ which contains positive frequencies $\exp[ik \cdot (z - y)]$, with $k_o > 0$. Thus, similarly to (221)

$$\langle \psi_\alpha(z) \overline{\psi}_\beta(y) \rangle_o = -i S^+_{\alpha\beta}(z - y) \qquad (325)$$

$$S^+(x) = \frac{1}{(2\pi)^3} \int e^{ik \cdot x} \, (\not{k} + i\mu) \, \delta(k^2 + \mu^2) \, \Theta(k) \, d^4k \qquad (326)$$

where

$$\Theta(x) = \begin{cases} = +1 & \text{for } x_0 > 0 \\ = 0 & \text{for } x_0 < 0 \end{cases}$$

writing as before

$$S^+ = \frac{1}{2}\left(S - iS^{(1)}\right) \qquad S^- = \frac{1}{2}\left(S + iS^{(1)}\right) \tag{327}$$

$$\left\langle \overline{\psi}_\beta(y)\psi_\alpha(z) \right\rangle_o = -iS^-_{\alpha\beta}(z-y) \tag{328}$$

$$\left\langle [\,\psi_\alpha(z), \overline{\psi}_\beta(y)\,] \right\rangle_o = -S^{(1)}_{\alpha\beta}(z-y) \tag{329}$$

$$S^{(1)}(x) = \frac{i}{(2\pi)^3}\int e^{ik\cdot x}\,(\slashed{k} + i\mu)\,\delta(k^2 + \mu^2)\,d^4k \tag{330}$$

$$S^-(x) = -\frac{1}{(2\pi)^3}\int e^{ik\cdot x}\,(\slashed{k} + i\mu)\,\delta(k^2 + \mu^2)\,\Theta(-k)\,d^4k \tag{330a}$$

These results for the Dirac theory without electromagnetic interaction will be used a lot when we come to the complete quantum electrodynamics with both Dirac and Maxwell fields quantized. Meanwhile, we should say a little about the theory of quantized Dirac particles in a given c-number Maxwell field.

5.4   Field Theory of Dirac Electron in External Field

Lagrangian

$$\mathscr{L} = \mathscr{L}_D - ie\overline{\psi}\slashed{A}^e\psi \tag{331}$$

Field equations

$$\left\{\sum_\lambda \gamma_\lambda\left(\frac{\partial}{\partial x_\lambda} + \frac{ie}{\hbar c}A^e_\lambda\right) + \mu\right\}\psi = 0 \tag{332}$$

$$\sum_\lambda\left(\frac{\partial\overline{\psi}}{\partial x_\lambda} - \frac{ie}{\hbar c}A^e_\lambda\overline{\psi}\right)\gamma_\lambda - \mu\overline{\psi} = 0 \tag{333}$$

These equations are still *linear*, the A^e_μ being given functions of position. This makes the theory still simple.

5.4.1 *Covariant Commutation Rules*

Because of the linearity, a change $\overline{v}\,\psi(z)$ in $\mathscr{L}$ makes no difference to the field equation for $\psi(y)$. Therefore for every two space-time points y and z we still have, as in (298),

$$\{\psi_\alpha(z),\,\psi_\beta(y)\} = \{\overline{\psi}_\alpha(z),\,\overline{\psi}_\beta(y)\} = 0 \tag{334}$$

Beyond this point not much can be done with the theory for general time-dependent potentials A_μ^e. In practice when we have time-dependent A_μ^e we always use perturbation theory starting from free field formalism, assuming the A_μ^e to be small, or else use special tricks for particular problems.

The important practical cases in which the A_μ^e are not small are always those in which the A_μ^e are *time-independent* in a particular Lorentz system. Examples: electrons bound in atoms by static Coulomb forces, electrons moving in constant macroscopic electric and magnetic fields.

So we assume $A_\mu^e = A_\mu^e(r)$ time-independent. We also assume the A_μ^e to be physically well-behaved so that the stationary eigenvalue equation

$$E_n\psi_n = \left\{ -e\Phi + \sum_{j=1}^{3}\left(-i\hbar c\frac{\partial}{\partial x_j} + eA_j^e\right)\alpha^j + mc^2\beta \right\}\psi_n \tag{335}$$

regarded as an equation for the c-number Dirac wave-function $\psi_n(r)$ has a complete set of eigenfunctions ψ_n with eigenvalues E_n; the spectrum may be either discrete or continuous or mixed. Equation (335) is derived from (332) by substituting into it the particular function

$$\psi = \psi_n(r)\exp\left\{-i\frac{E_n}{\hbar}t\right\} \tag{336}$$

We assume further the potentials to be such that the eigenfunctions ψ_n separate clearly into two classes, the ψ_{n+} with positive E_n and the ψ_{n-} with negative E_n. This is again true for all physically occurring potentials, though it would fail for the Coulomb field of a point nucleus with charge $Z > 137$.

The time-independent potentials make the problem essentially non-covariant, and so we shall use the non-covariant notations in developing the theory. We write the equations as if all the levels n are *discrete*, thus $\sum_n$ means a sum over discrete levels plus an integral over continuous levels suitably normalized. We are now chiefly interested in discrete levels, and

so we need not bother to write the formulae for normalizing the continuous levels explicitly. This makes the picture apparently simpler than the free-particle theory, where the normalization of continuous levels was done at every stage accurately; the simplicity is only apparent, because we shall just be ignoring the complications arising from continuous levels.

The general solution of the field equations (332) is

$$\psi(r, t) = \sum_n b_n \psi_n(r) \exp\left\{-i\frac{E_n}{\hbar}t\right\} \tag{337}$$

where the b_n are operators independent of r and t, and E_n can be both positive and negative.

By (334) we have

$$\{b_m,\, b_n\} = 0 \qquad \{b_m^*,\, b_n^*\} = 0$$

where b_m^* is the Hermitian conjugate of b_m. We take the ψ_n normalized so that

$$\int \psi_m^*(r)\, \psi_n(r)\, d^3\boldsymbol{r} = \delta_{nm} \tag{338}$$

i.e. in the usual non-covariant way. Note especially in (339) that we are treating all levels as if they were discrete; this can be done for example by enclosing our whole system in a finite box.

We want still to find the continuous rules between ψ and ψ^*, or between b_n and b_m^*.[46] Suppose we add to the Lagrangian (331)

$$\delta\mathscr{L}(r, t) = \epsilon\delta(t - t_o)\, \psi^*(r, t_o)\, \psi_n(r)\, u \tag{339}$$

where u is an operator anticommuting with ψ and ψ^* as in (288). This produces a change in the field equation for ψ which now becomes

$$\left\{\sum_\lambda \gamma_\lambda\left(\frac{\partial}{\partial x_\lambda} + \frac{ie}{\hbar c}A_\lambda^e\right) + \mu\right\}\psi = \frac{\epsilon}{\hbar c}\,\delta(t - t_o)\beta\psi_n(r)u \tag{340}$$

So the change $\delta\psi$ produced in ψ by the addition of $\delta\mathscr{L}$ satisfies (341) with the initial condition $\delta\psi(r, t) = 0$ for $t < t_o$. (Compare with (198) and (290).) Now the solution of (341) will obviously be of the form

$$\delta\psi = a(t)\psi_n(r) \tag{341}$$

where $a(t)$ is a function of t only, because the right side of the linear equation also has this form. Substituting (342) into (341) and using (335) we have

$$\left(i\hbar\frac{\partial}{\partial t} - E_n\right)a(t) = -\epsilon\delta(t - t_o)u \tag{342}$$

and hence

$$\delta\psi = \frac{i\epsilon}{\hbar}\Theta(t - t_o)\psi_n(r)\exp\left\{-i\frac{E_n}{\hbar}(t - t_o)\right\}u \qquad (343)$$

as we can verify using

$$\frac{d}{dt}\Theta(t - t_o) = \delta(t - t_o)$$

Integrating (340) over space-time gives by (339)

$$c\iint \delta\mathscr{L}(r,t)\,d^3\boldsymbol{r}\,dt = \epsilon\,c\,b_n^*\exp\left\{i\frac{E_n}{\hbar}t_o\right\}u \qquad (344)$$

For $t > t_o$, (343) is the change produced in $\psi(r, t)$ by adding (344) to the action integral. Hence by the Peierls commutation rule, using (193), (343) and (344),

$$[\,b_n^*u, \psi(r, t)\,] = -\psi_n(r)\exp\left\{-i\frac{E_n}{\hbar}t\right\}$$

and therefore

$$\{b_n^*, \psi(r, t)\} = \psi_n(r)\exp\left\{-i\frac{E_n}{\hbar}t\right\} \qquad (345)$$

because by assumption the u's anticommute with the ψ's. The time t_o no longer appears in (345), checking the consistency of the method.

Multiplying (345) by $\psi_n^*(r')\exp\{-iE_nt'/\hbar\}$ and summing over n, we have[47]

$$\{\psi_\alpha(r, t), \psi_\beta^*(r', t')\} = \sum_n \psi_{n\alpha}(r)\psi_{n\beta}^*(r')\exp\left\{-i\frac{E_n}{\hbar}(t - t')\right\} \qquad (346)$$

This is the general commutation rule which reduces to (299) in the special case of free particles.

Multiplying (345) by $\psi_m^*(r)$ and integrating over r, we have

$$\{b_m, b_n^*\} = \delta_{nm} \qquad (347)$$

which is identical to (311) in the case of free particles when normalizations are properly treated.

5.4.2   *The Hamiltonian*

As before, the b_{n+} are absorption operators for electrons and the b_{n-}^* absorption operators for positrons, only the electrons and positrons being defined by the bound wave-functions. The vacuum state Ψ_o is given by

$$b_{n+}\Psi_0 = 0 \qquad\qquad b_{n-}^*\Psi_0 = 0 \qquad\qquad (348)$$

And the total Hamiltonian of the system, in order to give the correct commutators with the b_n and the b_n^* and also to have zero expectation value in the vacuum, is

$$H = \sum_{n+} E_n b_n^* b_n - \sum_{n-} E_n b_n b_n^* \qquad\qquad (349)$$

$$= \sum_{n+} E_n b_n^* b_n + \sum_{n-} |E_n| b_n b_n^* \qquad\qquad (350)$$

From this Hamiltonian it is clear that the system is just a superposition of non-interacting particle-states. In each particle-state independently there is a number of particles given by

$$N_n = b_n^* b_n \qquad \text{for electron states}$$
$$N_n = b_n b_n^* \qquad \text{for positron states}$$

From the commutation rules (338) and (347)

$$N_n^2 = N_n \qquad\qquad (351)$$

so each N_n has just the two eigenvalues 0, 1. This describes exactly the physical situation in a many-electron atom where each atomic level may be full or empty independent of the others.

If each N is represented by the (2×2) diagonal matrix

$$N_n = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad\qquad (352)$$

then we have

$$b_{n+} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad b_{n-}^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$b_{n-} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad b_{n-}^* = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad\qquad (353)$$

This gives an explicit matrix representation for the operators. Each of the states n has its own two-valued row and column index. Thus for an atom with M levels altogether the operators would be represented by $(2^M \times 2^M)$ matrices.

Once having got the Hamiltonian (350) and the stationary states ψ_n the theory of many-electron systems is completely straightforward. We see that the levels of the hydrogen atom given by the one-electron Dirac theory are still exactly valid in this many-electron theory. Only now the Hamiltonian (350) having positive eigenvalues, the negative energy states give us no disasters. The positrons appear with positive energy, so that all the results of the Dirac theory are given to us simply and automatically.

5.4.3   *Antisymmetry of the States*

We know that in elementary quantum theory of many-electron systems we have to represent the system wave-functions by determinants of one-particle wave-functions, so as to make the system wave-functions always antisymmetric in the particle coordinates. We shall no longer need to make any such arbitrary choice of wave-functions in the field theory, all the results of the antisymmetry being given automatically by the theory.

For example, consider an atom with 2 electrons in states ψ_1 and ψ_2, all other states being empty. Then the state of the system is given by

$$\Psi = b_1^* b_2^* \Psi_o \tag{354}$$

where Ψ_o is the vacuum state. In (354) there is no arbitrariness; an interchange of the indices 1 and 2 will only change Ψ into $-\Psi$ which means no physical change. Now consider a 2-particle interaction operator

$$V = \frac{1}{2} \iint d^3\boldsymbol{r}_1 \, d^3\boldsymbol{r}_2 \, \{\psi^*(r_1)\psi(r_1)\} \, V(r_1 - r_2) \, \{\psi^*(r_2)\psi(r_2)\} \tag{355}$$

For example, V may be the Coulomb potential between two electrons, which is not included in the Lagrangian (331). The $\frac{1}{2}$ is put in so as to count each pair of points r_1, r_2 only once. We calculate the matrix element of V for a transition from Ψ to a state

$$\Psi' = b_3^* b_4^* \Psi_o$$

where the 2 electrons are in two other states ψ_3 and ψ_4. This matrix element is

$$M = (\Psi_o^*, b_4 b_3 V b_1^* b_2^* \Psi_o) \tag{356}$$

Expanding V by (337) into a sum of products of b_n and b_n^*, a contribution to (356) will come only from the 4 terms in V proportional to $b_1 b_2 b_3^* b_4^*$. Using the anticommutation rules, we have

$$\left(\Psi'^*, b_1 b_2 V b_3^* b_4^* \Psi\right) = -1$$

$$\left(\Psi'^*, b_1 b_2 V b_4^* b_3^* \Psi\right) = 1, \quad \text{etc.,} \tag{357}$$

Hence adding up the 4 terms

$$M = \iint d^3 \boldsymbol{r}_1 \, d^3 \boldsymbol{r}_2 \, V(r_1 - r_2)\{\{\psi_3^*(r_1)\psi_1(r_1)\}\{\psi_4^*(r_2)\psi_2(r_2)\}$$

$$- \{\psi_3^*(r_1)\psi_2(r_1)\}\{\psi_4^*(r_2)\psi_2(r_2)\}\} \tag{358}$$

This is exactly the result, direct minus exchange interaction, which would be given by using antisymmetrized wave-functions.

The field theory thus gives the full force of the Fermi statistics for electrons. And we could have shown in the same way that it gives the Bose statistics for photons.

5.4.4 *Polarization of the Vacuum*

Because of the possibility of exciting the vacuum by creating a positron-electron pair, the vacuum behaves like a dielectric, just as a solid has dielectric properties in virtue of the possibility of its atoms being excited to excited states by Maxwell radiation. This effect does not depend on the quantizing of the Maxwell field, so we calculate it using classical fields.

Like a real solid dielectric, the vacuum is both non-linear and dispersive, i.e. the dielectric constant depends on the field intensity and on the frequency. And for sufficiently high frequencies and field intensities it has a complex dielectric constant, meaning it can absorb energy from the Maxwell field by real creation of pairs.

We calculate the dielectric constant only in the linear region, i.e. assuming weak fields. The critical field for this problem is

$$E_c = \frac{m^2 c^3}{e\hbar} \approx 10^{16} \, \frac{\text{Volts}}{\text{cm}} \qquad \left(\text{from } eE \cdot \tfrac{\hbar}{mc} \approx mc^2\right) \tag{359}$$

and in fact the linear theory is good enough for almost all problems. The important case where it is *not* good enough is the propagation of photons through the intense Coulomb field round a heavy nucleus like lead. Then the

non-linearity produces a scattering of photons which is small but has been detected experimentally by Wilson. [15]

We calculate the dispersive effects exactly, i.e. with no restriction on the frequency. Since the treatment is linear, the imposed Maxwell field may be supposed to be given by the potentials of a plane wave whose amplitude increases slowly with time

$$A_\mu^e(x) = e_\mu \exp\{iq \cdot x + \delta_o x_o\} \tag{360}$$

Here e and q are given vectors, δ_o is a small positive number. This exponentially increasing amplitude is put in so that the potential A_μ^e effectively acts only for a finite time before any given time at which observations may be made. This enables us to fix the initial conditions of the problem unambiguously. At the end of the calculation we shall go to the limit $\delta_o = 0$.

The vacuum polarization is the effect of the fluctuations of the quantized electron-positron field on a given Maxwell field. The Lamb shift is the effect of the fluctuations of the quantized Maxwell field on a given electron. The two effects are just opposites of each other, the roles of the two fields being interchanged. Thus we can treat the vacuum polarization now conveniently with the theory of the quantized electron field alone. The treatment will be relativistic and so more correct than the treatment given for the Lamb shift. Later, to have a complete theory of both effects, we shall quantize both fields together and consider the reaction of each on the other.

Historically the electron self-energy (Lamb shift) and the vacuum polarization were the two problems on which the theory broke down because of divergences. Schwinger showed that the vacuum polarization could be calculated and was finite, if one used the same kind of renormalization idea that made the Lamb shift also finite.

The electron field operator ψ_H in the field (360) satisfies (332). Here ψ_H is the operator in the Heisenberg representation. Now a solution of (332) correct to first order in A_μ^e is

$$\psi_H(x) = \psi(x) + \frac{ie}{\hbar c} \int dx' \, S_R(x - x') A\!\!\!/^e(x')\psi(x') \tag{361}$$

Here S_R is given by (293), (301), and (303), and $\psi(x)$ is a solution of the free-field equation (286). In fact $\psi(x)$ is the field operator of the interaction representation, when the effects of A_μ are represented in the wave-function instead of in the operators. Using the retarded potential in (361) means that the unperturbed states are specified in the past, as the initial states upon

which A_μ^e later produces effects. Thus the vacuum state defined by (323a) is the state in which initially no electrons or positrons are present. This is the state which we wish to study, and we call it Ψ_o.

Using the interaction representation operators $\psi(x)$, Ψ_o is the vacuum state and remains so for all time; the physical state is initially Ψ_o but does not remain Ψ_o. Using the Heisenberg operators $\psi_H(x)$, Ψ_o is the physical state for all time; it is initially the vacuum state but does not remain so. In the remote past as $x_0 \to -\infty$, because S_R is a retarded potential $\psi_H(x)$ and $\psi(x)$ become identical.

The expression (361) is useful because we know how to calculate matrix elements of $\psi(x)$ from the state Ψ_o, whereas the matrix elements of ψ_H do not have any simple form. We also need the adjoint equation

$$\overline{\psi}_H(x) = \overline{\psi}(x) + \frac{ie}{\hbar c} \int dx' \, \overline{\psi}(x') \slashed{A}^e(x') S_A(x' - x) \tag{362}$$

where $S_A(x)$ is given by (295).

The total current operator to first order in A_μ is

$$j_{\mu H}(x) = -iec \, \overline{\psi}_H(x) \gamma_\mu \psi_H(x)$$

$$= j_\mu(x) + \frac{e^2}{\hbar} \int d^4x' \{ \overline{\psi}(x) \gamma_\mu S_R(x - x') \slashed{A}^e(x') \psi(x')$$

$$+ \overline{\psi}(x') \slashed{A}^e(x') S_A(x' - x) \gamma_\mu \psi(x) \} \tag{363}$$

Here

$$j_\mu(x) = -iec \, \overline{\psi}(x) \gamma_\mu \psi(x) \tag{364}$$

is the current operator of the interaction representation. The vacuum expectation value

$$(\Psi_o^* \, j_\mu(x) \, \Psi_o) = \langle j_\mu(x) \rangle_o$$

$$= -iec \, \left\langle \sum_{\alpha,\beta} \overline{\psi}_\beta(x) (\gamma_\mu)_{\beta\alpha} \psi_\alpha(x) \right\rangle_o$$

$$= -iec \sum_{\alpha,\beta} (\gamma_\mu)_{\beta\alpha} \left\langle \overline{\psi}_\beta(x) \psi_\alpha(x) \right\rangle_o \tag{365}$$

is given by (328)

$$\langle j_\mu(x)\rangle_o = -ec\,\mathrm{Tr}\left\{\gamma_\mu S^-(0)\right\}$$

$$= \frac{ec}{(2\pi)^3}\int d^3\boldsymbol{k}\,\delta(k^2+\mu^2)\,\Theta(-k)\,\mathrm{Tr}\left\{\gamma_\mu\left[\not{k}-i\mu\right]\right\}$$

$$= \frac{4ec}{(2\pi)^3}\int d^3\boldsymbol{k}\,\delta(k^2+\mu^2)\,\Theta(-k)\,k_\mu \tag{366}$$

This is a highly divergent integral and is mathematically meaningless. This is one of the difficulties of the theory about which one can argue for a long time.

However there is no doubt that correct physical results are obtained by putting $\langle j_\mu(x)\rangle_o = 0$ simply. There are two good reasons for doing this.

(1) Physical. $\langle j_\mu(x)\rangle_o$, being the expectation value of the charge-current in the vacuum in the absence of all external fields, is known to be zero experimentally. Therefore if we calculated $\langle j_\mu(x)\rangle_o$ and found it were not zero we should simply define the current operator to be $j_\mu - \langle j_\mu\rangle_o$. With this definition the expectation value would become zero automatically.

(2) Mathematical. $\langle j_\mu(x)\rangle_o$ as we have calculated it is a vector, each component of which is a number quite independent of the coordinate system. There does not exist such a vector invariant under Lorentz transformations, except for this zero vector. Therefore $\langle j_\mu(x)\rangle_o = 0$ is the only assumption we *can* make which would keep the theory invariant.

This is a simple example of a method which has often to be used in quantum electrodynamics. When a calculation leads to a divergent integral or a mathematically indeterminate expression, we use physical arguments or arguments of Lorentz invariance to find a definite value for the quantity we are unable to calculate. This is the reason for the great success of the covariant formulation of electrodynamics introduced by Schwinger.

So using this principle we have by (328)

$$\langle j_{\mu H}(x)\rangle_o = -\frac{ie^2}{\hbar}\int d^4x'\,\mathrm{Tr}\left\{\not{A}^e(x')S^-(x'-x)\gamma_\mu S_R(x-x')\right.$$

$$\left. + \not{A}^e(x')S_A(x'-x)\gamma_\mu S^-(x-x')\right\} \tag{367}$$

We use the momentum representation (303) for S_R. But instead of taking the contour along the real axis for k_0 we can use a path of integration running along a straight line parallel to the real axis a distance δ_o above it. That

gives the momentum representation

$$e^{-\delta_0} S_R(x) = \frac{i}{(2\pi)^4} \int e^{ik\cdot x} \frac{k\!\!\!/ + i\delta\!\!\!/ + i\mu}{(k+i\delta)^2 + \mu^2} d^4k \tag{368}$$

where δ_0 is any positive real number, δ is the vector with components $(0,0,0,\delta_0)$ and the integral is along the real axis. The poles of (368) in the k_0 plane are displaced away from the real axis and so the integrand is free from singularities on the path of integration. Similarly

$$e^{+\delta_0} S_A(x) = \frac{i}{(2\pi)^4} \int e^{ik\cdot x} \frac{k\!\!\!/ - i\delta\!\!\!/ + i\mu}{(k-i\delta)^2 + \mu^2} d^4k \tag{369}$$

When using (368) and (369) we shall usually make $\delta_0 \to 0$ after carrying out the integrations, so that the convergence factors $e^{\pm\delta_0 x_0}$ will tend to 1 for every finite x.

Thus the momentum representation of (367) will be

$$\langle j_{\mu H}(x)\rangle_o = -\frac{e^2}{\hbar} \frac{1}{(2\pi)^7} \int d^4x' \iint d^4k_1\, d^4k_2$$

$$\times \exp\{iq\cdot x' + i(k_1 - k_2)\cdot(x' - x) + \delta_0 x_0\}$$

$$\times \left\{ \mathrm{Tr}\left\{ \phi\!\!\!/(k\!\!\!/_1 + i\mu)\gamma_\mu(k\!\!\!/_2 + i\delta\!\!\!/ + i\mu)\right\} \frac{\delta(k_1^2 + \mu^2)\,\Theta(-k_1)}{(k_2 + i\delta)^2 + \mu^2} \right.$$

$$\left. + \mathrm{Tr}\left\{ \phi\!\!\!/(k\!\!\!/_1 - i\delta\!\!\!/ + i\mu)\gamma_\mu(k\!\!\!/_2 + i\mu)\right\} \frac{\delta(k_2^2 + \mu^2)\,\Theta(-k_2)}{(k_1 - i\delta)^2 + \mu^2} \right\}$$

The integration of x' is immediate and gives $(2\pi)^4\delta^4(k_1 - k_2 + q)$. Hence

$$\langle j_{\mu H}(x)\rangle_o = -\frac{e^2}{(2\pi)^3\hbar} e^{iq\cdot x + \delta_0 x_0}$$

$$\times \int d^4k \left\{ \mathrm{Tr}\left\{ \phi\!\!\!/(k\!\!\!/ + i\mu)\gamma_\mu(k\!\!\!/ + q\!\!\!/ + i\delta\!\!\!/ + i\mu)\right\} \right.$$

$$\times \frac{\delta(k^2 + \mu^2)\,\Theta(-k)}{(k+q+i\delta)^2 + \mu^2} + \mathrm{Tr}\left\{ \phi\!\!\!/(k\!\!\!/ - i\delta\!\!\!/ + i\mu)\gamma_\mu(k\!\!\!/ + q\!\!\!/ + i\mu)\right\}$$

$$\left. \times \frac{\delta\{(k+q)^2 + \mu^2\}\,\Theta(-k-q)}{(k - i\delta)^2 + \mu^2} \right\} \tag{370}$$

Now consider the function[48]

$$F_\nu(k) = \text{Tr}\left\{ \phi(\not{k}+i\mu)\gamma_\nu(\not{k}+\not{q}+i\not{\phi}+i\mu) \right\} \frac{1}{(k^2+\mu^2)[(k+q+i\delta)^2+\mu^2]} \quad (371)$$

This has poles at 4 points in the k_0 plane

$$k_0 = \pm\sqrt{|\boldsymbol{k}|^2+\mu^2} \qquad k_0 = -q_0 - i\delta_0 \pm \sqrt{|\boldsymbol{k}+\boldsymbol{q}|^2+\mu^2} \qquad (372)$$

The integral in the expression (370) is just the sum of the residues at the two points[49]

$$k_0 = -\sqrt{|\boldsymbol{k}|^2+\mu^2} \qquad k_0 = -q_0 - i\delta_0 - \sqrt{|\boldsymbol{k}+\boldsymbol{q}|^2+\mu^2} \qquad (373)$$
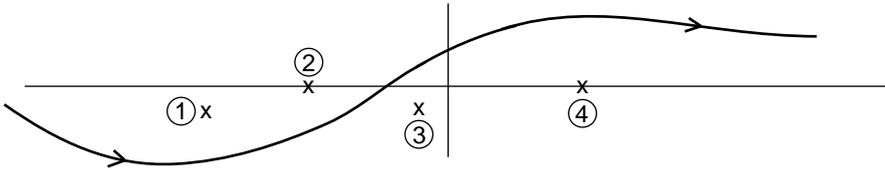
① $k_0 = -q_0 - i\delta - \sqrt{|\boldsymbol{k}+\boldsymbol{q}|^2+\mu^2}$ ② $k_0 = -\sqrt{|\boldsymbol{k}|^2+\mu^2}$

③ $k_0 = -q_0 - i\delta + \sqrt{|\boldsymbol{k}+\boldsymbol{q}|^2+\mu^2}$ ④ $k_0 = +\sqrt{|\boldsymbol{k}|^2+\mu^2}$

Hence

$$\langle j_{\nu H}(x)\rangle_o = \frac{ie^2}{(2\pi)^4\hbar} e^{iq\cdot x+\delta_0 x_0} \int_C F_\nu(k)\,dk \quad (374)$$

where $\int_C$ means a contour integral in the k_0 plane drawn as shown in the diagram, going from $-\infty$ to $+\infty$ below the two poles (373) and above the other two poles, and including the upper semicircle of the k_0 plane at infinity. So long as $\delta_0 > 0$ the poles are always clearly separated from each other. Now the calculation really starts, with the evaluation of this integral (374). This will be typical of all the calculations which have to be done in quantum electrodynamics using the modern methods.

5.4.5 *Calculation of Momentum Integrals*

Write $J_\nu = \int_C F_\nu(k)\,dk$. Then J_ν is a vector function of the variables μ (the mass of the electron) and

$$Q = q + i\delta \quad (375)$$

J_ν is certainly an analytic function of μ if $\delta > 0$, and it is an analytic function of Q if μ is large enough so that the poles (373) always lie to the left of the

imaginary axis and the other two poles (372) to the right. Therefore we can calculate J_ν for large values of μ and $\delta = 0$ so that $Q = q$.

To simplify (371) we take $\delta = 0$ and use Feynman's formula (*Phys. Rev.* **76** (1949) 785)

$$\frac{1}{ab} = \int_0^1 dz \, \frac{1}{[az + b(1-z)]^2} \tag{376}$$

which we can get at once simply from the definite integral of the RHS.[50] Hence

$$J_\nu = \int_0^1 dz \int_C dk \, \mathrm{Tr}\, \{\not{e}(\not{k} + i\mu)\gamma_\nu(\not{k} + \not{q} + i\mu)\} \frac{1}{\{k^2 + \mu^2 + z(2k \cdot q + q^2)\}^2}$$

$$= \int_0^1 dz \int_C dk \, \mathrm{Tr}\, \{\not{e}(\not{k} - z\not{q} + i\mu)\gamma_\nu(\not{k} + (1-z)\not{q} + i\mu)\}$$

$$\times \frac{1}{\{k^2 + \mu^2 + (z - z^2)q^2\}^2} \tag{377}$$

In the last step a shift of origin was made in the k–integration replacing k by $(k - zq)$. Again in (377) the poles in the k_0–plane for every z are well separated by the imaginary axis, provided that μ is large. Evaluating the spur and dropping terms are odd in k, and using (33),[51] $\mathrm{Tr}\, \gamma_\nu = 0$, $\mathrm{Tr}\, \gamma_\mu\gamma_\nu = 4\delta_{\mu\nu}$, we get

$$J_\nu = 4 \int_0^1 dz \int_C dk$$

$$\times \frac{e_\nu(-k^2 - \mu^2 + (z - z^2)q^2) + 2(e \cdot k)\, k_\nu - 2(z - z^2)(e \cdot q)\, q_\nu}{\{k^2 + \mu^2 + (z - z^2)q^2\}^2} \tag{378}$$

the odd terms going out because we can now take the k_0–integration straight up the imaginary axis from $-i\infty$ to $i\infty$ if we like. For the same reasons of symmetry we may replace

$$e \cdot k \, k_\nu \quad \text{by} \quad \frac{1}{4}k^2 e_\nu \qquad \text{because} \quad e \cdot k \, k_\nu = \sum_\lambda e_\lambda k_\lambda k_\nu \to e_\nu k_\nu k_\nu \to \frac{1}{4}e_\nu k^2$$

in the numerator, and obtain finally

$$J_\nu = 4 \int_0^1 dz \int_C dk \, \frac{e_\nu\{-\frac{1}{2}k^2 - \mu^2 + (z - z^2)q^2\} - 2(z - z^2)(e \cdot q)\, q_\nu}{\{k^2 + \mu^2 + (z - z^2)q^2\}^2} \tag{379}$$

This integral is still badly divergent. So we again use a physical argument to give a definite value to the most divergent part of it. The current operator both in interaction and in Heisenberg representation must satisfy

$$\sum_{\nu} \frac{\partial j_{\nu}(x)}{\partial x_{\nu}} = 0 \tag{380}$$

identically. Therefore (374) gives (since we now take $\delta = 0$)

$$\sum_{\nu} q_{\nu} J_{\nu} = 0 \tag{381}$$

which gives the relation

$$\int_0^1 dz \int_C dk \, \frac{-\frac{1}{2}k^2 - \mu^2 - (z - z^2)q^2}{\{k^2 + \mu^2 + (z - z^2)q^2\}^2} \equiv 0 \tag{382}$$

This equation (382) is a guiding equation telling us that a certain divergent expression appearing in (379) is to be given the value zero in order to make physical sense. We are left with

$$J_{\nu} = 8(q^2 e_{\nu} - e \cdot q \, q_{\nu}) \int_0^1 dz \, (z - z^2) \int_C \frac{dk}{\{k^2 + \mu^2 + (z - z^2)q^2\}^2} \tag{383}$$

For any positive Λ the integral

$$I_{\Lambda} = \int_C \frac{dk}{(k^2 + \Lambda)^3} \tag{384}$$

is convergent and can be evaluated by integrating for k_0 up the imaginary axis from $-i\infty$ to $+i\infty$. This gives (see the Appendix below)

$$I_{\Lambda} = i \iiiint \frac{dk_1 dk_2 dk_3 dk_0}{(k_1^2 + k_2^2 + k_3^2 + k_0^2 + \Lambda)^3} = 2\pi^2 i \int_0^{\infty} \frac{k^3 \, dk}{(k^2 + \Lambda)^3}$$

$$= \pi^2 i \int_0^{\infty} \frac{x \, dx}{(x + \Lambda)^3} = \frac{\pi^2 i}{2\Lambda} \tag{385}$$

Hence integrating with respect to Λ

$$\int_C dk \left\{ \frac{1}{(k^2 + \Lambda)^2} - \frac{1}{(k^2 + \mu^2)^2} \right\} = \pi^2 i \log\left(\frac{\mu^2}{\Lambda}\right) \tag{386}$$

this integral being also convergent. However

$$\int_C \frac{dk}{(k^2 + \mu^2)^2}$$

is logarithmically divergent for large k. Its value is

$$2i\pi^2 \log\left(\frac{k_{\max}}{\mu}\right) = 2i\pi^2 R \tag{387}$$

where R is the logarithmic factor and is independent of q.

Using (386) and (387) in (383), with $\Lambda = \mu^2 + (z - z^2)q^2$, we get

$$J_\nu = 8\pi^2 i(q^2 e_\nu - e \cdot q \, q_\nu)\left\{\frac{1}{3}R - \int_0^1 dz \, (z - z^2) \log\left[1 + \frac{(z - z^2)q^2}{\mu^2}\right]\right\} \tag{388}$$

This is the analytic formula for J_ν valid for large μ in which case the logarithm is real. We make the analytic continuation to small values of μ by writing $(q + i\delta)$ for q in (388), treating δ_0 as small and positive. Then q^2 becomes $q^2 - 2i\delta q_0$ and the logarithm becomes

$$\log\left|1 + \frac{(z - z^2)q^2}{\mu^2}\right| + \begin{cases} 0 & \text{for } \dfrac{(z - z^2)q^2}{\mu^2} > -1 \\[2ex] -i\pi \, \epsilon(q_0) & \text{for } \dfrac{(z - z^2)q^2}{\mu^2} < -1 \end{cases}$$

Writing now z for $4(z - z^2)$ and using (374), we go to the limit $\delta_0 = 0$ and find[52]

$$\langle j_{\nu H}(x)\rangle_o = -\frac{e^2}{2\pi^2\hbar}\left(q^2 e_\nu - e \cdot q \, q_\nu\right) e^{iq \cdot x}$$

$$\times \left\{\frac{1}{3}R - \frac{1}{8}\int_0^1 \frac{z \, dz}{\sqrt{1 - z}} \log\left|1 + \frac{zq^2}{4\mu^2}\right|\right.$$

$$\left. + \frac{i\pi}{8} \epsilon(q_0) \int_0^{-4\mu^2/q^2} \frac{z \, dz}{\sqrt{1 - z}}\right\} \tag{389}$$

the last term being zero except when

$$q^2 < -4\mu^2 \tag{390}$$

Now the external potential $A_\nu^e(x)$ is associated with a classical external charge-current density

$$j_{\nu E}(x) = -c \sum_\lambda \frac{\partial}{\partial x_\lambda} F_{\lambda\nu E}(x)$$

$$= -c \sum_\lambda \left\{\frac{\partial^2}{\partial x_\lambda^2} A_\nu^e(x) - \frac{\partial^2}{\partial x_\nu \partial x_\lambda} A_\lambda^e(x)\right\}$$

$$= c \left\{q^2 e_\nu - e \cdot q \, q_\nu\right\} e^{iq \cdot x} \tag{391}$$

Hence (389) gives the final result, with $\alpha = \frac{1}{137} = \frac{e^2}{4\pi\hbar c}$, (Heaviside units)

$$\langle j_{\nu H}(x)\rangle_o = -\alpha j_{\nu E}(x)\left\{\frac{2}{3\pi}R - \frac{1}{4\pi}\int_0^1 \frac{z\,dz}{\sqrt{1-z}}\log\left|1 + \frac{zq^2}{4\mu^2}\right|\right.$$

$$\left. + \frac{i}{4}\,\epsilon(q_0)\int_0^{-4\mu^2/q^2} \frac{z\,dz}{\sqrt{1-z}}\right\} \tag{392}$$

Appendix

The four dimensional volume element is (see (385)) $d\xi_1 d\xi_2 d\xi_3 d\xi_4 = 2\pi^2 r^3\,dr$ in four dimensional polar coordinates. To show this, we denote the surface of a p dimensional unit sphere by ω. Then the surface of a p dimensional sphere of radius R is $R^{p-1}\omega$, and hence the volume element in polar coordinates is $\omega R^{p-1}dR$.

To calculate the value of ω, we compute the p dimensional Laplace integral in Cartesian and polar coordinates. We have

$$J = \int\int \ldots \int \exp\left\{-\sum_{i=1}^p \xi_i^2\right\} d\xi_1 d\xi_2 \ldots d\xi_p = (\sqrt{\pi})^p$$

On the other hand

$$J = \omega \int_0^\infty e^{-\rho^2}\rho^{p-1}\,d\rho = \omega\frac{\Gamma(p/2)}{2}$$

and thus by comparison

$$\omega = \frac{2\pi^{p/2}}{\Gamma(p/2)}$$

For $p = 4$, we have $\Gamma(2) = 1$, $\omega = 2\pi^2$. For $p = 3$, we have $\Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$, $\omega = \frac{2\pi^{3/2}}{\sqrt{\pi}/2} = 4\pi$, etc.

5.4.6 *Physical Meaning of the Vacuum Polarization*

We now discuss the various physical effects arising from the calculation of

$$\langle j_{\nu H}(x)\rangle_o$$

(1) The result is completely gauge-invariant. This may be seen at once from (391): for if a gradient $\partial\Lambda/\partial x_\nu$ is added to A_ν^e, there is no change in $j_{\nu E}$.

(2) If the relation (382) had not been used in order to simplify (379), we should have had an addition to $\langle j_{\nu H}(x) \rangle_o$ of the form $K' e_\nu = K A_\nu^e(x)$ (see (360)) where K is an indeterminate numerical factor containing the divergent integral standing on the left of (382). That is to say, an induced current proportional to the inducing potential. This would have destroyed the gauge-invariance of the result if $K \neq 0$. Therefore we can also use the physical requirement that the results must be gauge-invariant in order to give the value zero to the indeterminate K.

(3) The energy density of the vacuum arising from the polarization by the potentials $A_\nu^e(x)$ is

$$d(x) = -\frac{1}{2c} \sum_\nu A_\nu^e(x) \langle j_{\nu E}(x) \rangle_o \tag{393}$$

Thus the term $K A_\nu^e(x)$ would give an energy density

$$-\frac{K}{2c} \sum_\nu A_\nu^e(x) A_\nu^e(x) \tag{394}$$

associated with the electromagnetic potentials. This would give to the photon a finite rest-mass, and for this reason K is often referred to as the "self-energy of the photon". There is a lot of discussion in the literature of this photon self-energy. But since physical arguments lead us definitely to give K the value zero, we have no choice but to say that the photon self-energy also is zero. This is of course the result that any consistent theory of electrodynamics must give.

(4) The logarithmic divergence R is a real divergence and cannot be given the value 0 from physical arguments. However, it gives only an induced charge exactly proportional to the inducing external charge. It is never possible experimentally to separate the external charge from the proportional induced charge. Therefore in all measurements of the external charge, the measured charge will be not $j_{\nu E}(x)$ but

$$j_{\nu R}(x) = \left(1 - \frac{2\alpha}{3\pi} R\right) j_{\nu E}(x) \tag{395}$$

$j_{\nu R}$ here meaning "renormalized charge". Thus the effect of the term R in (392) is only change the unit in which the external charge is measured. We write the results in terms of the observed external charge $j_{\nu R}$ instead of the unobservable $j_{\nu E}$, this change of units being called "charge renormalization".

Note the similarity between charge and mass renormalization. In both cases a divergent effect is recognized as producing no observable phenomenon, because it only changes the value of one of the fundamental constants, in the one case the electron mass m and in the other the unit of charge e. Since m and e are things which are observed directly, the divergent effects disappear completely when the results are written in terms of the observed m and e. Thus (392) becomes

$$\langle j_{\nu H}(x)\rangle_o = \alpha j_{\nu R}(x)\left\{\frac{1}{4\pi}\int_0^1 \frac{z\,dz}{\sqrt{1-z}}\log\left|1+\frac{zq^2}{4\mu^2}\right|\right.$$

$$\left. -\frac{i}{4}\,\epsilon(q_0)\int_0^{-4\mu^2/q^2}\frac{z\,dz}{\sqrt{1-z}}\right\} \tag{396}$$

Everything here is now finite and observable.

(5) When $A_\nu^e(x)$ is the potential of a pure radiation field without sources, $j_{\nu R} = 0$ and so there is no polarization. Thus for every photon or freely travelling wave, the vacuum behaves like a true vacuum; there are no dielectric effects of any kind. This agrees with the common-sense idea of the vacuum. The result is however only true so long as the polarization can be treated as linear; if non-linear effects were included then two beams of light crossing the same region would produce a polarization current, giving rise to a "scattering of light by light". The light-by-light scattering has been calculated, it is not zero but is much too small to be observable.

(6) The induced current (396) consists of two components, the first in phase with the potential $A_\nu^e(x)$ and the second out of phase by $\pi/2$. Speaking of the vacuum as a circuit being driven by the potential $A_\nu^e(x)$, the first is an inductive effect, the second a resistive one. So only the second term gives an absorption of energy by the vacuum from the driving potentials.

We study the energy balance, remembering that the classical potentials $A_\nu^e(x)$ must always be real quantities like

$$A_\nu^e(x) = e_\nu \cos(q\cdot x) \tag{397}$$

and we take without loss of generality q_0 positive. Then (396) gives, taking $e\cdot q = 0$,

$$\langle j_{\nu H}(x)\rangle_o = e_\nu \left\{A\cos(q\cdot x) + B\sin(q\cdot x)\right\} \tag{398}$$

where A and B are real and

$$B = \frac{1}{4}\alpha cq^2 \int_0^{-4\mu^2/q^2}\frac{z\,dz}{\sqrt{1-z}} \tag{399}$$

The energy supplied by the potentials to the vacuum per unit volume and per unit time is

$$E = -\frac{1}{c} \sum_\nu \langle j_{\nu H}(x) \rangle_o \frac{\partial A_\nu^e(x)}{\partial t}$$

$$= -q_0 \sum_\nu e_\nu^2 \left[A \sin(q \cdot x) \cos(q \cdot x) + B \sin^2(q \cdot x) \right] \qquad (400)$$

From (400) one sees that the in-phase current gives no net absorption of energy, while the out-of-phase current gives a mean energy supply per unit time

$$\overline{E} = -\frac{1}{2} q_0 e^2 B = -\frac{\alpha c e^2 q^2 q_0}{8} \int_0^{-4\mu^2/q^2} \frac{z \, dz}{\sqrt{1-z}} \qquad (401)$$

If q does not satisfy (390), i.e. if

$$q_0 < \sqrt{4\mu^2 + |\boldsymbol{q}|^2} \qquad (402)$$

then $B = 0$ and $E = 0$, there is not enough energy in the vibrations of the field to create a real positron-electron pair whose rest-mass alone requires $2mc^2$, given that the field supplies in the interaction a momentum $\hbar k$ together with the energy $\hbar c q_0$.

However if (390) is satisfied, then there is enough energy for creation of real pairs, each real pair carrying energy $\hbar c q_0$. Since q is time-like and $e \cdot q = 0$, e is space-like and $(e^2) > 0$. This can be seen as follows:

$$e \cdot q = 0 = \boldsymbol{e} \cdot \boldsymbol{q} - e_0 q_0$$

Since q is timelike, we can use a Lorentz transformation which makes $\boldsymbol{q} = 0$; then, of course, $q_0 \neq 0$. But then we must have $e_0 = 0$, which means that e is spacelike. Therefore $\overline{E} > 0$, verifying that the potentials can never extract energy from the vacuum. And we have the probability per unit volume and per unit time that the potential (397) will create a real pair

$$w = \frac{\overline{E}}{\hbar c q_0} = -\frac{\alpha(e^2)(q^2)}{8\hbar} \int_0^{-4\mu^2/q^2} \frac{z \, dz}{\sqrt{1-z}} \qquad (403)$$

The result could of course have been found more easily by elementary methods. The thing I want to stress here is that the elementary real pair-creation processes are necessarily tied up with the less elementary vacuum-polarization effect given by the A term in (398), and the A term will exist

whether or not real pair production is possible. The situation is quite the same as the tying up of the elementary line-width effect in atomic spectra with the less elementary line-shifts; we discussed these effects in detail earlier. Therefore we have quite as much reason to take seriously the vacuum-polarization effect given by the in-phase current in (396), as we have to take seriously the Lamb shift itself. Because physicists were unwilling to take these two effects seriously, physics was held up for quite a number of years.

5.4.7 *Vacuum Polarization for Slowly Varying Weak Fields. The Uehling Effect*

Let now the external potential $A_\nu^e(x)$ be not only weak but also slowly varying in both space and time, i.e. let it be a superposition of Fourier components (360) with

$$|q^2| \ll \mu^2 \tag{404}$$

Then by (390) the second term in (396) is zero[53] and the logarithm may be expanded in terms of (q^2/μ^2). Keeping only the term of order q^2

$$\langle j_{\nu H}(x)\rangle_o = \alpha \frac{q^2}{16\pi\mu^2} j_{\nu R}(x) \int_0^1 \frac{z^2\,dz}{\sqrt{1-z}} = \frac{\alpha q^2}{15\pi\mu^2} j_{\nu R}(x)$$

But in each Fourier component of $j_{\nu R}(x)$, operating with the D'Alembertian $\Box^2$ gives a factor $(-q^2)$. Hence the result, independent of the Fourier decomposition, valid for slowly varying fields

$$\langle j_{\nu H}(x)\rangle_o = -\frac{\alpha}{15\pi\mu^2}\left\{\Box^2 j_{\nu R}(x)\right\} \tag{405}$$

Consider the effect of (405) in the case of the hydrogen atom. The proton[54] is represented by the static charge-density $\rho_P(r)$, and this induces a charge in the vacuum whose density is

$$\rho_{IN}(r) = -\frac{\alpha}{15\pi\mu^2}\nabla^2\rho_P(r) \tag{406}$$

The electrostatic potential of the proton is thus $V(r) + V_{IN}(r)$, where

$$\nabla^2 V(r) = -\rho_P(r)$$
$$\nabla^2 V_{IN}(r) = -\rho_{IN}(r) = \frac{\alpha}{15\pi\mu^2}\nabla^2\rho_P(r)$$

and hence

$$V_{IN}(r) = +\frac{\alpha}{15\pi\mu^2}\rho_P(r) \tag{407}$$

Thus for a point proton, the potential added to the Coulomb potential by vacuum-polarization is

$$V_{\text{IN}}(r) = +\frac{\alpha e}{15\pi\mu^2}\delta^3(\boldsymbol{r}) \tag{408}$$

And the change in the energy of a state of the hydrogen atom with wave-function $\psi(r)$ is

$$\Delta E_P = -\frac{\alpha e^2}{15\pi\mu^2}|\psi(0)|^2 = -\frac{1}{5}\left\{\frac{e^4\hbar}{12\pi^2 m^2 c^3}|\psi(0)|^2\right\} \tag{409}$$

This is just the same as the formula for the Lamb shift, (284), with $(-1/5)$ instead of the logarithm. Thus it is a factor 40 smaller than the Lamb shift, and in the opposite direction, -27 megacycles in the total of 1062. Still the experiments are good enough to show that the effect is there all right.

The result (409) was calculated many years ago by Uehling [17] using old-fashioned methods.

5.5 Field Theory of Dirac and Maxwell Fields in Interaction

5.5.1 *The Complete Relativistic Quantum Electrodynamics*

We now take the combined system of Dirac and Maxwell fields in interaction, and make out of it a relativistic quantum theory using the methods we have already developed. This will then be the complete theory of quantum electrodynamics, applicable to all problems in which electrons, positrons and photons are concerned. We also include in the theory a classical Maxwell field, which acts on the electrons and positrons and represents the effects of external charges such as protons which may happen to be present.

Lagrangian

$$\mathscr{L} = \mathscr{L}_D + \mathscr{L}_M - ie\overline{\psi}\slashed{A}\psi - ie\overline{\psi}\slashed{A}^e\psi \tag{410}$$

Here we use $A_\nu(x)$ for the Maxwell potential operators and $A_\nu^e(x)$ for the potentials of the classical external field.

Field equations

$$\left\{\sum_\lambda \gamma_\lambda\left\{\frac{\partial}{\partial x_\lambda} + \frac{ie}{\hbar c}(A_\lambda + A_\lambda^e)\right\} + \mu\right\}\psi = 0 \tag{411}$$

$$\sum_\lambda \left\{ \frac{\partial}{\partial x_\lambda} - \frac{ie}{\hbar c} \left(A_\lambda + A_\lambda^e \right) \right\} \overline{\psi} \gamma_\lambda - \mu \overline{\psi} = 0 \tag{412}$$

$$\Box^2 A_\nu = ie\overline{\psi}\gamma_\nu\psi \tag{413}$$

(See (384).)[55]

These equations are *non-linear*. And so there is no possibility of finding the general commutation rules of the field operators in closed form. We cannot find any solutions of the field equations, except for the solutions which are obtained as formal power series expansions in the coefficient e which multiplies the non-linear interaction terms. It is thus a basic limitation of the theory, that it is in its nature a perturbation theory stating from the non-interacting fields as an unperturbed system. Even to write down the general commutation laws of the fields, it is necessary to use a perturbation theory of this kind.

Since the perturbation theory treatment is forced on us from the beginning, it is convenient not to set up the theory in the Heisenberg representation but to use the interaction representation. The IR is just designed for a perturbation theory in which the radiation interaction is treated as small. In the IR the commutation rules can be obtained simply in closed form, and so the theory can be set up with a minimum of trouble.

There are two different interaction representations which we can use. The first may be called the *Bound Interaction Representation* (BIR). It is exactly the representation we used in discussing the radiation from an atom in the non-relativistic theory. We take all field operators to have the time-dependence of the Heisenberg operators in the theory of the free Maxwell field and the electron field interacting with external potential, only the interaction between the two fields being omitted. Thus the field equations in the BIR are (332), (333), and

$$\Box^2 A_\nu = 0 \tag{414}$$

The wave-function $\Phi(t)$ in the BIR satisfies the Schrödinger equation

$$i\hbar \frac{\partial \Phi}{\partial t} = H_R(t)\Phi \tag{415}$$

$$H_R(t) = ie \int \overline{\psi}(r,t) \slashed{A}(r,t)\psi(r,t)d^3\boldsymbol{r} \tag{416}$$

This $H_R(t)$ is just the difference between the Hamiltonians of the theories with and without the radiation interaction. Because no derivatives of field operators occur in H_R, the difference is just minus the difference between

the corresponding Lagrangians, and so has the simple form given by (416). (Compare with (243).)

Using the BIR we can discuss the radiation of light by an atom, as it was done before, but now treating the atom relativistically. In fact this representation must be used, as soon as we wish to calculate effects accurately enough to require exact Dirac wave-functions for the unperturbed atomic states. However, the BIR is not convenient to use, because the commutation rules for the electron field are given by (346) and are still too complicated for all but the simplest problems. Therefore we shall use the BIR only when we are compelled to, and then usually only in the final stages of a problem. Generally we can do the main part of the work, with the major calculations, in the second type of interaction representation.

5.5.2 *Free Interaction Representation*

Here we take all field operators $\overline{\psi}$, ψ and A_μ to satisfy the free field equations (286) and (414). The commutation rules are then also given by the free-field formulae (203) and (298), (299). The wave-function satisfies the Schrödinger equation

$$i\hbar\frac{\partial \Phi}{\partial t} = \left\{H^e(t) + H_R(t)\right\}\Phi \tag{417}$$

$$H^e(t) = ie\int \overline{\psi}(\boldsymbol{r},t)\slashed{A}^e(\boldsymbol{r},t)\psi(r,t)d^3\boldsymbol{r} \tag{418}$$

with H_R formally again given by (416). But here H_R is not the same operator as in (415) because of the time-dependence of the $\overline{\psi}$ and ψ in the two cases.

This FIR is the interaction representation normally used in quantum electrodynamics, and will be called henceforth the Interaction Representation or IR simply. It is very well suited to relativistic calculations because it makes the field-commutators and expectation values invariant functions. Thus the calculations can be explicitly and formally invariant, even when the potentials A_ν^e are given in a special Lorentz frame as in the hydrogen atom.

It was Schwinger and Feynman who first discovered the importance of making calculations formally invariant when using a relativistic theory. They made this discovery in completely different ways, characteristically. Feynman simply found that the calculations become much easier and simpler when done in a way which did not conceal the invariance of the theory. This is still true, in fact the main reason why we can tackle now more difficult problems than we could 10 years ago is just that the calculations with the

new methods are so much shorter. But the greater and essential advantage of the covariant calculations, pointed out by Schwinger, is that they enable the separation of finite observable effects from infinite renormalization terms to be made in a clear and unambiguous way. We had an example of this in the treatment we gave of vacuum polarization, where the covariant type of calculation was used. The divergent term (382) could be separated clearly from (379), because of the way in which (379) depends formally on the vectors e_ν and q_ν. If the calculation had been done in a non-covariant way we could not have used (381) as we did.

So now we shall apply the covariant methods, working in the IR, to solve a number of the standard problems of electrodynamics, in increasing order of difficulty.

This page intentionally left blank

CHAPTER 6

Free Particle Scattering Problems

In this extensive class of problems we are interested in calculating the over-all transition matrix element M between an initial state A and a final state B, A and B being specified by assigning the spins and momenta of the free particles which are present in these states. The scattering process is supposed to occur as follows: The free particles which are specified by state A in the remote past, converge and interact, and other free particles emerge or are created in the interaction and finally constitute the state B in the remote future. We wish to calculate the matrix element M for this process, without studying the equations of motion or investigating the behavior of the system at intermediate times while the interaction is in progress.

The unperturbed states A and B are supposed to be states of free particles without interaction and are therefore represented by constant state-vectors Φ_A and Φ_B in the interaction representation. The actual initial and final states in a scattering problem will consist of particles each having a self-field with which it continues to interact even in the remote future and past, hence Φ_A and Φ_B do not accurately represent the initial and final states. However, so long as we are using perturbation theory and not including the higher order effects arising form the self-fields of the particles, it is consistent to use constant Φ_A and Φ_B representing bare particles without radiation interaction. Even when self-field effects are considered, it turns out that the bare-particle Φ_A and Φ_B can still be used, although in this case some careful justification for it is needed.

The matrix element M is

$$M = (\Phi_B^* U \Phi_A) \tag{419}$$

Here $U\Phi_A$ is the state which is obtained at $t = +\infty$ by solving the equation of motion (417) with the initial condition $\Phi = \Phi_A$ at $t = -\infty$. U can easily be written down as a perturbation expansion in the operators H^e and H_R,

$$U = 1 + \left(-\frac{i}{\hbar}\right) \int_{-\infty}^{\infty} dt_1 \left\{H^e(t_1) + H_R(t_1)\right\}$$

$$+ \left(-\frac{i}{\hbar}\right)^2 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \left\{H^e(t_1) + H_R(t_1)\right\}$$

$$\times \left\{H^e(t_2) + H_R(t_2)\right\} + \ldots \tag{420}$$

$$= \sum_{n=0}^{\infty} \left(-\frac{i}{\hbar}\right)^n \frac{1}{n!} \int_{-\infty}^{\infty} dt_1 \ldots \int_{-\infty}^{\infty} dt_n$$

$$\times P\left\{\left\{H^e(t_1) + H_R(t_1)\right\} \ldots \left\{H^e(t_n) + H_R(t_n)\right\}\right\} \tag{421}$$

Here the P signifies a chronological product, the factors in it being multiplied not in the order in which they are written but in the order of the times $t_1, t_2, \ldots, t_n$, the factors with later times standing to the left of those with earlier times. This accounts for the factor $1/n!$ after we change all the limits to cover the whole range from $-\infty$ to $+\infty$. The operator U is generally referred to as "the S-matrix".

Before discussing the general analysis of the series expansion (421), we shall use it to solve some standard problems.

6.1   Møller Scattering of Two Electrons

We have in the initial state A two electrons in states $(p_1u_1)(p_2u_2)$ and in the final state B two electrons $(p'_1u'_1)(p'_2u'_2)$. The electron (p_1u_1) is given by the one-particle wave-function

$$u_1 e^{ip_1 \cdot x} \tag{422}$$

normalized by $(\overline{u}_1u_1) = 1$. With the wave function normalized in this way, (422) is just the matrix element of the operator $\psi(x)$ between the vacuum state and the state containing electron 1. We can see this from

$$\psi(x) = \sum_{p,u} b_{pu} u e^{ip \cdot x}$$

where $\{b_{pu}, b^*_{p'u'}\} = \delta_{pp'}\delta_{uu'}$. Then $(\Phi_o^*, \psi(x)\Phi_{pu}) = (\Phi_o^*, b_{pu}\Phi_{pu})ue^{ip \cdot x} = (\Phi_o^*, \Phi_o)ue^{ip \cdot x} = ue^{ip \cdot x}$.

Thus we consider the states 1, 2 and $1'$, $2'$ as if they were discrete states, the ψ operator being given by the expansion (337). It would also be possible to use the continuous-state expansion (306) for ψ, but then the normalization of the initial and final states would have to be considered afresh. Since we fixed the normalization (472) when we previously derived the Møller formula (144), we shall now stick to it.

We shall calculate in the Born approximation as before, thus keeping only the term $n = 2$ in (421) which gives the matrix element M proportional to e^2. In this problem the external potential A^e is zero. The term $n = 2$ in (421) is

$$U_2 = \frac{+e^2}{2\hbar^2 c^2} \iint dx_1\, dx_2\, P\left\{\overline{\psi}(x_1)A\!\!\!/(x_1)\psi(x_1), \overline{\psi}(x_2)A\!\!\!/(x_2)\psi(x_2)\right\} \quad (423)$$

the integration going over all space-time. To obtain the matrix element $M = (\Phi_B^* U_2 \Phi_A)$ we only have to replace according to (377)

$$\psi(x_i) = u_1 e^{ip_1 \cdot x_i} b_1 + u_2 e^{ip_2 \cdot x_i} b_2 + u' e^{ip_1' \cdot x_i} b_1' + u_2' e^{ip_2' \cdot x_i} b_2' \quad \text{by}$$

$$\times\, u_1 e^{ip_1 \cdot x_i} b_1 + u_2 e^{ip_2 \cdot x_i} b_2 \quad (424)$$

and

$$\overline{\psi}(x_i) = \overline{u}_1 e^{-ip_1 \cdot x_i} b_1^* + \overline{u}_2 e^{-ip_2 \cdot x_i} b_2^* + \overline{u}_1' e^{-ip_1' \cdot x_i} b_1^{*\prime} + \overline{u}_2' e^{-ip_2' \cdot x_i} b_2^{*\prime} \quad \text{by}$$

$$\times\, \overline{u}_1' e^{-ip_1' \cdot x_i} b_1^{*\prime} + \overline{u}_2' e^{-ip_2' \cdot x_i} b_2^{*\prime}$$

because we only absorb 1 and 2 and only create $1'$ and $2'$. Then we want to pick out the coefficient of

$$(b_1^{*\prime} b_1)(b_2^{*\prime} b_2) \quad (425)$$

in the resulting expansion. There are no photons in the initial and final states and so the vacuum expectation value is taken for the Maxwell potential operators. We thus find, taking account of the fact that the b and $\overline{b}$ anticommute with each other as in the derivation of (358), the result

$$M = \sum_{\mu,\lambda} \frac{e^2}{\hbar^2 c^2} \iint dx_1\, dx_2\, \left\{\exp\left[i(p_1 - p_1') \cdot x_1 + i(p_2 - p_2') \cdot x_2\right]\right.$$

$$\times (\overline{u}_1' \gamma_\lambda u_1)(\overline{u}_2' \gamma_\mu u_2) - \exp\left[i(p_1 - p_2' \cdot x_1 + i(p_2 - p_1') \cdot x_2\right]$$

$$\left. \times (\overline{u}_2' \gamma_\lambda u_1)(\overline{u}_1' \gamma_\mu u_2)\right\} \langle P\{A_\lambda(x_1), A_\mu(x_2)\}\rangle_o \quad (426)$$

The expectation-value of the chronological product brings in an important new function,

$$\langle P\{A_\lambda(x_1), A_\mu(x_2)\}\rangle_o = \frac{1}{2}\hbar c\, \delta_{\lambda\mu}\, D_F(x_1 - x_2) \tag{427}$$

where F stands for Feynman. Also called by Stueckelberg D^c, c for causality. [18]

6.1.1   *Properties of the D_F Function*

Since

$$P\{A_\lambda(x_1), A_\mu(x_2)\} = \frac{1}{2}\{A_\lambda(x_1), A_\mu(x_2)\} + \frac{1}{2}\epsilon(x_1 - x_2)[A_\lambda(x_1), A_\mu(x_2)] \tag{428}$$

we have by (203) and (205)

$$D_F(x) = D^{(1)}(x) + i\epsilon(x)D(x) = \frac{1}{2\pi^2}\left[\frac{1}{x^2} - i\pi\delta(x^2)\right]$$

$$= D^{(1)}(x) - i\{D_A(x) + D_R(x)\}$$

$$= D^{(1)}(x) - 2i\overline{D}(x) \tag{429}$$

according to (228). Obviously D_F is an even function. It has the property that asymptotically as $x_0 \to \infty$ in the future, $D_F = 2iD^+$ contains only positive frequencies, while as $x_0 \to -\infty$ in the past, $D_F = -2iD^-$ contains only negative frequencies. See Fierz, *Helv. Phys. Acta* **23** (1950) 731 for a full discussion.

Thus D_F is the potential arising from a point source disturbance at the origin, when all the potential travelling out in the future represents particles created, and all the potential travelling in from[56] the past represents particles to be absorbed, all particles having positive energy. It is thus the potential which maintains the correct causal time-sequence of events, and in this way it was discovered by Stueckelberg. But the definition (427) is easier to understand and use.

The momentum representation of D_F is

$$D_F(x) = \frac{-2i}{(2\pi)^4}\int_F e^{ik\cdot x}\,\frac{d^4k}{k^2} \tag{430}$$

The contour of integration here is along the real axis, below the pole at $k_0 = -|\boldsymbol{k}|$ and above the pole at $k_0 = +|\boldsymbol{k}|$, in the k_0 plane:

We can see this using (429), (207), (208), and comparing (210), (226) and (209) one also knows that

$$D^{(1)}(x) = \frac{1}{(2\pi)^4} \int_C \exp(ik \cdot x) \frac{d^4k}{k^2} \quad \text{with}$$

(430) is always known as a "Feynman integral". We can also write

$$D_F(x) = \frac{-2i}{(2\pi)^4} \int e^{ik \cdot x} \frac{d^4k}{k^2 - i\epsilon} \tag{431}$$

Here the integration is along the real axis for all four components of k. ϵ is a small positive real number and the limit $\epsilon \to 0$ is supposed to be taken in (431) after the integration is done. Before going to the limit, the ϵ term just shifts the poles off the real axis, the pole $|\boldsymbol{k}|$ going done and the pole $-|\boldsymbol{k}|$ going up, and so the integral is well-defined and non-singular.

6.1.2   The Møller Formula, Conclusion

Using (427) and (431) in (426), the integrations over x_1 and x_2 can be done at once, giving a δ-function involving k, and the k integration can then also be done immediately. The result is

$$M_2 = \sum_\lambda \frac{-ie^2}{\hbar c} (2\pi)^4 \delta^4(p_1 + p_2 - p_1' - p_2')$$

$$\times \left\{ \frac{(\overline{u}_1' \gamma_\lambda u_1)(\overline{u}_2' \gamma_\lambda u_2)}{(p_1 - p_1')^2 - i\epsilon} - \frac{(\overline{u}_2' \gamma_\lambda u_1)(\overline{u}_1' \gamma_\lambda u_2)}{(p_1 - p_2')^2 - i\epsilon} \right\} \tag{432}$$

Now p_1 and p_1' are both electron momentum 4-vectors and hence $(p_1 - p_1')$ is a space-like vector and its square cannot be zero. Hence we can go directly to the limit by putting $\epsilon = 0$ in (432). This gives us the Møller formula (144) allowing for a difference in units in p and e.

It is clear that the formula comes very directly out of (423), once we know the momentum expansion (431) of the D_F function. And we shall find that things are equally simple in other free-particle scattering problems.

6.1.3   *Electron-Positron Scattering*

Exactly the same formula (432) also gives the matrix element of scattering
of an electron by a positron. We must assume that the electron is initially
in state 1 given by

$$u_1 e^{ip_1 \cdot x} \tag{433}$$

and finally in state $1'$ by

$$u_1' e^{ip_1' \cdot x} \tag{434}$$

But now the initial state of the positron is given by the wave-function

$$\overline{u}_2' e^{-ip_2' \cdot x} \tag{435}$$

and the final state by

$$\overline{u}_2 e^{-ip_2 \cdot x} \tag{436}$$

using the negative-energy electron wave-functions and *not* the charge-
conjugate functions to represent the positron. The correctness of (435) and
(436) is clear since b_2 is the *emission* operator and b_2' the *absorption* operator
for this positron.

The second term in (432) now represents not an ordinary exchange ef-
fect but a special short-range scattering due to a virtual annihilation of the
positron and electron. This term has been observed experimentally by mea-
suring the fine-structure constant of positronium. (M. Deutsch and E. Dulit,
Phys. Rev. **84** (1951) 601, (Nov. 1, 1951).)

6.2   Scattering of a Photon by an Electron.   The Compton Effect.   Klein-Nishina Formula

Again we use the same operator U_2 given by (423). We only have to calculate
its matrix element M_2 between an initial state A and a final state B, where
A consists of an electron with wave-function

$$u e^{ip \cdot x} \tag{437}$$

and a photon with potentials given by

$$A_\mu = e_\mu e^{ik \cdot x} \tag{438}$$

and B consists of the electron in state

$$u' e^{ip' \cdot x} \tag{439}$$

and the photon with potentials

$$A_\mu = e'_\mu e^{ik' \cdot x} \tag{440}$$

The operator $A_\lambda(x_1)$ appearing in (423) contains both photon emission and absorption components according to (211), and so does $A_\mu(x_2)$. Hence the matrix element M_2 is a sum of contributions; we can either take $e_\lambda e^{ik \cdot x_1}$ from $A_\lambda(x_1)$ and[57] $e'_\mu e^{ik' \cdot x_2}$ from $A_\mu(x_2)$ or vice versa. Likewise the electron can be absorbed by $\psi(x_2)$ or emitted again by $\overline{\psi}(x_1)$, or vice versa. Thus altogether we find for M_2, after taking into account that the whole expression is symmetrical in x_1 and x_2,

$$M_2 = \frac{e^2}{\hbar^2 c^2} \sum_{\lambda,\mu,\alpha,\beta} \iint dx_1 \, dx_2 \, \left\{ \exp(ip \cdot x_2 - ip' \cdot x_1) \left(\overline{u}'_1 \gamma_\lambda \right)_\alpha \right.$$

$$\times \left\langle \epsilon(x_1 - x_2) P \left\{ \psi_\alpha(x_1), \overline{\psi}_\beta(x_2) \right\} \right\rangle_o (\gamma_\mu u)_\beta \Big\}$$

$$\times \left\{ e_\lambda e'_\mu \exp(ik \cdot x_1 - ik' \cdot x_2) + e_\mu e'_\lambda \exp(ik \cdot x_2 - ik' \cdot x_1) \right\} \tag{441}$$

Note that with anticommuting fields the expression $\epsilon(x_1 - x_2) P\{\psi_\alpha(x_1), \overline{\psi}_\beta(x_2)\}$ is a relativistic invariant whereas the P-product by itself is not. Thus in analogy with (427) we write

$$\left\langle \epsilon(x_1 - x_2) P \left\{ \psi_\alpha(x_1), \overline{\psi}_\beta(x_2) \right\} \right\rangle_o = -\frac{1}{2} S_{F\alpha\beta}(x_1 - x_2) \tag{442}$$

where S_F is a new invariant function. Since

$$\epsilon(x_2 - x_1) P \left\{ \psi_\alpha(x_1), \overline{\psi}_\beta(x_2) \right\}$$

$$= \frac{1}{2} [\psi_\alpha(x_1), \overline{\psi}_\beta(x_2)] + \frac{1}{2} \epsilon(x_2 - x_1) \left\{ \psi_\alpha(x_1), \overline{\psi}_\beta(x_2) \right\} \tag{443}$$

we have by (299) and (329)

$$S_F(x) = S^{(1)} + i\epsilon(x)S(x) \tag{444}$$

in exact analogy to (429). We write also

$$S_F(x) = \sum_\lambda \left(\gamma_\lambda \frac{\partial}{\partial x_\lambda} - \mu \right) \Delta_F(x) \tag{445}$$

and we find from (444) the momentum representations

$$\Delta_F(x) = \frac{-2i}{(2\pi)^4} \int e^{ik \cdot x} \frac{d^4 k}{k^2 + \mu^2 - i\epsilon} \tag{446}$$

$$S_F(x) = \frac{2}{(2\pi)^4} \int e^{ik \cdot x} \left(\frac{\not{k} + i\mu}{k^2 + \mu^2 - i\epsilon} \right) d^4k \qquad (447)$$

We can also write (447) conveniently as

$$S_F(x) = \frac{2}{(2\pi)^4} \int_F e^{ik \cdot x} \frac{1}{\not{k} - i\mu} d^4k \qquad (448)$$

Here the Dirac matrix in the denominator means that we have to multiply above and below by $(\not{k} + i\mu)$ in order to evaluate the integral. Thus (448) is not a real simplification of (447), only it saves writing. The Feynman integration in (448) is defined as a contour integral exactly as in (430).

Now substituting from (448) into (441), just like for (432)

$$M_2 = \frac{-e^2}{\hbar^2 c^2} (2\pi)^4 \delta^4(p + k - p' - k') \, \overline{u}' \left[\not{e} \frac{1}{\not{p} - \not{k}' - i\mu} \not{e}' + \not{e}' \frac{1}{\not{p} + \not{k} - i\mu} \not{e} \right] u \qquad (449)$$

Again the denominators $(p - k)^2 + \mu^2$ can never vanish so the ϵ can be put equal to zero in (447). In fact if the electron is initially at rest

$$(p - k')^2 + \mu^2 = (p^2 + \mu^2) + k'^2 - 2p \cdot k' = 2p_0 k'_0 = 2\mu k'_0$$

and similarly

$$(p + k)^2 + \mu^2 = -2\mu k_0 \qquad (450)$$

because $k'^2 = 0$ and $p^2 = -\mu^2$; $\boldsymbol{p} = 0$ because the electron is at rest.

$$M_2 = \frac{-e^2}{2\hbar^2 c^2 \mu} (2\pi)^4 \delta^4(p + k - p' - k')$$

$$\times \overline{u}' \left[\frac{1}{k'_0} \not{e}(\not{p} - \not{k}' + i\mu)\not{e}' - \frac{1}{k_0} \not{e}'(\not{p} + \not{k} + i\mu)\not{e} \right] u \qquad (451)$$

Now we can simplify (451) further. Since the photon is not polarized in time, $e_4 = e'_4 = 0$; since the electron is at rest, $\boldsymbol{p} = 0$. Hence $e \cdot p = 0$, and thus $\not{e}\not{p} = -\not{p}\not{e} + 2\mathbb{I} e \cdot p = -\not{p}\not{e}$, i.e. $\not{p}$ and $\not{e}$ anticommute. This plus the fact that u is a spin state of momentum $\hbar p$, i.e.

$$(\not{p} - i\mu)u = 0 \qquad (452)$$

mean that the term $\not{p} + i\mu$ in (451) can be omitted. Thus we get

$$M_2 = \frac{e^2}{2\hbar^2 c^2 \mu} (2\pi)^4 \delta^4(p + k - p' - k') \, \overline{u}' \left[\not{e} \frac{\not{k}'}{k'_0} \not{e}' + \not{e}' \frac{\not{k}}{k_0} \not{e} \right] u \qquad (453)$$

6.2.1 *Calculation of the Cross-Section*

We write as in (145)

$$M_2 = K(2\pi)^4 \delta^4(p + k - p' - k') \tag{454}$$

Then the scattering probability per unit volume and per unit time for the single final state is as before

$$w_\delta = c|K|^2(2\pi)^4 \delta^4(p + k - p' - k') \tag{455}$$

The number of final states for the electron is

$$\frac{1}{(2\pi)^3} \left(\frac{mc^2}{E'}\right) dp'_1 \, dp'_2 \, dp'_3 \tag{456}$$

The photon with potentials (440), when e'_μ is a space-like vector having $(e'_\mu)^2 = 1$ is normalized to one particle per volume $\hbar c/2k'_0$. This can be seen at once, comparing (440) with (211) and (214), and taking into account the difference of $(2\pi)^3$ between the continuous and discrete normalizations. Hence the number of final states for the photon is

$$\frac{1}{(2\pi)^3} \left(\frac{\hbar c}{2k'_0}\right) dk'_1 \, dk'_2 \, dk'_3 \tag{457}$$

The total transition probability is thus

$$w = c|K|^2 \frac{1}{(2\pi)^2} \left(\frac{mc^2}{E'} \frac{\hbar c}{2k'_0}\right) \frac{dk'_1 \, dk'_2 \, dk'_3}{d(p'_0 + k'_0)} \tag{458}$$

We write this as a probability for scattering the *photon* with frequency k'_0 into a unit solid angle $d\Omega$. Then using momentum conservation, we have

$$\frac{dp'_0}{dk'_0} = \frac{p'_1 \, dp'_1 + \cdots}{p'_0 \, dk'_0} = -\frac{p'_1 \, dk'_1 + \cdots}{p'_0 \, dk'_0} = -\frac{p'_1 k'_1 + \cdots + p'_3 k'_3}{p'_0 k'_0}$$

$$\frac{dk'_0}{d(p'_0 + k'_0)} = \frac{p'_0 k'_0}{-(p' \cdot k')} = \frac{p'_0 k'_0}{-p \cdot k} = \frac{p'_0 k'_0}{p_0 k_0} \tag{459}$$

Hence finally

$$w = c|K|^2 \frac{1}{(2\pi)^2} \frac{\hbar c}{2k_0} k'^2_0 \, d\Omega \tag{460}$$

The differential cross-section for scattering the photon of frequency k_0 into solid angle $d\Omega$ is then

$$\sigma = \frac{wV_1V_2}{c}$$

where V_1 is according to (149) the electron normalization volume $\frac{mc^2}{E} = 1$
and V_2 is the photon volume, $V_2 = \frac{\hbar c}{2k_0}$. Thus

$$\sigma = \left(\frac{\hbar c}{4\pi k_0} \right)^2 |K|^2 k_0'^2 \, d\Omega \tag{461}$$

$$K = \frac{e^2}{2\hbar^2 c^2 \mu} \, \overline{u}' \left[\phi \frac{\slashed{k}'}{k_0'} \phi' + \phi' \frac{\slashed{k}}{k_0} \phi \right] u \tag{462}$$

This gives the cross-section for a known electron spin in the initial and final
states.

6.2.2   *Sum Over Spins*

Experimentally we do not observe the electron spins. Hence we observe only
the cross-section $\overline{\sigma}$ obtained by averaging σ over the two spin states u and
summing over the two spin states u'. The summing and averaging we do by
the method of projection operators according to (109), (114).

Here are some rules for trace and dagger operators:

ξ is a Dirac matrix in general.

(1) Tr $\left(\xi^{(1)} \xi^{(2)} \dots \xi^{(2k-1)} \right) = 0$, i.e. the trace of an odd number of factors is
0.
(2) Tr $\left(\xi^{(1)} \xi^{(2)} \dots \xi^{(2k)} \right) = $ Tr $\left\{ P \left(\xi^{(1)} \xi^{(2)} \dots \xi^{(2k)} \right) \right\}$ where P is any cyclic
permutation.
This is clear, because any cyclic permutation consists of steps of the form

$$\left(\xi^{(1)} \xi^{(2)} \dots \xi^{(s)} \right) \xi^{(m)} \rightarrow \xi^{(m)} \left(\xi^{(1)} \xi^{(2)} \dots \xi^{(s)} \right)$$

and for two square matrices[58] A, B

$$\text{Tr } AB = \sum_{i,j} a_{ij} b_{ji} = \sum_{j,i} b_{ji} a_{ij} = \text{Tr } BA$$

(3) Tr $\left(\xi^{(1)} \xi^{(2)} \dots \xi^{(2k-1)} \xi^{(2k)} \right) = $ Tr $\left(\xi^{(2k)} \xi^{(2k-1)} \dots \xi^{(2)} \xi^{(1)} \right)$
To show this, it is enough to assume that all $\xi^{(i)}$ are different; by the
commutation rules of the γ's we can always reduce the product to this
form. Then, since each inversion (of neighbors) brings in a minus sign,
and since there are an even number of inversions, we get our rule imme-
diately.

(4) $\displaystyle a\!\!\!/\,b\!\!\!/ = -b\!\!\!/\,a\!\!\!/ + 2\,\mathbb{I}\,(a\cdot b)$

In particular

$$\phi\!\!\!/\,\phi\!\!\!/ = \mathbb{I}\,(e\cdot e) \qquad \phi\!\!\!/\,k\!\!\!/ = -k\!\!\!/\,\phi\!\!\!/ \qquad \phi\!\!\!/'\,k\!\!\!/' = -k\!\!\!/'\,\phi\!\!\!/'$$

(5) By (2), one can cyclically permute a product of dagger operators without changing its trace.

Now we can proceed to evaluate the sum over spins. We have

$$\frac{1}{2}\sum_u\sum_{u'}|K|^2 = -\frac{e^4}{8\hbar^4 c^4\mu^2}\sum_u\sum_{u'}\left\{\bar u'\left(\phi\!\!\!/\,\frac{k\!\!\!/'}{k_0'}\phi\!\!\!/' + \phi\!\!\!/'\,\frac{k\!\!\!/}{k_0}\phi\!\!\!/\right)u\right\}$$

$$\times\left\{\bar u\left(\phi\!\!\!/\,\frac{k\!\!\!/}{k_0}\phi\!\!\!/' + \phi\!\!\!/'\,\frac{k\!\!\!/'}{k_0'}\phi\!\!\!/\right)u'\right\}$$

$$= \frac{e^4}{32\hbar^4 c^4\mu^4}\,\mathrm{Tr}\left\{(p\!\!\!/+i\mu)\left(\phi\!\!\!/\,\frac{k\!\!\!/}{k_0}\phi\!\!\!/' + \phi\!\!\!/'\,\frac{k\!\!\!/'}{k_0'}\phi\!\!\!/\right)\right.$$

$$\left.\times(p\!\!\!/'+i\mu)\left(\phi\!\!\!/\,\frac{k\!\!\!/'}{k_0'}\phi\!\!\!/' + \phi\!\!\!/'\,\frac{k\!\!\!/}{k_0}\phi\!\!\!/\right)\right\} \tag{463}$$

Now $\dfrac{k\!\!\!/}{k_0} = i\beta + \dfrac{k_1\gamma_1 + k_2\gamma_2 + k_3\gamma_3}{k_0} = i\beta + \gamma_k$, say.

Similarly $\dfrac{k\!\!\!/'}{k_0'} = i\beta + \gamma_{k'}$.

Since $p\!\!\!/$ anticommutes with $\phi\!\!\!/$, $\phi\!\!\!/'$, γ_k and $\gamma_{k'}$, (compare with remark after (451)) we may write (463) in the form

$$\frac{e^4}{32\hbar^4 c^4\mu^4}\,\mathrm{Tr}\left[\left\{\left(\phi\!\!\!/\,\frac{k\!\!\!/}{k_0}\phi\!\!\!/' + \phi\!\!\!/'\,\frac{k\!\!\!/'}{k_0'}\phi\!\!\!/\right)(i\mu - p\!\!\!/) + 4\mu(e\cdot e')\right\}\right.$$

$$\left.\times(p\!\!\!/'+i\mu)\left\{\phi\!\!\!/\,\frac{k\!\!\!/'}{k_0'}\phi\!\!\!/' + \phi\!\!\!/'\,\frac{k\!\!\!/}{k_0}\phi\!\!\!/\right\}\right]$$

because, using 4), and again compare with remark after (451),[59]

$$p\!\!\!/\,\phi\!\!\!/\,\frac{k\!\!\!/}{k_0}\phi\!\!\!/' = -\phi\!\!\!/\,p\!\!\!/\,\frac{k\!\!\!/}{k_0}\phi\!\!\!/' = +\phi\!\!\!/\,\frac{k\!\!\!/}{k_0}p\!\!\!/\,\phi\!\!\!/' - 2\phi\!\!\!/\,\frac{k\cdot p}{k_0}\phi\!\!\!/' = -\phi\!\!\!/\,\frac{k\!\!\!/}{k_0}\phi\!\!\!/'\,p\!\!\!/ + 2\mu\phi\!\!\!/\,\phi\!\!\!/'$$

and similarly[60]

$$p\!\!\!/\,\phi\!\!\!/'\,\frac{k\!\!\!/'}{k_0'}\phi\!\!\!/ = -\phi\!\!\!/'\,\frac{k\!\!\!/'}{k_0'}\phi\!\!\!/\,p\!\!\!/ + 2\mu\phi\!\!\!/'\,\phi\!\!\!/$$

so adding the terms gives

$$-\left(\not{e}\frac{\not{k}}{k_0}\not{e}' + \not{e}'\frac{\not{k}'}{k_0'}\not{e}\right)\not{p} + 2\mu\{\not{e},\not{e}'\} = -\left(\not{e}\frac{\not{k}}{k_0}\not{e}' + \not{e}'\frac{\not{k}'}{k_0'}\not{e}\right)\not{p} + 4\mu(e\cdot e')$$

Now since $\not{k}\not{k} = \not{k}'\not{k}' = \not{p}\not{p}+\mu^2 = 0$ (for the photons because they are on the lightcone, for the electron because $p^2 = -\mu^2$), we get for $|K|^2$

$$\frac{e^4}{32\hbar^4 c^4 \mu^4}\,\mathrm{Tr}\left[4\mu(e\cdot e')(\not{p}'+i\mu)(\not{e}\frac{\not{k}'}{k_0'}\not{e}' + \not{e}'\frac{\not{k}}{k_0}\not{e}) + (\not{e}\frac{\not{k}'}{k_0'}\not{e}'\not{e}\frac{\not{k}}{k_0}\not{e}'\right.$$

$$\left. + \not{e}'\frac{\not{k}}{k_0}\not{e}\not{e}'\frac{\not{k}'}{k_0'}\not{e})(i\mu - \not{p})(\not{k} - \not{k}')\right] \tag{464}$$

because

$$(i\mu - \not{p})(\not{p}' + i\mu) = i\mu(\not{p}' - \not{p}) - \mu^2 - \not{p}\not{p}' = i\mu(\not{k} - \not{k}') - \mu^2 - \not{p}(\not{p} - \not{k}' + \not{k})$$

$$= (i\mu - \not{p})(\not{k} - \not{k}')$$

We consider the second part of (464) first:

$$\mathrm{Tr}\left[\not{e}\frac{\not{k}'}{k_0'}\not{e}'\not{e}\frac{\not{k}}{k_0}\not{e}'\,\{-\not{p}(\not{k} - \not{k}') - (\not{k} - \not{k}')\not{p}\}\right]$$

$$= 2p_0(k_0 - k_0')\,\mathrm{Tr}\left[\not{e}\not{e}'\frac{\not{k}'}{k_0'}\frac{\not{k}}{k_0}\not{e}\not{e}'\right]$$

$$= 2\mu(k_0 - k_0')\,\mathrm{Tr}\left[-2(e\cdot e')\left(\frac{\not{k}'}{k_0'}\not{e}'\frac{\not{k}}{k_0}\not{e}\right) - \frac{\not{k}'}{k_0'}\frac{\not{k}}{k_0}\right]$$

$$= -8\mu\frac{k_0 - k_0'}{k_0 k_0'}(k\cdot k') + 4\mu(e\cdot e')\,\mathrm{Tr}\left[-\not{k}\not{e}\frac{\not{k}'}{k_0'}\not{e}' + \not{k}'\not{e}'\frac{\not{k}}{k_0}\not{e}\right]$$

since

$$\mathrm{Tr}\left[\not{e}\not{e}'\not{k}'\not{k}\not{e}\not{e}'\right] = \mathrm{Tr}\left[\not{e}\not{e}'\not{e}\not{e}'\not{k}'\not{k}\right] = \mathrm{Tr}\left[-\not{e}\not{e}\not{e}'\not{e}'\not{k}'\not{k} + 2\,\mathbb{I}\,\not{e}\not{e}'\not{k}'\not{k}(e\cdot e')\right]$$

$$= \mathrm{Tr}\left[-\not{k}'\not{k} - 2(e\cdot e')\not{k}'\not{e}'\not{k}\not{e}\right]$$

Hence altogether (464) becomes

$$\frac{e^4}{32\hbar^4 c^4 \mu^4} \left\{ -8\mu \frac{k_0 - k_0'}{k_0 k_0'} (k \cdot k') + 4\mu(e \cdot e') \right.$$

$$\left. \times \text{Tr} \left[(\not{p} - \not{k}') \not{e} \frac{\not{k}'}{k_0'} \not{e}' + (\not{p} + \not{k}) \not{e}' \frac{\not{k}}{k_0} \not{e} \right] \right\}$$

But $\not{k}\not{e}\not{k} = \not{k}'\not{e}'\not{k}' = 0$ and

$$(p' - p)^2 = (k - k')^2 = k^2 + k'^2 - 2k \cdot k' = -2k \cdot k'$$

$$(p' - p)^2 = p'^2 + p^2 - 2p' \cdot p = 2\mu^2 + 2\mu p_0' = 2\mu^2 + 2\mu(-\mu + k_0 - k_0') = 2\mu(k_0 - k_0')$$

Hence $k \cdot k' = -\mu(k_0 - k_0')$ and then

$$\text{Tr} \left[(\not{p} - \not{k}') \not{e} \frac{\not{k}'}{k_0'} \not{e}' + (\not{p} + \not{k}) \not{e}' \frac{\not{k}}{k_0} \not{e} \right]$$

$$= \text{Tr} \left[\not{p}\not{e} \frac{\not{k}'}{k_0'} \not{e}' - \not{k}'\not{e}' \frac{\not{k}'}{k_0'} \not{e} + \not{p}\not{e}' \frac{\not{k}}{k_0} \not{e} + \not{k}\not{e} \frac{\not{k}}{k_0} \not{e}' \right]$$

$$= \text{Tr} \left[\not{p}\not{e} \frac{\not{k}'}{k_0'} \not{e}' + \not{e} \frac{\not{k}}{k_0} \not{e}' \not{p} \right] = \text{Tr} \left[\not{p}\not{e} \frac{\not{k}'}{k_0'} \not{e}' + \not{p}\not{e} \frac{\not{k}}{k_0} \not{e}' \right]$$

and so (464) is

$$\frac{e^4}{8\hbar^4 c^4 \mu^4} \left\{ \frac{2\mu^2 (k_0 - k_0')^2}{k_0 k_0'} + 2\mu(e \cdot e') \, \text{Tr} \left[\not{p}\not{e} i\beta\not{e}' \right] \right\}$$

$$= \frac{e^4}{4\hbar^4 c^4 \mu^2} \left\{ \frac{(k_0 - k_0')^2}{k_0 k_0'} + 4(e \cdot e')^2 \right\} \tag{465}$$

Hence by (461) the cross-section averaged over electron spins is

$$\overline{\sigma} = \frac{e^4 k_0'^2 \, d\Omega}{64\pi^2 \hbar^2 c^2 \mu^2 k_0^2} \left\{ \frac{(k_0 - k_0')^2}{k_0 k_0'} + 4(e \cdot e')^2 \right\}$$

The classical electron radius is

$$r_o = \frac{e^2}{4\pi mc^2} = \frac{e^2}{4\pi \hbar c \mu}$$

Hence

$$\overline{\sigma} = \frac{1}{4} r_o^2 \, d\Omega \left(\frac{k_0'}{k_0}\right)^2 \left\{ \frac{(k_0 - k_0')^2}{k_0 k_0'} + 4\cos^2\phi \right\} \tag{466}$$

where ϕ is the angle between the polarizations of the incident quantum k_0 and the emitted quantum k_0'.

This is the famous *Klein-Nishina formula*.

To put $\overline{\sigma}$ explicitly as a function of the scattering angle θ, we must use the equations[61]

$$k \cdot k' = -\mu(k_0 - k_0')$$

$$k \cdot k' = |\boldsymbol{k}||\boldsymbol{k}'|\cos\theta - k_0 k_0' = k_0 k_0'(\cos\theta - 1)$$

$$k_0 k_0'(1 - \cos\theta) = \mu(k_0 - k_0')$$

$$\frac{k_0}{k_0'} = 1 + (1 - \cos\theta)\frac{k_0}{\mu}$$

Put

$$\epsilon = \frac{k_0}{\mu} = \left(\frac{\text{Photon energy}}{mc^2}\right)$$

Then

$$\overline{\sigma} = \frac{1}{4} r_o^2 \, d\Omega \, \frac{\left(\frac{(1-\cos\theta)^2 \epsilon^2}{1 + \epsilon(1-\cos\theta)} + 4\cos^2\phi\right)}{[1 + \epsilon(1 - \cos\theta)]^2} \tag{467}$$

Thus for large ϵ the scattered photons are mainly unpolarized and concentrated in the forward direction.

For small ϵ (non-relativistic problem) we have simply[62]

$$\overline{\sigma} = r_o^2 \cos^2\phi \, d\Omega \tag{468}$$

the classical result. Summing over the two polarizations of the photon k' and averaging over all polarizations of k, this gives the cross-section for all polarizations

$$\overline{\overline{\sigma}} = \frac{1}{2} r_o^2 (1 + \cos^2\theta) \, d\Omega \tag{469}$$

We get this by evaluating $\frac{1}{2}\sum_e \sum_{e'}(e \cdot e')^2$. First we have to sum over the two polarization directions of photon k'. This summation for three directions would give

$$\sum(e \cdot e')^2 = e^2 = 1$$

Hence for the two directions perpendicular to k' we can write

$$\sum_{e'}(e \cdot e')^2 = 1 - (e \cdot \hat{k}')^2$$

Now we perform the other summation over the two polarization directions of photon k, using the same argument. This gives

$$\sum_{e}\sum_{e'}(e \cdot e')^2 = \sum_{e}\left[1 - (e \cdot \hat{k}')^2\right] = 2 - \left[\hat{k}'^2 - (\hat{k} \cdot \hat{k}')^2\right] = 1 + \cos^2\theta$$

This with the averaging factor of $\frac{1}{2}$ gives (469).

The total cross-section then is

$$\sigma = \frac{8}{3}\pi r_o^2 \tag{470}$$

This non-relativistic scattering given by (468) – (470) is called *Thomson scattering*.[63]

6.3 Two Quantum Pair Annihilation

Consider a process by which an electron in the state (p, u) and a positron associated with the wave-function (439) are both annihilated, with the emission of two photons given by the potentials (438) and (440). The positron momentum-energy 4-vector is then $(-\hbar p')$ so we write $p_+ = -p'$. The positron spinor in the charge-conjugate representation is $v = Cu'^+$.

This annihilation process will again be effected by the operator U_2 given by (423). And the matrix element for the transition is exactly as before given by an expression identical with (449) except that k is now replaced by $-k$, namely

$$M_2 = -\frac{e^2(2\pi)^4}{\hbar^2 c^2}\delta^4(p + p_+ - k - k')\,\overline{u}'\left\{\not{e}\frac{1}{\not{p} - \not{k}' - i\mu}\not{e}' + \not{e}'\frac{1}{\not{p} - \not{k} - i\mu}\not{e}\right\}u$$

$$= K(2\pi)^4\delta^4(p + p_+ - k - k') \tag{471}$$

We consider the probability for this process for an electron and positron *both at rest*. The result will then apply to the decay of a positronium atom, where the velocities are only of the order of αc and can be treated as zero to a good approximation.

Then

$$p = p_+ = (0, 0, 0, i\mu)$$
$$k_0 = k_0' = \mu \tag{472}$$

As in (453) we have

$$K = \frac{e^2}{2\hbar^2 c^2 \mu^2} \overline{u}'(\phi k' \phi' + \phi' k \phi)u \tag{473}$$

The decay probability per unit volume per unit time into a solid angle $d\Omega$ for *one* of the photons is

$$w = c|K|^2 \frac{1}{(2\pi)^2} \left(\frac{\hbar c}{2\mu}\right)^2 \frac{dk_1 dk_2 dk_3}{d(k_0 + k_0')}$$

(because $k_0 = \mu$ here) in analogy to (458). But now $d(k_0 + k_0') = 2dk_0'$ and so

$$w = c|K|^2 \frac{1}{(2\pi)^2} \frac{1}{8} \hbar^2 c^2 \, d\Omega \tag{474}$$

For *parallel* polarizations, $e = e'$, and

$$(\phi k' \phi' + \phi' k \phi) = -(k' + k) = -2i\mu\beta$$

But β has zero matrix element between the spin-states u and u', which are positive and negative-frequency states both of zero momentum. Hence for parallel polarizations

$$w = 0 \tag{475}$$

For *perpendicular* polarizations, take coordinate axes 1 along e, 2 along e', and 3 along k. Then

$$(\phi k' \phi' + \phi' k \phi) = \mu\{\gamma_1(-\gamma_3 + i\beta)\gamma_2 + \gamma_2(\gamma_3 + i\beta)\gamma_1\} = 2\mu\gamma_1\gamma_2\gamma_3 \tag{476}$$

Hence for perpendicular polarizations[64]

$$\overline{u}[\phi k' \phi' + \phi' k \phi]u = 2\mu v^T C \gamma_4 \gamma_1 \gamma_2 \gamma_3 u = -2\mu i v^T \gamma_4 \sigma_2 u$$

$$= 2\mu v^T \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} u$$

$$= \begin{cases} 0 & \text{when spins } u \text{ and } v \text{ are parallel} \\ 2\mu\sqrt{2} & \text{when spins } u \text{ and } v \text{ are antiparallel} \end{cases}$$

We get this latter result by observing that for antiparallel spins the initial wave function is

$$\psi = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

(neglecting the "small components") and therefore

$$v^T \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} u = \frac{1}{\sqrt{2}} (0\ 1) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} (1\ 0) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$= \frac{1}{\sqrt{2}} \left[(1\ 0) \begin{pmatrix} 1 \\ 0 \end{pmatrix} - (0\ -1) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] = \frac{2}{\sqrt{2}} = \sqrt{2}$$

This is one place where charge-conjugate spinors are useful and necessary!

Summarizing, we find for electron and positron with spins parallel, in triplet state, the 2-photon decay is forbidden. This selection rule is in fact an exact one for positronium in the ground state $1s$ triplet. Only 3-photon decay can occur and this makes the lifetime ~ 1100 times longer. For electron and positron in singlet state, the 2-photon decay occurs with the photons always polarized perpendicularly to each other. The probability for the decay, integrating (474) over a solid angle 2π since the photons are indistinguishable, is

$$w = \frac{\hbar^2 c^3}{16\pi} 2|K|^2 = \frac{2e^4}{\hbar^2 c \mu^2 8\pi} = 4\pi c r_o^2 \qquad (477)$$

Formula (477) is for electron and positron normalized to one particle per unit volume. If the density of the electron probability relative to the position of the positron is ρ, then the mean annihilation life-time will be[65] (the "classical electron radius" $r_o = e^2/(4\pi mc^2)$ in Heaviside units)

$$\tau = \frac{1}{4\pi c r_o^2 \rho} \qquad (478)$$

For the positronium $1s$ singlet state

$$\rho = \frac{1}{8\pi a_o^3} \qquad a_o = \text{Bohr radius} = 137^2 r_o$$

$$\tau = 2 \times 137^4 \times \frac{a_o}{c} = 2 \times 137^5 \times \frac{\hbar}{mc^2} \approx 1.2 \times 10^{-10} \text{ sec.} \qquad (479)$$

For slowly-moving electrons and positrons with relative velocity v, the annihilation cross-section according to (477) will be

$$4\pi r_o^2 \left(\frac{c}{v}\right) \qquad \text{singlet state,} \qquad (480)$$

proportional to $\frac{1}{v}$ just like neutron cross-sections at low (thermal) energies.

6.4   Bremsstrahlung and Pair Creation in the Coulomb Field of an Atom

We consider these two important processes together. Given an external potential A_μ^e representing the Coulomb field, the processes are:

Bremsstrahlung
    Electron $(pu) \to$ Electron $(p'u')$ + Photon $(k'e')$

Pair-creation
    Photon $(k'e') \to$ Electron (pu) + Positron $(p'_+ u')$

We treat not only the photon (ke) but also the potential A^e in *Born approximation.* This is valid so long as

Potential energy $\times$ time of transit $\ll \hbar$

$$\text{or} \quad \frac{Ze^2}{4\pi r} \times \frac{r}{v} \ll \hbar$$

$$\text{or} \quad \frac{Ze^2}{4\pi \hbar v} = \frac{Z}{137}\frac{c}{v} \ll 1 \qquad (481)$$

The treatment will only be good for *relativistic velocities* $v \sim c$, and for *lighter atoms* $Z \ll 137$. In fact for heavy atoms ($Z = 82$ for lead) and $v \sim c$ the error from the Born approximation is about 10%.

The processes arise in the Born approximation just from the term linear in A_μ and linear in A_μ^e in (421). This term is

$$U_1 = \frac{e^2}{\hbar^2 c^2} \iint dx_1\, dx_2\, P\left\{\overline{\psi}(x_1)\slashed{A}(x_1)\psi(x_1), \overline{\psi}(x_2)\slashed{A}^e(x_2)\psi(x_2)\right\} \qquad (482)$$

The factor $\frac{1}{2}$ in (423) is now missing, otherwise everything is the same as before. We suppose $A_\mu^e(x_2)$ is a superposition of Fourier components

$$A_\mu^e(x_2) = \frac{1}{(2\pi)^4} \int dk\, f(k)\, e_\mu e^{ik\cdot x_2} \qquad (483)$$

where $f(k)$ is a known function of k. For a static Coulomb field all the vectors k appearing in (483) have zero 4th component, and e_μ is the constant vector

$(0, 0, 0, i)$. We calculate the matrix element M_1 for bremsstrahlung or pair creation with A_μ^e given by the pure Fourier component (438) ; the results are then to be afterwards superposed to give the actual potential according to (483).

For bremsstrahlung the formula for M_1 is (449), identically the same as for the Compton effect, or integrating over the frequency k

$$M_1 = -\frac{e^2}{\hbar^2 c^2} f(p' + k' - p) \, \overline{u}' \left\{ \not{e} \frac{1}{\not{p} - \not{k}' - i\mu} \not{e}' + \not{e}' \frac{1}{\not{p}' + \not{k}' - i\mu} \not{e} \right\} u \quad (484)$$

The factor 2 difference between (482) and (423) is just compensated by the fact that the photon $k'e'$ can be emitted by two operators A_μ in (423) and by only one in (482). The bremsstrahlung cross section is then calculated by squaring (484) and integrating over k' and p' with appropriate normalization factors. For the details see Heitler's book §17.[66]

For pair creation the same formula (449) gives the matrix element M_1, allowing for the fact that the roles of the particles are now changed around so that electron (pu) instead of $(p'u')$ is created etc. Thus

$$M_1 = -\frac{e^2 (2\pi)^4}{\hbar^2 c^2} \delta^4(k + k' - p - p_+)$$

$$\times \overline{u} \left\{ \not{e} \frac{1}{\not{k}' - \not{p}_+ - i\mu} \not{e}' + \not{e}' \frac{1}{\not{k} - \not{p}_+ - i\mu} \not{e} \right\} u' \quad (485)$$

and integrating over the components of the potential

$$M_1 = -\frac{e^2}{\hbar^2 c^2} f(p + p_+ - k') \, \overline{u} \left\{ \not{e} \frac{1}{\not{k}' - \not{p}_+ - i\mu} \not{e}' + \not{e}' \frac{1}{\not{p} - \not{k}' - i\mu} \not{e} \right\} u' \quad (486)$$

For the cross-section calculation see again Heitler, §20.[67]

This page intentionally left blank

General Theory of Free Particle Scattering

We have shown how (421) leads to matrix elements for standard scattering processes, from which cross-sections can be calculated. In each case we used only the term $n = 2$ in (421), which happened to be the lowest term giving a contribution to these processes. The higher terms $n = 4, 6, \ldots$ will also give contributions to the matrix elements for these processes, contributions which are collectively called "radiative corrections". It turns out that the results without radiative corrections agree with the experimental scattering cross-sections in all cases. The experiments are never accurate to better than a few percent, and the radiative corrections are always smaller than the lowest-order terms by at least one power of $(e^2/4\pi\hbar c) = (1/137)$. Thus the study of the radiative corrections for scattering processes will not lead to any directly observable effects.

Nevertheless we shall work out a method of calculating the high-order terms of (421). This method turns out to be simplest and easiest to explain when we are discussing scattering problems. Incidentally we shall see what the radiative corrections to scattering look like, and we shall learn something about the nature of radiative corrections in general. Finally at the end we shall be able to use the method of calculation in order to find the radiative corrections to the motion of an electron in a hydrogen atom, which is the case in which these small effects can be accurately observed, but where the pure scattering theory is not directly applicable.

To avoid unnecessary complications we suppose there is no external field A^e. Problems in which there is an external field, so long as it can be treated in the Born approximation, can always be simply related to problems without

external field, just as the bremsstrahlung matrix element (484) is related to the Compton effect (449). When there is no external field the matrix element for any scattering process is

$$M = (\Phi_B^* S \Phi_A) \qquad (487)$$

$$S = \sum_{n=0}^{\infty} \left(\frac{e}{\hbar c}\right)^n \frac{1}{n!} \int \ldots \int dx_1 \ldots dx_n P\left\{\overline{\psi}A\psi(x_1), \ldots, \overline{\psi}A\psi(x_n)\right\} \qquad (488)$$

The operators in (488) are field-operators of the interaction representation, the integrations over the points $x_1, \ldots x_n$ extend over all space-time, and the initial and final states A and B are entirely arbitrary.

We wish to calculate the matrix element M of S for a particular scattering process, in which the states A and B are specified by enumerating the particles present in the two states. We must now take properly into account the fact that the particles in states A and B, although well separated and no longer interacting with one another, are real particles interacting with their self-fields and with the vacuum-fluctuations of the fields in their neighborhood. Thus A and B are really time-dependent states in the interaction representation and will not be given by time-independent vectors Φ_A and Φ_B, except in the lowest-order approximation. (See page 81.) Let $\Psi_B(t)$ be the actual time-dependent state-vector of the state B in the IR. We are not interested in the dependence of $\Phi_B(t)$ on t. In an actual scattering experiment the particles in state B are observed in counters or photographic plates or cloud-chambers and the time of their arrival is not measured precisely. Therefore it is convenient to use for B not the state-function $\Psi_B(t)$ but a state-function Φ_B which is *by definition* the state-function describing a set of bare particles without radiation interaction, the bare particles having the same momenta and spins as the real particles in the state B. In the IR the state-function Φ_B is time-independent. The question is only, what is the connection between $\Psi_B(t)$ and Φ_B?

Suppose t_B to be a time so long in the future after the scattering process is over, that from t_B to $+\infty$ the state B consists of separated outward-traveling particles. Then the relation between $\Psi_B(t)$ and Φ_B is simple. We imagine a fictitious world in which the charge e occurring in the radiation interaction decreases infinitely slowly (adiabatically) from its actual value at time t_B to zero at time $(+\infty)$. In the fictitious world, the state $\Psi_B(t_B)$ at time t_B will grow into the bare-particle state Φ_B at time $+\infty$. Thus

$$\Phi_B = \Omega_2(t_B)\Psi_B(t_B) \qquad (489)$$

where[68]

$$\Omega_2(t_B) = \sum_{n=0}^{\infty} \left(\frac{e}{\hbar c}\right)^n \frac{1}{n!} \int_{t_B}^{\infty} \dots \int_{t_B}^{\infty} dx_1 \dots dx_n$$

$$\times P\left\{\overline{\psi} A \psi(x_1), \dots, \overline{\psi} A \psi(x_n)\right\} g_B(t_1) \dots g_B(t_n) \qquad (490)$$

and $g_B(t)$ is a function decreasing adiabatically from the value 1 at $t = t_B$ to zero at $t = \infty$. Similarly, when t_A is a time so far in the past that the state A consists of separated converging particles from $t = -\infty$ to $t = t_A$ we have

$$\Psi_A(t_A) = \Omega_1(t_A)\Phi_A \qquad (491)$$

$$\Omega_1(t_A) = \sum_{n=0}^{\infty} \left(\frac{e}{\hbar c}\right)^n \frac{1}{n!} \int_{-\infty}^{t_A} \dots \int_{-\infty}^{t_A} dx_1 \dots dx_n$$

$$\times P\left\{\overline{\psi} A \psi(x_1), \dots, \overline{\psi} A \psi(x_n)\right\} g_A(t_1) \dots g_A(t_n) \qquad (492)$$

where $g_A(t)$ is a function increasing adiabatically from $t = -\infty$ to $t = t_A$.

The scattering matrix element between states A and B is given exactly by

$$M = \left(\Psi_B^*(t_B) S_{t_A}^{t_B} \Psi_A(t_A)\right) \qquad (493)$$

$$S_{t_A}^{t_B} = \sum_{n=0}^{\infty} \left(\frac{e}{\hbar c}\right)^n \frac{1}{n!} \int_{t_A}^{t_B} \dots \int_{t_A}^{t_B} P\left\{\overline{\psi} A \psi(x_1), \dots, \overline{\psi} A \psi(x_n)\right\} dx_1 \dots dx_n$$

$$(494)$$

Of course (493) is independent of the times t_A and t_B. When t_A and t_B are chosen so far in the past that (489) and (491) are satisfied, then (493) may be written in the form (487), where now

$$S = \Omega_2(t_B) S_{t_A}^{t_B} \Omega_1(t_A)$$

$$= \sum_{n=0}^{\infty} \left(\frac{e}{\hbar c}\right)^n \frac{1}{n!} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} dx_1 \dots dx_n$$

$$\times P\left\{\overline{\psi} A \psi(x_1), \dots, \overline{\psi} A \psi(x_n)\right\} g(t_1) \dots g(t_n) \qquad (495)$$

and $g(t)$ is a function increasing adiabatically from 0 to 1 for $-\infty < t < t_A$, equal to 1 for $t_A \leq t \leq t_B$, and decreasing adiabatically from 1 to 0 for $t_A < t < \infty$. Thus we come to the important conclusion that formula (487)

for the matrix element is correct, using the bare particle state-functions Φ_A and Φ_B, provided that formula (488) for S is interpreted by putting in the slowly varying cut-off function $g(t_i)$ to make the integrals converge at $t_i = \pm\infty$. The cut-off functions are to be put in as they appear in (495), and then S is defined as the limit to which (495) tends as the rate of variation of $g(t)$ is made infinitely slow.

The main practical effect of this limiting process in the definition of S is to justify us in throwing away all terms in the integrals which oscillate finitely at $t_i = \pm\infty$. There are however certain cases in which the integral (488) is in a more serious way ambiguous due to bad convergence at $t_i = \pm\infty$. In these cases the cut-off functions have to be kept explicitly until a late stage of the calculations before going to the limit $g(t) = 1$. In all cases, if the limiting process is done in this way, the matrix element M is obtained correctly and unambiguously.

The use of bare-particle wave-functions Φ_A and Φ_B in (487) is thus justified. This makes the calculation of M in principle simple. It is only necessary to pick out from (488) the terms which contain the right combination of elementary emission and absorption operators to annihilate the particles in A and to create those in B. We shall next describe a general method of systematically picking out those terms, which is due originally to Feynman. It was first explained in published form by G. C. Wick, *Phys. Rev.* **80** (1950) 268. Feynman and Wick have applied the method only to chronologically ordered products such as appear in (488). However the method applies in the same way to all products whether chronological or not, and we shall describe it in full generality.

7.1   The Reduction of an Operator to Normal Form

Given any operator $\mathcal{O}$ which is a product of field operators, for example

$$\mathcal{O} = \overline{\psi}(x_1)\slashed{A}(x_1)\psi(x_1)\overline{\psi}(x_2)\slashed{A}(x_2)\psi(x_2) \tag{496}$$

we want to pick out the matrix element of $\mathcal{O}$ for a transition between states A and B in which there is a known distribution of bare particles. For example A may be a state with only one electron in state 1, and B a state with one electron in state 2. Then we wish to pick out from (496) terms in which there appear the operators b_1 and b_2^*. In order to pick out all such terms systematically, we make a complete analysis of $\mathcal{O}$ into a sum of terms $\mathcal{O}_n$, each $\mathcal{O}_n$ being a sum of products of emission and absorption operators in

which all emission operators stand to the left of all absorption operators.[69] Any operator in which the emission and absorption operators are arranged in this special way is called "Normal"[70] [19]. The $\mathcal{O}_n$ will be called the "Normal constituents" of $\mathcal{O}$. Once $\mathcal{O}$ has been analyzed in this way, then we find the matrix element simply by taking the coefficient of $b_2^* b_1$ in the expansion $\sum \mathcal{O}_n$. No other term in the expansion can give any contribution to the matrix element. In $\mathcal{O}$ itself there might appear a term such as

$$b_2^* b_3 b_3^* b_1 \tag{497}$$

which would give a contribution to the matrix element, since the operator b_3^* could create a particle in an intermediate state 3 which the operator b_3 would then annihilate. The expansion of $\mathcal{O}$ into normal constituents eliminates all terms such as (497) and replaces them by sums of normal products with numerical coefficients. Thus using the anticommutation rule for b_3 and b_3^*, (497) becomes replaced by

$$A b_2^* b_1 - b_2^* b_3^* b_3 b_1 \tag{498}$$

where A is a numerical coefficient. The second term in (498) gives no contribution to the matrix element.

It is clear that by using the commutation rules of the operators in this way, every $\mathcal{O}$ can be written as a sum of normal products, and that the analysis leads to a unique expansion of $\mathcal{O}$. But we do not need to go through the tedious algebra of using the commutation rules, instead we can write down the normal constituents $\mathcal{O}_n$ directly by following simple rules.

First, we define the notation $N(\mathcal{Q})$, where $\mathcal{Q}$ is any product of emission and absorption operators, to be the product obtained by simply rearranging the factors of $\mathcal{Q}$ in a normal order, irrespective of the commutation rules, with a factor (-1) if the rearrangement involves an odd permutation of the electron-positron operators. Similarly if $\mathcal{Q}$ is any sum of products, $N(\mathcal{Q})$ is defined by rearranging factors in each term of the sum in the same way. Thus we have for example (see (211))

$$N(A_\lambda(x) A_\mu(y)) = A_\lambda^+(x) A_\mu^+(y) + A_\lambda^-(x) A_\mu^-(y) + A_\lambda^-(x) A_\mu^+(y) + A_\mu^+(y) A_\lambda^+(x) \tag{499}$$

where $A_\mu^+(x)$ is the positive-frequency part of $A_\mu(x)$, i.e. the part containing absorption operators. Observe that the order of factors in the first two products in (499) is immaterial, only the third and fourth products have

their order fixed by the condition of being normal. Similarly (see (306) and (309))

$$N\big(\psi_\alpha(x)\overline{\psi}_\beta(y)\big) = \psi_\alpha^+(x)\overline{\psi}_\beta^+(y) + \psi_\alpha^-(x)\overline{\psi}_\beta^-(y) + \psi_\alpha^-(x)\overline{\psi}_\beta^+(y) - \overline{\psi}_\beta^-(y)\psi_\alpha^+(x) \tag{500}$$

With this notation, every product of *two* field operators can immediately be written as a sum of normal constituents. Using the commutation rules (213) and the vacuum expectation values given by (219), (220),

$$A_\lambda(x)A_\mu(y) = \langle A_\lambda(x)A_\mu(y)\rangle_o + N(A_\lambda(x)A_\mu(y)) \tag{501}$$

Similarly, using (310), (311), (324),

$$\psi_\alpha(x)\overline{\psi}_\beta(y) = \big\langle \psi_\alpha(x)\overline{\psi}_\beta(y)\big\rangle_o + N\big(\psi_\alpha(x)\overline{\psi}_\beta(y)\big) \tag{502}$$

And in fact for *any* two field operators $\mathcal{P}, \mathcal{Q}$ we have

$$\mathcal{P}\mathcal{Q} = \langle \mathcal{P}\mathcal{Q}\rangle_o + N(\mathcal{P}\mathcal{Q}) \tag{503}$$

provided $\mathcal{P}$ and $\mathcal{Q}$ are both linear in emission and absorption operators. The proof of (503) has in effect been done by proving (501) and (502), because these include all of the possible products of two boson or two fermion operators, and (503) is trivial for the product of one boson and one fermion operator, because they commute. Equations (501) – (503) are operator identities, and hold whether or not the physical problem is directly concerned with the vacuum state of the fields. In fact we could if we wished define the "vacuum expectation values" as the functions appearing in (501) – (503) and so avoid speaking about the vacuum state at all.

   Next we shall state the generalization of the rule (503) to any product $\mathcal{O}$ of field operators, for example the $\mathcal{O}$ given by (496). We define a "factor pairing" of $\mathcal{O}$ by picking out from $\mathcal{O}$ a certain even number of factors, either all or none or any intermediate number, and associating them together in pairs. For the product $\mathcal{P}\mathcal{Q}$ there are only two factor-pairings, either we choose the pair $\mathcal{P}\mathcal{Q}$ together or we choose no pairs at all. To each factor-pairing n corresponds to a normal constituent $\mathcal{O}_n$ obtained as follows: For each pair of factors $\mathcal{P}\mathcal{Q}$ which is paired in n, $\mathcal{O}_n$ contains the numerical factor $\langle \mathcal{P}\mathcal{Q}\rangle_o$, the order of $\mathcal{P}$ and $\mathcal{Q}$ being maintained as it was in $\mathcal{O}$. The unpaired factors $\mathcal{R}_1\mathcal{R}_2\ldots\mathcal{R}_m$ in $\mathcal{O}$ appear in $\mathcal{O}_n$ rearranged in normal form. Thus the complete form of $\mathcal{O}_n$ is

$$\mathcal{O}_n = \pm\,\langle \mathcal{P}\mathcal{Q}\rangle_o\,\langle \mathcal{P}'\mathcal{Q}'\rangle_o\ldots N(\mathcal{R}_1, \mathcal{R}_2,\ldots \mathcal{R}_m) \tag{504}$$

the sign in front being $+$ or $-$ according to the even or odd character of the permutation of the electron-positron operators from the order in which they are written in $\mathcal{O}$ to the order in which they are written in (504). With this definition of (504) of $\mathcal{O}_n$, we have the following theorem:

> *Every operator product $\mathcal{O}$ is identically equal to the sum of the $\mathcal{O}_n$*
>
> *obtained from all its factor-pairings.*

This theorem gives the decomposition of $\mathcal{O}$ into its normal constituents. Equations (501) – (503) are just special cases of it. Clearly non-zero $\mathcal{O}_n$ are only obtained when each pair of factors is either a $\overline{\psi}$ and a ψ operator or two A_μ operators. We shall therefore suppose that the factor-pairings are always restricted in this way.

The proof of the theorem is very simple, by induction on m, the number of factors in $\mathcal{O}$. The theorem is true when $m = 1$ or 2, so we need only prove it true for m assuming it true for $m - 2$. Let then $\mathcal{O}'$ be a product of $(m - 2)$ factors. First we show that the theorem is true for

$$\mathcal{O} = (\mathcal{P}\mathcal{Q} \pm \mathcal{Q}\mathcal{P})\mathcal{O}' \tag{505}$$

where $\mathcal{P}$ and $\mathcal{Q}$ are field operators and the plus sign appears only if $\mathcal{P}$ and $\mathcal{Q}$ are both electron-positron operators. In fact, the normal constituents of $\mathcal{P}\mathcal{Q}\mathcal{O}'$ and of $(\pm\mathcal{Q}\mathcal{P}\mathcal{O}')$ will be identical, so long as $\mathcal{P}$ and $\mathcal{Q}$ are not paired together. Therefore the sum of the normal constituents of $\mathcal{O}$ reduces to

$$\sum \mathcal{O}_n = \{\langle \mathcal{P}\mathcal{Q}\rangle_o \pm \langle \mathcal{Q}\mathcal{P}\rangle_o\} \sum \mathcal{O}'_n \tag{506}$$

But $\sum \mathcal{O}'_n = \mathcal{O}'$, and

$$\langle \mathcal{P}\mathcal{Q}\rangle_o \pm \langle \mathcal{Q}\mathcal{P}\rangle_o = (\mathcal{P}\mathcal{Q} \pm \mathcal{Q}\mathcal{P}) \tag{507}$$

this being a number and not an operator. Therefore (506) gives $\sum \mathcal{O}_n = 0$, and the theorem is proved for $\mathcal{O}$ given by (505). Next let $\mathcal{O}$ be any product of m factors. Then by using the commutation relations we can write

$$\mathcal{O} = N(\mathcal{O}) + \Sigma \tag{508}$$

where Σ is a sum of terms of the form (505). The theorem is true for each term (505) and so is true for Σ. The theorem is trivially true for $N(\mathcal{O})$, because $\langle \mathcal{P}\mathcal{Q}\rangle_o = 0$ for every pair of factors $\mathcal{P}$, $\mathcal{Q}$ in the order in which they occur in $N(\mathcal{O})$, and so all normal constituents (504) of $N(\mathcal{O})$ are zero except the constituent $N(\mathcal{O})$ itself. Therefore the theorem holds for every $\mathcal{O}$ given by (508), and this completes the proof.

7.2   Feynman Graphs

We use a method of Feynman to enumerate the possible factor-pairings of $\mathcal{O}$. Each pairing is pictured in a diagram or graph G. G consists of a certain number of vertices with lines joining them. The vertices represent simply the different field-points at which the factors of $\mathcal{O}$ operate. Thus for $\mathcal{O}$ given by (496) each G has the two vertices x_1, x_2. The lines in G are either dotted, representing photon operators, or undotted, representing electron-positron operators. The rules for drawing these lines are the following:

(1) For each factor-pair $\overline{\psi}(x)\psi(y)$, an undotted line is drawn in G running from x to y, the direction being marked by an arrow in the line.
(2) For each unpaired factor $\overline{\psi}(x)$, an undotted line is drawn running from x out of the diagram, the other end of the line being free and not being a vertex of G.
(3) For each unpaired factor $\psi(y)$, an undotted line is drawn running into y, the other end of the line being free.
(4) For each factor-pair $A_\mu(x)A_\nu(y)$, a dotted line joins x and y.
(5) For unpaired factor $A_\mu(x)$, a dotted line is drawn with one end of x and the other end free.
(6) Every undotted line has a definite direction marked by an arrow. A dotted line has no direction and no arrow.

In general, we must allow factor-pairings in which two operators at the same field-point are paired together. This will give a line in G with both ends at the same point. However in the case of operators such as (496) or more generally (488), a pair of factors taken from the same point will always give rise to a factor

$$\langle j_\mu(x) \rangle_o = -iec \left\langle \overline{\psi}(x)\gamma_\mu\psi(x) \right\rangle_o \tag{509}$$

in the corresponding normal constituents (504). We saw in the discussion following (366) that the vacuum expectation value (509) is zero, the operators being IR operators. Therefore factor-pairings in which two factors at the same field-point are paired, in the analysis of quantum-electrodynamical operators such as (488), always give zero contributions. So we may add to our list of rules for the construction of G:

(7) Lines joining a point to itself are forbidden.

The possible factor pairings of (496) are then represented by the following

G's:
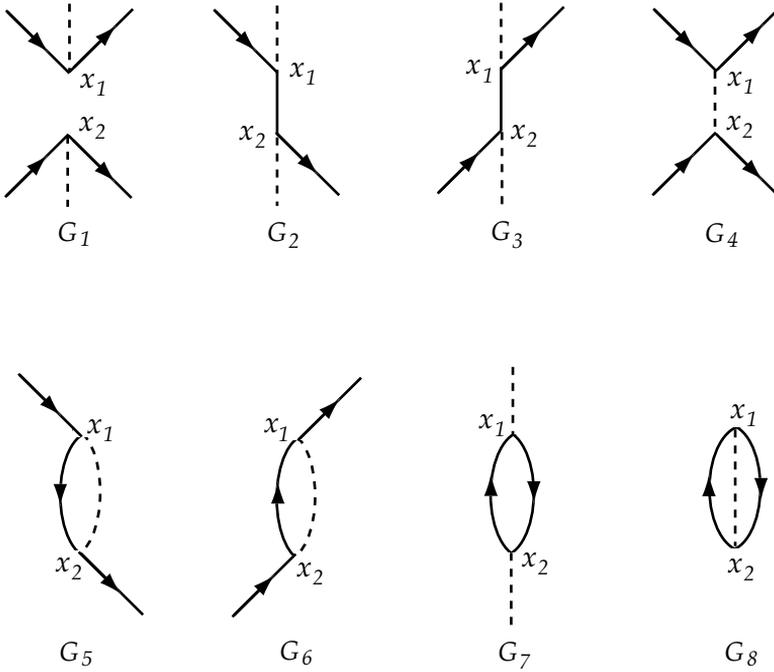
Corresponding to these G's there will be just 8 normal constituents of $\mathcal{O}$, which we call $\mathcal{O}_1 \ldots \mathcal{O}_8$. These normal constituents are, in their respective order:

$$\mathcal{O}_1 = \pm N\{\overline{\psi}(x_1)\slashed{A}(x_1)\psi(x_1)\overline{\psi}(x_2)\slashed{A}(x_2)\psi(x_2)\}$$

$$\mathcal{O}_2 = \pm \left\langle \overline{\psi}(x_1)\psi(x_2)\right\rangle_o N\{\slashed{A}(x_1)\psi(x_1)\overline{\psi}(x_2)\slashed{A}(x_2)\}$$

$$\mathcal{O}_3 = \pm \left\langle \overline{\psi}(x_2)\psi(x_1)\right\rangle_o N\{\overline{\psi}(x_1)\slashed{A}(x_1)\slashed{A}(x_2)\psi(x_2)\}$$

$$\mathcal{O}_4 = \pm \left\langle \slashed{A}(x_1)\slashed{A}(x_2)\right\rangle_o N\{\overline{\psi}(x_1)\psi(x_1)\overline{\psi}(x_2)\psi(x_2)\}$$

$$\mathcal{O}_5 = \pm \left\langle \overline{\psi}(x_1)\psi(x_2)\right\rangle_o \left\langle \slashed{A}(x_1)\slashed{A}(x_2)\right\rangle_o N\{\overline{\psi}(x_2)\psi(x_1)\}$$

$$\mathcal{O}_6 = \pm \left\langle \overline{\psi}(x_2)\psi(x_1)\right\rangle_o \left\langle \slashed{A}(x_1)\slashed{A}(x_2)\right\rangle_o N\{\overline{\psi}(x_1)\psi(x_2)\}$$

$$\mathcal{O}_7 = \pm \left\langle \overline{\psi}(x_1)\psi(x_2)\right\rangle_o \left\langle \overline{\psi}(x_2)\psi(x_1)\right\rangle_o N\{\slashed{A}(x_1)\slashed{A}(x_2)\}$$

$$\mathcal{O}_8 = \pm \left\langle \overline{\psi}(x_1)\psi(x_2)\right\rangle_o \left\langle \overline{\psi}(x_2)\psi(x_1)\right\rangle_o \left\langle \slashed{A}(x_1)\slashed{A}(x_2)\right\rangle_o$$

This type of process to which $\mathcal{O}_i$ gives matrix elements can be seen at once by looking at the *external* lines of G_i, i.e. the lines which have one end free. Thus, for Møller scattering only G_4 contributes. For Compton scattering only G_2 and G_3. And for a transition from a one-electron atom to a one-electron state, which requires an operator of the form $b_2^* b_1$, and G_5 and G_6 will contribute.

A ψ operator not only annihilates electrons but also creates positrons. And a $\overline{\psi}$ not only creates electrons but annihilates positrons. Thus an undotted external line with the arrow pointing in represents *either* an electron in the initial state *or* a positron in the final state. And an undotted external line with the arrow pointing out represents either an outgoing electron or an incoming positron. Dotted external lines likewise represent a photon either in the initial or the final state, because A_μ can both annihilate and create a photon. Thus G_4 will give not only electron-electron scattering, but also electron-positron scattering. G_2 and G_3 give not only the Compton effect, but also two-quantum annihilation of a positron-electron pair, and also the inverse process of pair-creation by the collision of two photons.

We have introduced the Feynman graphs simply as a convenient pictorial way of visualizing the analysis of an operator into its normal constituents. The graphs are just diagrams drawn on the paper. But according to Feynman[71], "Space-time Approach to Quantum Electrodynamics", *Phys. Rev.* **76** (1949) 769, the graphs are more than this. He regards the graphs as a picture of an actual process which is occurring physically in space-time. Thus G_2 represents an electron and a photon coming together and interacting at the space-time point x_1, where the photon is absorbed; then the electron propagates through space-time along the line $x_1 x_2$, until at x_2 it emits a photon, and the electron and photon then travel out along the outgoing lines at x_2. According to Feynman an internal undotted line running from x_1 to x_2 represents an electron propagating from x_1 to x_2 if the time x_2 is later than the time x_1, and represents a positron propagating from x_2 to x_1 if the time x_2 is earlier. In this sense a positron can be regarded as precisely equivalent to an electron which happens to propagate itself backwards instead of forwards in time.

The space-time picture of Feynman is a perfectly consistent and reasonable one. It gives a correct account of everything that happens, including pair creation and annihilation and all the other phenomena connected with positrons. It is in fact equivalent mathematically to the field-theory treatment we follow in this course.

The disadvantage of the Feynman theory is that it is constructed as a particle theory. The fact that there are many particles, indistinguishable from each other and obeying quantum statistics, has to be put into the theory as a special assumption. And the equations of motion of the particles become quite complicated when interactions between several particles, not to mention vacuum-polarization effects, are included. Thus the logical basis of the Feynman theory is much less simple than that of the field theory, where everything follows from general principles once the form of the Lagrangian is chosen.

In this course we follow the pedestrian route of logical development, starting from the general principles of quantizations applied to covariant field equations, and deriving from these principles first the existence of particles and later the results of the Feynman theory. Feynman by the use of imagination and intuition was able to build a correct theory and get the right answers to problems much quicker than we can. It is safer and better for us to use the Feynman space-time pictures not as the basis for our calculations but only as a help in visualizing the formulae which we derive rigorously from the field-theory. In this way we have the advantages of the Feynman theory, its concreteness and its simplification of calculations, without its logical disadvantages.

7.3   Feynman Rules of Calculation

The Feynman rules of calculation arise when we analyze into normal constituents a *chronologically ordered* operator such as (488). In this case the vacuum expectation values in (504) are always taken for pairs of operators which are already chronologically ordered. Hence the numerical factors in (504) are all either

$$\langle P(A_\lambda(x), A_\mu(y)) \rangle_o = \frac{1}{2} \, \hbar c \, D_F(x-y) \, \delta_{\lambda\mu} \tag{510}$$

or

$$\epsilon(x-y) \, \langle P\big(\psi_\alpha(x), \overline{\psi}_\beta(y)\big) \rangle_o = -\frac{1}{2} \, S_{F\alpha\beta}(x-y) \tag{511}$$

using (427) and (442). The factor ϵ is put into (511) so that the $\pm$ sign still characterizes the permutation of electron-positron operators in going from the order *as written* in (504) to the order as written in $\mathcal{O}$. For the same reason, we shall follow Wick and generally use for chronological products the notation

$$T(\mathcal{R}_1 \mathcal{R}_2 \ldots \mathcal{R}_n) = \pm P(\mathcal{R}_1 \mathcal{R}_2 \ldots \mathcal{R}_n) \tag{512}$$

where the sign is plus or minus according to the even or odd character of the permutation of electron-positron operators involved in going from the written order to the chronological order in (512). Hence in particular we have

$$
\begin{aligned}
T\{A_\lambda(x), A_\mu(y)\} &= P\{A_\lambda(x), A_\mu(y)) \\
T\big(\psi_\alpha(x), \overline{\psi}_\beta(y)\big) &= \epsilon(x - y)\, P\big(\psi_\alpha(x), \overline{\psi}_\beta(y)\big)
\end{aligned}
\tag{513}
$$

and for every set of field operators $\mathcal{R}_1 \mathcal{R}_2 \ldots \mathcal{R}_n$ the quantity (512) is a relativistic invariant although the P-product by itself is not. In (488) itself the P-product may be written as a T-product, the sign in (512) in this case always being plus.

The rules for writing down the normal constituents of (488) are therefore extremely simple. We are generally only interested in those normal constituents which give matrix elements for some specified type of scattering process. Then the rules are

(1) Draw all the graphs which have the right set of external lines corresponding to the particles absorbed and emitted in the process considered. Each graph G will have the same external lines, but the number of vertices and of internal lines will vary from graph to graph. We shall always calculate only up to some definite order N in the series (488), and so we draw only graphs with not more than N vertices. The total number of such graphs is finite. Each vertex in each graph must have precisely 3 lines ending at it, one incoming electron line, one outgoing electron line, and one photon line.

(2) To each graph G with n vertices corresponds one normal constituent S_G of S.

(3) Choosing a particular G, write down the nth term S_n of the series (488) and pair off the factors of S_n as indicated by G. Replace each factor-pair $A_\lambda(x)A_\mu(y)$ by (510), and replace each factor-pair $\psi_\alpha(x)\overline{\psi}_\beta(y)$ by (511). Apply an N-ordering to the remaining unpaired factors of S_n, and multiply the whole expression by (± 1) following the rule given for equation (504). The result of applying these operations to S_n is the normal constituent S_G.

If we wish to calculate the matrix element for the scattering process, then we have only to add one more rule to the three already given.

(4) In each S_G, substitute for the unpaired operators the wave-functions of the absorbed and emitted particles, for example writing (437) for $\psi(x)$ when an electron (p, u) is absorbed, and writing (438) for $A_\mu(x)$ when a photon (k, e) is absorbed. These substitutions may sometimes be made in

more than one way (for example in the Compton effect when the absorbed and emitted photon may be assigned in two ways to the two unpaired photon operators.) In such cases the substitutions are to be made in all possible ways, and the results added together, taking account of Fermi statistics by putting in a minus sign when two electron or positron wave-functions are interchanged.

The rules (1)–(4) constitute the Feynman rules for calculating the matrix elements of all processes in electrodynamics. According to Feynman they have an intermediate concrete interpretation. Thus (510) is the probability amplitude for a photon emitted at x with polarization λ to propagate itself and arrive at y with polarization μ, plus the amplitude for a photon having been emitted at y to arrive at x. And (511) is the amplitude for an electron emitted at y to arrive at x, plus the amplitude for a positron emitted at x to arrive at y, with the assigned spins α and β. In this way the matrix element is just the probability amplitude for the succession of events, interactions and propagations, that are depicted in the vertices and lines of G. The total probability amplitude for a process is just the sum of the amplitudes derived by from the various graphs G which contribute to the process.

The Feynman rules of calculation take their most practical form when we use the momentum representations (430) and (448) for the D_F and S_F functions, carry out the integrations over the points $x_1 \ldots x_n$, and so obtain the matrix elements as integrals of rational functions in momentum-space. In this way for example the simple matrix elements (432) and (449) were obtained.

In the momentum space integral for S_G, there will appear

(1)   A factor $\dfrac{1}{k^2}$ corresponding to each internal photon line of G,

$$(514)$$

(2)   A factor $\dfrac{1}{k\!\!\!/ - i\mu}$ corresponding to each internal electron line of G,

$$(515)$$

(3)   A factor $(2\pi)^4 \delta^4(k_1 + k_2 + k_3)$ $\hspace{2cm}$ (516)

   corresponding to each internal photon line of G at which the 3 lines associated with momenta (k_1, k_2, k_3) meet. This factor arises from the integration over the space-time position of the vertex.

In addition to these factors there will be numerical factors and Dirac matrices γ_α arising from the particular form of S_n. In practice it is easiest not to write down the S_G directly in momentum space, but to use the rules (1) – (4) to obtain formulae in configuration space with the right numerical constants, and then transform to momentum space by (430) and (448).

We shall now show how these general methods work by calculating in detail *the* historic problem, the second-order radiative correction to the scattering of an electron by a weak external potential. This problem has been first satisfactorily treated by Schwinger,[72] *Phys. Rev.* **76** (1949) 790. Schwinger's paper is outstandingly difficult to read, and I hope you will find my treatment at least slightly easier. But the problem is in its nature complicated and cannot be done without some fairly heavy mathematics. Once the calculations are done for this problem of scattering, it turns out that the results can be used without much further trouble for the relativistic calculation of the Lamb shift too. The scattering and Lamb shift problems are very closely related: in both cases one is calculating the second-order radiative corrections to the motion of an electron, only in one case the electron is in a high continuum state so that the external field can be treated as weak, in the other case the electron is in a discrete state and the potential must be treated as strong.

7.4 The Self-Energy of the Electron

Before we can study the effect of radiation interaction on an electron scattered by an external potential, we must first consider the effect of the radiation interaction on a single free electron in the absence of external potentials. Let the free electron be given initially in the state (pu). The effect of the radiation interaction acting alone is given by the scattering matrix (488). If the initial state is Φ_A, then the final state, reached after the radiation interaction has been acting for an infinitely long time, will be $S\Phi_A$. Now S has matrix elements only for transitions which conserve momentum and energy. Starting from a one-electron state, it is impossible to make a transition to a many-particle state, for example by emitting one or more photons, while conserving momentum and energy. Therefore the only non-zero matrix elements of S from the state Φ_A will be given by (487), where Φ_B is also a one-electron state. In Φ_B let the electron have the momentum and spin $(p'u')$.

We consider radiative effects only up to the second order. The term of order 1 in (488) gives transitions only with emission an absorption of photons,

and hence gives no contribution to the transition $\Phi_A \to \Phi_B$. Therefore we may write simply

$$S = 1 + U_2 \tag{517}$$

with U_2 given by (423). We have to calculate the matrix element M_2 of U_2 between the states (pu) and $(p'u')$.

To write down M_3 we use the Feynman rules. The factor-pairings of U_2 are represented in the 8 graphs on pp 98-99. Of these only G_5 and G_6 contribute to M_2, and they contribute equally since the integral (423) is symmetrical in the variables x_1 and x_2. The normal constituent of U_2 arising from G_5 and G_6 is, using (510) and (511)[73]

$$U_{2N} = \sum_{\lambda,\mu} \frac{e^2}{\hbar^2 c^2} \iint dx_1\, dx_2\, N\big(\overline{\psi}(x_1)\gamma_\lambda \left\langle T\big(\psi(x_1), \overline{\psi}(x_2)\big)\right\rangle_o \gamma_\mu \psi(x_2)\big)$$

$$\times \left\langle T(A_\lambda(x_1), A_\mu(x_2))\right\rangle_o$$

$$= -\frac{e^2}{4\hbar c} \sum_{\lambda} \iint dx_1\, dx_2\, N\big(\overline{\psi}(x_1)\gamma_\lambda S_F(x_1 - x_2)\gamma_\lambda \psi(x_2)\big) D_F(x_2 - x_1) \tag{518}$$

To obtain M_2 from (518) we substitute for $\psi(x_2)$ and $\overline{\psi}(x_1)$ the wave-functions of the initial and final states, and use the momentum integrals (430), (448). Then the integration over x_1 and x_2 can be carried out and we find

$$M_2 = \sum_{\lambda} \frac{ie^2}{\hbar c} \int_F \int_F dk_1\, dk_2\, \left(\overline{u}'\gamma_\lambda \frac{1}{\slashed{k}_1 - i\mu}\gamma_\lambda u\right) \frac{1}{k_2^2}$$

$$\times \delta(k_1 - k_2 - p')\, \delta(k_2 - k_1 + p)$$

$$= \sum_{\lambda} \frac{ie^2}{\hbar c} \delta(p - p') \int_F dk\, \left(\overline{u}'\gamma_\lambda \frac{1}{\slashed{k} + \slashed{p} - i\mu}\gamma_\lambda u\right) \frac{1}{k^2} \tag{519}$$

We consider the Dirac operator

$$\Sigma(p) = \sum_{\lambda} \int_F dk\, \left(\gamma_\lambda \frac{1}{\slashed{k} + \slashed{p} - i\mu}\gamma_\lambda\right) \frac{1}{k^2} \tag{520}$$

appearing in (519). Since (p, u) are the momentum and spin of a real electron, we may use the relations

$$p^2 + \mu^2 = 0, \qquad (\slashed{p} - i\mu)u = 0 \tag{521}$$

when we evaluate $\Sigma(p)$ in (519). So using (376), (585) and following the same method that was used in evaluating (377)[74]

$$\Sigma(p) = \int_F dk \sum_\lambda \frac{\gamma_\lambda(\not{k} + \not{p} + i\mu)\gamma_\lambda}{k^2(k^2 + 2p \cdot k)} = \int_F dk \frac{4i\mu - 2\not{k} - 2\not{p}}{k^2(k^2 + 2p \cdot k)}$$

$$= 2\int_0^1 dz \int_F dK \frac{i\mu - \not{k}}{[k^2 + 2zp \cdot k]^2} = 2\int_0^1 \int_F dk \frac{(i\mu - \not{k} + z\not{p})}{[k^2 - z^2 p^2]^2}$$

$$= 2\int_0^1 dz \int_F dk \frac{i\mu(1+z)}{[k^2 + z^2\mu^2]^2} \tag{522}$$

where we changed the origin of the k integration by the replacement $k \to k - zp$ and we eliminated the odd terms. Using (386) and introducing the logarithmic divergence R again according to (387),

$$\Sigma(p) = 2\int_0^1 dz\, i\mu(1+z)\{2i\pi^2(R - \log z)\} = -\pi^2\mu\,[6R + 5] = -6\pi^2\mu R' \tag{523}$$

Thus $\Sigma(p)$ is a logarithmically divergent constant, depending only on the electron mass and independent of the state of the electron. The difference $5/6$ between R and R' is of course not significant. Substituting (523) into (519) gives the value of M_2

$$M_2 = -6\pi^2 i\frac{e^2\mu}{\hbar c} R'\delta(p - p')\,(\overline{u}'u) \tag{524}$$

Thus U_2 does not give any transitions between different one-electron states. It has only the diagonal matrix elements given by (524) .

Now (524) has precisely the correct relativistic form to be identified with a pure self-energy effect. Suppose that in consequence of the radiation interaction the mass of a real electron is

$$m = m_o + \delta m \tag{525}$$

where m_o is the mass of the bare electron without interaction and δm is the electromagnetic contribution to the mass. The mass-change δm would be represented by a term

$$\mathscr{L}_S = -\delta m\, c^2\, \overline{\psi}\psi \tag{526}$$

in the Lagrangian density (410). This would produce an interaction energy

$$H_S(t) = \delta m c^2 \int \overline{\psi}(r,t)\psi(r,t)\, d^3\boldsymbol{r} \tag{527}$$

in the Schrödinger equation (415), and finally a contribution

$$U_S = -i\frac{\delta mc}{\hbar} \int \overline{\psi}\psi(x)\,dx \tag{528}$$

in the scattering matrix (421) or (488).

The matrix element of (528) between the states (pu) and $(p'u')$ is

$$M_S = -i\frac{\delta mc}{\hbar}(2\pi)^4\,\delta(p-p')\,(\overline{u}'u) \tag{529}$$

This is identical with (524) if we identify the self-mass δm by the equation

$$\delta m = \frac{3}{8\pi^2}\frac{e^2 m}{\hbar c}R' = \frac{3\alpha}{2\pi}R'm \tag{530}$$

For all one-electron matrix elements, U_2 is identical with U_S. That is to say, the whole effect of the radiation interaction upon a free electron is to change its mass by the amount (530). This is a most satisfactory conclusion. It means that an electron with its self-field still has the correct relationship between momentum and energy for a relativistic particle, only the value of the rest-mass being changed by the self-field. It was always one of the central difficulties of the classical electron theory, that a classical extended electron did not have the right relativistic behavior.

The self-mass δm is an unobservable quantity. The observed mass of an electron is m, and neither m_o nor δm can be measured separately. Thus it is unsatisfactory that δm appears in the scattering matrix S which is supposed to represent the results of experiments.

The reason why δm still appears explicitly is just that we have *not* used the observed mass m in defining the initial and final states of the system. We defined these states as states of a free electron with the bare mass m_o. Wherever we used the letter m for the electron mass in the theory up to this point, in fact it was an inconsistency of notation and we meant by m the mass of a *bare* electron.

It is much better not to change the notation, but to keep the notation and change the interpretation, so that m everywhere in the theory is now intended to mean the mass of a real electron. In particular, we set up the interaction representation operators with the real electron mass m, and the initial and final states of scattering problems are defined as free particles with the correct mass m. With this changed interpretation, the whole of the theory up to this point is correct, except that in $\mathcal{L}_D$ which appears in the

Lagrangian (410) of quantum electrodynamics, and in the field equations (411), (412) satisfied by the Heisenberg operators, the bare mass m_o must be used instead of m. We prefer to keep the observed mass m in $\mathscr{L}_D$, and correct for it by writing instead of (410)[75]

$$\mathscr{L} = \mathscr{L}_D + \mathscr{L}_M - ie\overline{\psi}A\!\!\!/\psi - ie\overline{\psi}A\!\!\!/^e\psi - \mathscr{L}_S \qquad (531)$$

with $\mathscr{L}_S$ given by (526). The radiation interaction becomes then

$$H_R(t) - H_S(t) = H^I(t) \qquad (532)$$

with H_R given by (416) and H_S by (527). After making the changes (537) and (532), the whole theory becomes consistent with the interpretation that m is everywhere the observed electron mass.

In particular, one result of (532) is that for one-electron states the scattering operator S becomes

$$S = 1 + U_2 - U_S \qquad (533)$$

instead of (517), keeping only terms of order e^2. The matrix elements of $(U_2 - U_S)$ for one-electron states are all zero. Thus, if we use the correct mass m in defining the states of an electron, *there are no longer any observable effects of the radiation interaction on the motion of a free electron.* This shows that the mass-renormalization, the procedure of inserting the term $(-\mathscr{L}_S)$ in (531), is consistent and is likely to give sensible results.

7.5   Second-Order Radiative Corrections to Scattering

Let an electron be scattered from the initial state (pu) to the final state $(p'u')$ by the external potential

$$A^e_\mu(x) = \frac{1}{(2\pi)^4} \int e^{iq\cdot x}\, e_\mu(q)\, dq \qquad (534)$$

At the same time the electron is interacting with the quantized Maxwell field with the interaction (532), since we suppose the initial and final states to be defined with the observed mass of a free electron. The scattering matrix element M is then given by (419), with U given by (421) after replacing each H_R by H^I according to (532).

We treat A^e_μ in the linear Born approximation. Thus we keep only terms of order 0 and 1 in A^e_μ. The terms of order 0 give the effects of the radiation

interaction alone; as we have seen, these effects are zero for an initial state consisting of a single electron.

The scattering matrix is thus given effectively by the terms of order 1 in A_μ^e taken from (421), namely

$$U = \sum_{n=0}^{\infty} \left(\frac{-i}{\hbar} \right)^n \frac{1}{n!} \int \cdots \int dt\, dt_1 \ldots dt_n\, P\left\{ H^e(t), H^I(t_1), \ldots, H^I(t_n) \right\}$$

(535)

We shall calculate radiative effects only up to the second order in the radiation interaction. Since δm is itself of second order, this means that we go to second order in H_R and to first order in H_S. Thus

$$U = U_0 + U_1 + U_2 + U_2'$$

(536)

$$U_0 = \frac{e}{\hbar c} \int dx\, \overline{\psi} A^e \psi(x)$$

(537)

$$U_1 = \frac{e^2}{\hbar^2 c^2} \iint dx\, dx_1 P\left\{ \overline{\psi} A^e \psi(x), \overline{\psi} A \psi(x_1) \right\}$$

(538)

$$U_2 = \frac{e^3}{2\hbar^3 c^3} \iiint dx\, dx_1\, dx_2\, P\left\{ \overline{\psi} A^e \psi(x), \overline{\psi} A \psi(x_1), \overline{\psi} A \psi(x_2) \right\}$$

(539)

$$U_2' = \frac{ie\, \delta m}{\hbar^2} \iint dx\, dx_1\, P\left\{ \overline{\psi} A^e \psi(x), \overline{\psi} \psi(x_1) \right\}$$

(540)

The matrix element we wish to calculate is then correspondingly

$$M = M_0 + M_1 + M_2 + M_2'$$

(541)

The wave-functions of initial and final states are

$$u e^{ip\cdot x} \qquad u' e^{ip'\cdot x}$$

(542)

Then by (534) we have

$$M_0 = \frac{e}{\hbar c} (u' \not q u)$$

(543)

where q is the constant vector

$$q = p' - p$$

(544)

and

$$e_\mu = e_\mu(q)$$

(545)

The operator U_1 gives transitions from a one-electron state only to states consisting of an electron and a photon. This is just the bremsstrahlung process, scattering of the electron with real photon emission, and the matrix element for it is given by (484). In any scattering experiment, this process will of course go on at the same time as the scattering without radiation. Experimentally, the scattering with photon emission will only be separable from the radiationless scattering if the emitted photon has an energy greater than some limit ΔE, roughly equal to the energy resolution with which the energy of the electron can be measured. The scattering with emission of soft quanta (low frequency, k' small) will always be included in the radiationless scattering cross-section. Therefore we shall be interested in the value of M_1 for a final state consisting of the electron $(p'u')$ and a photon with potentials (440), in the case where k' is so small as to be negligible in comparison with p, p' and q. In this case (484) gives

$$M_1 = \frac{e^2}{\hbar^2 c^2}\left[\frac{p\cdot e'}{p\cdot k'} - \frac{p'\cdot e'}{p'\cdot k'}\right](\overline{u}'\!\!\not{e}u) = \frac{e}{\hbar c}\left[\frac{p\cdot e'}{p\cdot k'} - \frac{p'\cdot e'}{p'\cdot k'}\right]M_0 \qquad (546)$$

where we used (521) and rule (4) on page 87.

We now come to the calculation of M_2, the matrix element of (539) between the states (542). This is the main part of the problem. To do it we use the Feynman rules. There are just 9 graphs giving contributions to M_2, namely
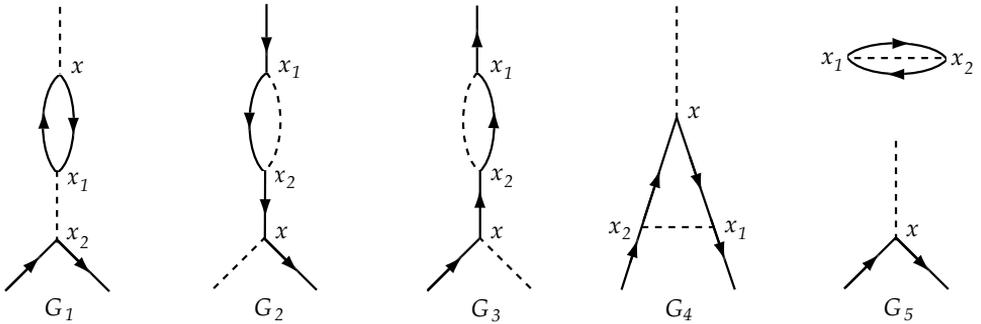
and $\{G_6, G_7, G_8, G_9\}$ obtained by interchanging the labels (x_1, x_2) in $\{G_1, G_2, G_3, G_4\}$, respectively. We can see this in the following way: The process we are interested in calls for one external photon line and two external electron lines. Hence $\not{A}^e(x)$, one $\overline{\psi}$ and one ψ must be unpaired. Therefore the two $\not{A}$'s are always coupled. The free electron lines can be

$\overline{\psi}_0, \psi_1; \overline{\psi}_1, \psi_0; \overline{\psi}_1, \psi_1; \overline{\psi}_1, \psi_2$ and four more cases obtained by the substitution $1 \leftrightarrow 2$. In each case, the rest is uniquely determined by the rules. The ninth case is the one with external electron lines $\overline{\psi}_0, \psi_0$.

The effect of $\{G_6, G_7, G_8, G_9\}$ is only to double the contribution from $\{G_1, G_2, G_3, G_4\}$, since (539) is symmetrical in the variables x_1 and x_2. Also G_5 will give only a numerical phase factor multiplying M_0, the phase factor being the same for all final states; it is in fact precisely the phase-shift between the initial and final states due to the self-energy of the vacuum. Similar phase shift factors would also multiply the contributions of G_1, G_2, etc., coming from disconnected graphs in higher order contributions. In this sense therefore, we can consider G_5 to be really the contribution of U_0 with one of the many possible disconnected additions. A numerical phase factor of this kind, the same for all final states, is entirely unobservable and without physical meaning, since it can be cancelled by changing the phase of all wave-functions by the same amount. Hence we can always ignore graphs such as G_5 having a disconnected part without external lines. There remain to be considered only $\{G_1, G_2, G_3, G_4\}$.

Using the Feynman rules, the contribution of G_1 to M_2 is (with the factor 2 from G_6)

$$M_{21} = -\frac{e^3}{\hbar^3 c^3} \iiint dx\, dx_1\, dx_2 \sum_{\mu} \text{Tr}$$
$$\times \left\{ \mathcal{A}^e(x) \left\langle T\{\psi(x), \overline{\psi}(x_1)\} \right\rangle_o \gamma_\mu \left\langle T\{\psi(x_1), \overline{\psi}(x)\} \right\rangle_o \right\}$$
$$\times \overline{\psi}(x_2) \left\langle T\{A_\mu(x_1), \mathcal{A}(x_2)\} \right\rangle_o \psi(x_2) \tag{547}$$

where the spur appears because of the contraction according to rule 3, and the minus sign comes from the change in order of $\overline{\psi}$ and ψ factors between (539) and (547). By (510) and (511)

$$M_{21} = -\frac{e^3}{8\hbar^2 c^2} \iiint dx\, dx_1\, dx_2 \sum_{\mu} \text{Tr} \left\{ \mathcal{A}^e(x) S_F(x - x_1)\gamma_\mu S_F(x_1 - x) \right\}$$
$$\times D_F(x_1 - x_2)\overline{\psi}(x_2)\gamma_\mu \psi(x_2)$$

Hence, using the momentum integrals (430), (448), (534), (542) and carrying out the integration over (x, x_1, x_2),

$$M_{21} = \frac{ie^3}{(2\pi)^4 \hbar^2 c^2} \iiiint dk_1 \, dk_2 \, dk_3 \, dq$$

$$\times \sum_\mu \text{Tr} \left\{ \phi(q) \frac{1}{\slashed{k}_1 - i\mu} \gamma_\mu \frac{1}{\slashed{k}_2 - i\mu} \right\} \frac{1}{k_3^2} \left(\overline{u}' \gamma_\mu u \right)$$

$$\times \delta(q + k_1 - k_2) \, \delta(-k_1 + k_2 + k_3) \, \delta(-k_3 - p' + p)$$

$$= \frac{ie^3}{(2\pi)^4 \hbar^2 c^2} \int_F dk \sum_\mu \text{Tr} \left\{ \phi \frac{1}{\slashed{k} - i\mu} \gamma_\mu \frac{1}{\slashed{k} + \slashed{q} - i\mu} \right\} \frac{1}{q^2} \left(\overline{u}' \gamma_\mu u \right)$$

$$= \frac{ie^3}{(2\pi)^4 \hbar^2 c^2} \sum_\mu \frac{1}{q^2} \left(\overline{u}' \gamma_\mu u \right) J_\mu \tag{548}$$

with q given by (544) and[76]

$$J_\mu = \int_F F_\mu(k) \, dk \tag{549}$$

the function $F_\mu(k)$ being identical with (371) for $\delta = 0$. Note that (549) is a Feynman integral, which is precisely the same thing as the contour integral (374) with the contour drawn in the diagram. The effect of the ϵ in (431) is just equivalent to the contour C. Hence[77] using (388) for J_μ,

$$M_{21} = -\frac{e^3}{2\pi^2 \hbar^2 c^2} \left(\overline{u}' \phi u \right) \left\{ \frac{1}{3} R - \int_0^1 (z - z^2) \log \left(1 + (z - z^2) \frac{q^2}{\mu^2} \right) dz \right\} \tag{550}$$

where we have also dropped the term q_μ in (388), since

$$\left(\overline{u}' \slashed{q} u \right) = \left\{ \overline{u}' (\slashed{p}' - i\mu) u \right\} - \left\{ \overline{u}' (\slashed{p} - i\mu) u \right\} = 0 \tag{551}$$

Writing $\alpha = \frac{e^2}{4\pi \hbar c}$, (550) becomes

$$M_{21} = \alpha M_0 \left\{ -\frac{2}{3\pi} R + \frac{2}{\pi} \int_0^1 (z - z^2) \log \left(1 + (z - z^2) \frac{q^2}{\mu^2} \right) dz \right\} \tag{552}$$

This is just the scattering which is produced by the charge-current density induced in the vacuum by the potential A_μ^e according to (392). As before, the term in R is unobservable since it can never be separated experimentally from the simple scattering M_0 to which it is proportional. The observed external potential, measured in any way whatever, will not be A_μ^e but $A_\mu^e(1 - \frac{2\alpha}{3\pi} R)$,

which we may call the "renormalized external potential". Hence in terms of the observed A_μ^e the total contribution from G_1 will be

$$M_{21} = \frac{2\alpha}{\pi} M_0 \int_0^1 (z - z^2) \log \left(1 + (z - z^2) \frac{q^2}{\mu^2} \right) dz \tag{553}$$

The integral will be in general complex as before. But for small q it will be real, and neglecting terms of order higher than q^2, we have

$$M_{21} = \frac{2\alpha}{\pi} \frac{q^2}{\mu^2} M_0 \int_0^1 (z - z^2)^2 dz = \frac{\alpha}{15\pi} M_0 \frac{q^2}{\mu^2} \tag{554}$$

Next we consider the contribution to M_2 from G_2. This is

$$M_{22} = \frac{e^3}{8\hbar^2 c^2} \iiint dx\, dx_1\, dx_2 \sum_\lambda \overline{\psi}(x) A^e(x)$$

$$\times S_F(x - x_2)\gamma_\lambda S_F(x_2 - x_1)\gamma_\lambda \psi(x_1) D_F(x_1 - x_2) \tag{555}$$

$$= -\frac{ie^3}{(2\pi)^4 \hbar^2 c^2} \iiiint dk_1\, dk_2\, dk_3\, dq$$

$$\times \sum_\lambda \left\{ \overline{u}' \slashed{\phi}(q) \frac{1}{\slashed{k}_1 - i\mu} \gamma_\lambda \frac{1}{\slashed{k}_2 - i\mu} \gamma_\lambda u \right\} \frac{1}{k_3^2} \times$$

$$\times \delta(k_1 + q - p')\, \delta(k_2 - k_1 - k_3)\, \delta(k_3 + p - k_2) \tag{556}$$

$$= -\frac{ie^3}{(2\pi)^4 \hbar^2 c^2} \sum_\lambda \int_F dk \left(\overline{u}' \slashed{\phi} \frac{1}{\slashed{p} - i\mu} \gamma_\lambda \frac{1}{\slashed{k} + \slashed{p} - i\mu} \gamma_\lambda u \right) \frac{1}{k^2} \tag{557}$$

$$= -\frac{ie^3}{(2\pi)^4 \hbar^2 c^2} \left(\overline{u}' \slashed{\phi} \frac{1}{\slashed{p} - i\mu} \Sigma(p)\, u \right) \tag{558}$$

with $\Sigma(p)$ given by (520).

Before discussing $\Sigma(p)$ we must look at the factor $\frac{1}{\slashed{p} - i\mu}$ which appears in (558). This factor is

$$\frac{\slashed{p} + i\mu}{p^2 + \mu^2} \tag{559}$$

But since p is the momentum vector of a real electron, $p^2 + \mu^2 = 0$ and the factor (559) is singular. This means that the integrals over x_1 and x_2 are really divergent and not merely finitely oscillating at $t = \pm\infty$, and the

transformation into momentum integrals is not allowable. Eq. (558) as it stands is strictly meaningless.

This is the place where we have to take explicitly into our calculations the slowly-varying cut-off functions $g(t_i)$ appearing in (495), which are put in for the purpose of defining unambiguously the initial and final states of the problem. So we write instead of (555)

$$M_{22} = \frac{e^3}{8\hbar^2 c^2} \iiint dx\, dx_1\, dx_2 \sum_\lambda \overline{\psi}(x) \slashed{A}^e(x) S_F(x - x_2) \gamma_\lambda S_F$$

$$\times (x_2 - x_1) \gamma_\lambda \psi(x_1) D_F(x_1 - x_2) g(t_1) g(t_2) \tag{560}$$

Here the $g(t_1)g(t_2)$ factors are attached to the radiation interaction operating at x_1 and x_2. It is supposed that the time T over which $g(t)$ varies appreciably is long compared with the duration of the scattering process. Let the Fourier integral representation of $g(t)$ be

$$g(t) = \int_{-\infty}^{\infty} G(\epsilon_0) e^{-i\epsilon_0 ct}\, d\epsilon_0$$

$$= \int_{-\infty}^{\infty} G(\epsilon_0) e^{i\epsilon \cdot x}\, d\epsilon_0 \tag{561}$$

where ϵ_0 is a real variable and ϵ is the vector

$$\epsilon = (0, 0, 0, \epsilon_0) \tag{562}$$

We have the normalization

$$g(0) = \int_{-\infty}^{\infty} G(\epsilon_0)\, d\epsilon_0 = 1 \tag{563}$$

and we suppose that $G(\epsilon_0)$ is "almost" a δ-function, that is to say a function which is large only for values of ϵ_0 in a range of about $(cT)^{-1}$ on either side of zero. Substituting (561) into (560), we obtain instead of (558) the corrected formula

$$M_{22} = -\frac{ie^3}{(2\pi)^4 \hbar^2 c^2} \iint G(\epsilon_0)\, G(\epsilon_0')\, d\epsilon_0\, d\epsilon_0'$$

$$\times \left\{ \overline{u}\, \slashed{\phi}(q - \epsilon - \epsilon') \frac{1}{\slashed{p} + \slashed{\phi} + \slashed{\phi}' - i\mu} \Sigma(p + \epsilon)\, u \right\} \tag{564}$$

In (564) the inadmissible factor (559) is replaced by something finite and mathematically well-defined. There will be a singularity in the integration of (564) over ϵ_0, but this is an ordinary pole and the integration over ϵ_0 will give a well-determined result when taken as a Feynman integral. We have as $T \to \infty$ and ϵ_0 and $\epsilon_0' \to 0$

$$\frac{1}{\not{p} + \not{\epsilon} + \not{\epsilon}' - i\mu} = \frac{\not{p} + \not{\epsilon} + \not{\epsilon}' + i\mu}{2p \cdot (\epsilon + \epsilon') + (\epsilon + \epsilon')^2} \sim -\frac{\not{p} + i\mu}{2p_0(\epsilon_0 + \epsilon_0')} \tag{565}$$

Hence when evaluating $\Sigma(p + \epsilon)$ we need retain only terms of order zero and one in ϵ_0; the terms of order two and higher are negligible, because even when multiplied by (565) they tend to zero as $T \to \infty$.

Keeping only terms of order zero and one in ϵ, $\Sigma(p + \epsilon)$ becomes

$$\Sigma(p + \epsilon) = \Sigma(p) - \sum_{\alpha} \epsilon_\alpha I_\alpha(p) \tag{566}$$

$$I_\alpha(p) = \int_F dk \sum_\lambda \left(\gamma_\lambda \frac{1}{\not{k} + \not{p} - i\mu} \gamma_\alpha \frac{1}{\not{k} + \not{p} - i\mu} \gamma_\lambda \right) \frac{1}{k^2} \tag{567}$$

Here we have used the identity

$$\frac{1}{A + B} = \frac{1}{A} - \frac{1}{A} B \frac{1}{A} + \frac{1}{A} B \frac{1}{A} B \frac{1}{A} - \dots \tag{568}$$

which is valid for any two operators A and B, commuting or not, if the series on the RHS converges in some sense. This can be seen at once by multiplying through by $A + B$; then the condition becomes $(B/A)^n \to 0$ in some sense.

In (564) we may use the conditions (521), and these give for $\Sigma(p)$ the constant value (523). The integral $I_\alpha(p)$ is like $\Sigma(p)$ logarithmically divergent for large k, and it is also logarithmically divergent for small k, which $\Sigma(p)$ is not. We shall not attempt to evaluate $I_\alpha(p)$ mathematically. From general principles of covariance we can say what is its form as a function of p. For general p not satisfying (521), $I_\alpha(p)$ is a Dirac matrix transforming like a vector under Lorentz transformations, and therefore must be of the form

$$I_\alpha(p) = F_1(p^2)\gamma_\alpha + F_2(p^2)(\not{p} - i\mu)\gamma_\alpha + F_3(p^2)\gamma_\alpha(\not{p} - i\mu)$$

$$+ F_4(p^2)(\not{p} - i\mu)\gamma_\alpha(\not{p} - i\mu) \tag{569}$$

where $F_1, \dots, F_4$ are functions of the scalar p^2. Hence using (521) and (523), we see that in (564) we may put

$$\Sigma(p + \epsilon) = -6\pi^2 \mu R' - I_1 \not{\epsilon} - I_2(\not{p} - i\mu)\not{\epsilon} \tag{570}$$

where I_1 and I_2 are new absolute constants, and in particular

$$I_1 = F_1(-\mu^2) \tag{571}$$

But in (564) the term

$$\left(\frac{1}{\not{p} + \not{\epsilon} + \not{\epsilon}' - i\mu} \right) (\not{p} - i\mu)\not{\epsilon} = \not{\epsilon} - \frac{1}{(\not{p} + \not{\epsilon} + \not{\epsilon}' - i\mu)}(\not{\epsilon} + \not{\epsilon}')\not{\epsilon}$$

is of order ϵ and tends to zero as $T \to 0$. This term may be dropped, and then (564) becomes[78]

$$M_{22} = -\frac{ie^3}{(2\pi)^4\hbar^2c^2} \iint G(\epsilon_0)\, G(\epsilon_0')\, d\epsilon_0\, d\epsilon_0'$$

$$\times \left\{ \overline{u}\,\not{\epsilon}(q - \epsilon - \epsilon')\frac{1}{\not{p} + \not{\epsilon} + \not{\epsilon}' - i\mu}(-6\pi^2\mu R' - I_1\not{\epsilon})\, u \right\} \tag{572}$$

Note that if $I_\alpha(p)$ given by (567) were to be evaluated assuming that $p^2+\mu^2 = 0$ and $\not{p} - i\mu = 0$ operating both to the left and to the right, instead of only to the right as in (521), the result obtained would be just

$$I_\alpha(p) = I_1\gamma_\alpha \tag{573}$$

This is a convenient definition of I_1 for future reference.

Now it is clear that the term R' in M_{22} represents some kind of effect of the electron self-energy, which ought not to be observable. We may expect that this term will be cancelled by the term M_2' arising from the self-energy correction H_S in (532). This is all the more plausible, because the graph G_2 on page 106 contains as a part the graph G_5 on page 99 which represents the self-energy of a free electron. We now turn to the calculation of M_2'.

M_2' is the sum of two contributions arising from the two graphs shown below.
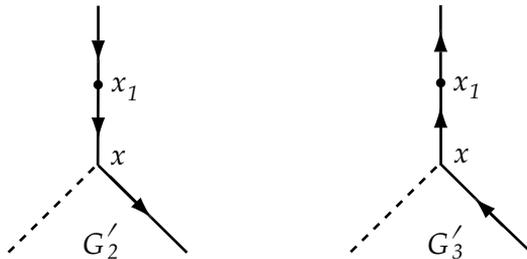
From G_2' the contribution is (compare with (528))

$$M_{22}' = -\frac{ie\,\delta m}{2\hbar^2} \iint dx\,dx_1\,\overline{\psi}(x)\slashed{A}^e S_F(x - x_1)\,\psi(x_1) \tag{574}$$

Like (555) this integral does not oscillate but diverges at $t_1 = \pm\infty$. Therefore we must take explicitly into account the cut-off factor multiplying the radiation interaction. At time t_1 the radiation interaction $H_R(t_1)$ will carry the cut-off factor $g(t_1)$. But the self-energy δm at time t_1 is an effect of second order in H_R, and therefore becomes multiplied by $[g(t_1)]^2$ if $g(t_1)$ varies slowly enough. In the definition of the scattering matrix element (487), the cut-off factors $g(t)$ were introduced in order to represent the initial and final states by simple bare-particle wave-functions in an unambiguous way. We now require that the bare-particles' wave-functions should always have the same mass m as a real electron. This is achieved if we add the term $(-H_S)$ to the radiation interaction appearing in (495), each $H_S(t_i)$ being multiplied by the cut-off factor $[g(t_i)]^2$ so long as we are dealing only with second order terms in δm. If we were calculating effects to fourth order in e, then the fourth order part of δm should be multiplied by $[g(t_i)]^4$, and so on.

The effect of the cut-off factors is therefore to replace (574) by

$$M_{22}' = -\frac{ie\,\delta m}{2\hbar^2} \iint dx\,dx_1\,\overline{\psi}(x)\slashed{A}^e S_F(x - x_1)\,\psi(x_1)\,[g(t_1)]^2 \tag{575}$$

Using (561) and performing the integrations as before, this becomes

$$M_{22}' = -\frac{ie\,\delta m}{2\hbar^2} \iint G(\epsilon_0)G(\epsilon_0')\,d\epsilon_0\,d\epsilon_0'$$

$$\times \left\{ \overline{u}'\slashed{\phi}(q - \epsilon - \epsilon')\left(\frac{1}{\slashed{p} + \slashed{\epsilon} + \slashed{\epsilon}' - i\mu}\right) u \right\} \tag{576}$$

In virtue of (530), this term does precisely cancel the term in R' in (572), as was to be expected.

To simplify the term in I_1 in (572), we may replace $\slashed{\epsilon}$ by $\frac{1}{2}(\slashed{\epsilon} + \slashed{\epsilon}')$ since the integrand is otherwise symmetrical between ϵ and ϵ'. And using (521) we may replace this in turn by $\frac{1}{2}(\slashed{p} + \slashed{\epsilon} + \slashed{\epsilon}' - i\mu)$. This cancels the denominator of (572) precisely. After the denominator is cancelled, the expression is non-singular and we may go to the limit $T \to \infty$, using (563) to perform the integration over ϵ_0 and ϵ_0'. Since it is assumed that the external potential is of a limited duration not tending to infinity with T, the factor $\slashed{\phi}(q - \epsilon - \epsilon')$

is a continuous function of $\epsilon + \epsilon'$ and tends to $\rlap{/}{\phi}(q)$ as $T \to \infty$.[79] Hence in the limit $T \to \infty$ we have

$$M_{22} + M_{22}' = \frac{ie^3}{(2\pi)^4 \hbar^2 c^2} \frac{1}{2} I_1 \left(\overline{u}' \rlap{/}{\phi} u \right) = \frac{i\alpha}{(2\pi)^3} I_1 M_o \tag{577}$$

The graphs G_3 and G_3' give an exactly equal contribution. Hence

$$M_{22} + M_{23} + M_{22}' + M_{23}' = \frac{i\alpha}{4\pi^3} I_1 M_o \tag{578}$$

It turns out that I_1 is pure imaginary, and the factor multiplying M_o in (578) is real and negative.

What is the physical interpretation of the divergent term (578)? It is just a divergent constant multiple of M_o, like the charge-renormalization term in (552). So one is tempted at first to call it an additional charge-renormalization effect. But this cannot be right, because the whole charge-renormalization was calculated in (392) and the result agreed with (552). In fact (578) has a much more elementary interpretation. When the electron arrives at the point x of the external potential where it undergoes scattering, there will be a certain probability P that it will have previously emitted and not yet reabsorbed a photon, as pictured in the Feynman graph G_4. There will be a probability $(1 - P)$ that it will arrive at x not accompanied by a photon, as pictured in G_2 or G_3.

Consider then the contribution M_{NP} to the matrix element M, produced by scattering processes in which the electron arrives at x with no photon present. In the zero-order approximation $M_{NP} = M_o$ simply. But in the second-order approximation we must take account of the reduced probability that the electron comes to x without a photon present; this is done by multiplying the wave-function of the electron in both initial and final states by the factor

$$(1 - P)^{1/2} \tag{579}$$

Therefore in second-order approximation

$$M_{NP} = (1 - P)M_o \tag{580}$$

Since in second order

$$M_{NP} = M_o + M_{22} + M_{22}' + M_{23} + M_{23}', \tag{581}$$

the result (578) agrees with (580) provided that we identify P with

$$P = -\frac{i\alpha}{4\pi^3} I_1 \tag{582}$$

The factor (579) represents a renormalization of the amplitude of the wave-function, and for this reason (578) is usually called a "wave-function renormalization" effect. But this does not mean that the term (578) is to be removed by a process analogous to mass or charge renormalization. No difficulties arise if we simply retain (578) as it stands. Finally it will cancel out against a term $(+PM_o)$ which comes from G_4.

The contribution from G_4 to M is, with a factor 2 to allow for G_9, using the Feynman rules,

$$M_{24} = \frac{e^3}{8\hbar^2 c^2} \iint dx\, dx_1\, dx_2 \sum_\lambda \overline{\psi}(x)\gamma_\lambda S_F(x_1 - x)\slashed{A}^e(x)S_F$$

$$\times (x - x_2)\gamma_\lambda\psi(x_2)D_F(x_2 - x_1)$$

$$= -\frac{ie^3}{(2\pi)^4\hbar^2 c^2} \left(\overline{u}'\Lambda(p,p')u\right) \tag{583}$$

where

$$\Lambda(p,p') = \int_F dk \sum_\lambda \left(\gamma_\lambda \frac{1}{\slashed{k} + \slashed{p}' - i\mu}\slashed{e}\frac{1}{\slashed{k} + \slashed{p} - i\mu}\gamma_\lambda\right)\frac{1}{k^2} \tag{584}$$

There is no singular factor in (584) such as we had in (558). To perform the sum over λ in (584), we use the table

$$\sum_\lambda \gamma_\lambda\gamma_\lambda = 4$$

$$\sum_\lambda \gamma_\lambda\slashed{a}\gamma_\lambda = -2\slashed{a}$$

$$\sum_\lambda \gamma_\lambda\slashed{a}\slashed{b}\gamma_\lambda = 4(a \cdot b) \tag{585}$$

$$\sum_\lambda \gamma_\lambda\slashed{a}\slashed{b}\slashed{c}\gamma_\lambda = -2\slashed{c}\slashed{b}\slashed{a}$$

valid for any vectors a, b, c. These formulae can be deduced from the following recursive formula:

Denote $\slashed{q}_{(n)} = \slashed{q}_1\slashed{q}_2\cdots\slashed{q}_n$ where q_i are arbitrary vectors, and $\chi_n = \sum_\lambda \gamma_\lambda\slashed{q}_{(n)}\gamma_\lambda$, $\chi_o = 4$. Then we have

$$\chi_{n+1} = \sum_\lambda \gamma_\lambda \not{q}_{(n)} \not{q}_{n+1} \gamma_\lambda = \sum_\lambda \sum_\mu \gamma_\lambda \not{q}_{(n)} \gamma_\mu \gamma_\lambda (q_{n+1})_\mu$$

$$= \sum_\lambda \sum_\mu \gamma_\lambda \not{q}_{(n)} \left[2\delta_{\lambda\mu} - \gamma_\lambda \gamma_\mu \right] (q_{n+1})_\mu$$

$$= 2 \not{q}_{n+1} \not{q}_{(n)} - \chi_n \not{q}_{n+1}$$

which then gives (585) for $n = 1, 2, 3$. Thus

$$\Lambda(p, p') = -2 \int_F dk \, \frac{(\not{k} + \not{p}) \not{e} (\not{k} + \not{p}') - 2i\mu (2e \cdot k + e \cdot p + e \cdot p') - \mu^2 \not{e}}{k^2 \left[k^2 + 2k \cdot p' \right] \left[k^2 + 2k \cdot p \right]}$$

$$(586)$$

In (583) we may use the relations

$$p^2 + \mu^2 = p'^2 + \mu^2 = 0 \qquad (\not{p} - i\mu)u = 0 \qquad \overline{u}'(\not{p}' - i\mu) = 0 \qquad (587)$$

We also assume that the external potential satisfies the Lorentz condition

$$\sum_\lambda \frac{\partial A_\mu^e}{\partial x_\mu} = 0 \quad \text{so that} \quad e \cdot q = 0 \tag{588}$$

To evaluate (586) we use the 3-variable generalization of (376)

$$\frac{1}{abc} = 2 \int_0^1 dx \int_0^1 x \, dy \, \frac{1}{[a(1-x) + bxy + cx(1-y)]^3} \tag{589}$$

which one can verify at once by direct integration. We write

$$p_y = py + p'(1-y)$$

$$p_y^2 = [-(p'-p)y + p']^2 = q^2 y^2 - (2p'^2 - 2p \cdot p')y - \mu^2$$

$$= q^2 y^2 - (p'^2 - 2p \cdot p' + p^2) - \mu^2$$

$$= -\mu^2 - (y - y^2)q^2 \tag{590}$$

Then changing the origin of the k-integration by the substitution $k \to k - xp_y$ (586) and (589) give

$$\Lambda(p, p') = -4 \iint x \, dx \, dy \int_F dk$$

$$\times \frac{(\not{k} - x\not{p}_y + \not{p}) \not{e} (\not{k} - x\not{p}_y + \not{p}') - 2i\mu e \cdot (2k - 2xp_y + p + p') - \mu^2 \not{e}}{\left[k^2 - x^2 p_y^2 \right]^3}$$

$$(591)$$

In (591) we may drop terms which are odd in k. Also using (587) and (588) we may put

$$e \cdot p = e \cdot p' = e \cdot p_y = i\mu \not{e} + \frac{1}{2} \not{e} \not{q} \tag{592}$$

$$(\not{p} - x \not{p}_y) \not{e} (\not{p}' - x \not{p}_y)$$
$$= \left\{ (1-x) i\mu - (1-xy) \not{q} \right\} \not{e} \left\{ (1-x) i\mu + (1-x+xy) \not{q} \right\}$$
$$= -(1-x)^2 \mu^2 \not{e} + (1-x) i\mu \not{e} \not{q} (2-x) + (1-xy)(1-x+xy) q^2 \tag{593}$$

Collecting terms from (592), (593), we have

$$\Lambda(p, p') = -4 \iint x \, dx \, dy \int_F dk$$
$$\times \frac{\not{k} \not{e} \not{k} + (1-xy)(1-x+xy) q^2 \not{e} - (x - x^2) i\mu \not{e} \not{q} + (2 - 2x - x^2) \mu^2 \not{e}}{[k^2 + x^2 (\mu^2 + (y - y^2) q^2)]^3}$$
$$\tag{594}$$

Now we saw earlier that the integral (567), evaluated using the conditions $p^2 + \mu^2 = 0, \not{p} - i\mu = 0$, has the value (573). Comparing (567) with (584), this implies that

$$\Lambda(p, p) = \Lambda(p', p') = I_1 \not{e} \tag{595}$$

when the conditions (587) with $p' = p$ are assumed to hold. Thus $(I_1 \not{e})$ is just the value of (594) for $p' = p$. Adding together (583) and (578), this gives

$$M_{2T} = M_{24} + M_{22} + M_{22}' + M_{23} + M_{23}' = -\frac{ie^3}{(2\pi)^4 \hbar^2 c^2} \left(\overline{u}' \Lambda_c(p, p') u \right) \tag{596}$$

$$\Lambda_c(p, p') = \Lambda(p, p') - \frac{1}{2} \left\{ \Lambda(p, p) + \Lambda(p', p') \right\}$$

$$= -4 \iint x \, dx \, dy \left\{ \int_F dk \, \left[\not{k} \not{e} \not{k} + (2 - 2x - x^2) \mu^2 \not{e} \right] \right.$$

$$\times \left(\frac{1}{[k^2 + x^2 (\mu^2 + (y - y^2) q^2)]^3} - \frac{1}{[k^2 + x^2 \mu^2]^3} \right)$$

$$+ \left[(1 - xy)(1 - x + xy)q^2 \slashed{\phi} - (x - x^2)i\mu\slashed{\phi}\slashed{q}\right] \int_F dk$$

$$\times \left. \frac{1}{[k^2 + x^2 \left(\mu^2 + (y - y^2) q^2\right)]^3} \right\}$$

(597)

The k-integrals in (597) are now convergent. Thus the effect of the "wave-function renormalization" term (578) is just to cancel out the part of M_{24} which is independent of q and divergent at high frequencies.

To evaluate (597) we use (385) and (386). First, in the term $\slashed{k}\slashed{\phi}\slashed{k}$ we may replace $k_\mu k_\nu$ by $\frac{1}{4}\delta_{\mu\nu}k^2$ because of the symmetry of the integral in k-space. So we may write, using (585)

$$\slashed{k}\slashed{\phi}\slashed{k} = \sum_\alpha \frac{1}{4}k^2 \gamma_\alpha \slashed{\phi} \gamma_\alpha = -\frac{1}{2}k^2 \slashed{\phi}$$

(598)

Then by (385) and (386)

$$\int_F dk\, k^2 \left\{\frac{1}{[k^2 + \Lambda]^3} - \frac{1}{[k^2 + \Lambda']^3}\right\}$$

$$= \int_F dk\, k^2 \left\{\frac{1}{[k^2 + \Lambda]^2} - \frac{1}{[k^2 + \Lambda']^2} - \frac{\Lambda}{[k^2 + \Lambda]^3} + \frac{\Lambda'}{[k^2 + \Lambda']^3}\right\}$$

$$= \pi^2 i \log \frac{\Lambda'}{\Lambda}$$

(599)

and (597) becomes

$$\Lambda_c(p, p') = 2\pi^2 i \iint x\, dx\, dy \left\{-\slashed{\phi} \log\left[1 + (y - y^2)\frac{q^2}{\mu^2}\right]\right.$$

$$\left. + \frac{\{x - 1 + 2(y - y^2)(1 - x - x^2)\}q^2\slashed{\phi} + (x - x^2)i\mu\slashed{\phi}\slashed{q}}{x^2(\mu^2 + (y - y^2) q^2}\right\}$$

Integrating the logarithmic term by parts with respect to y,

$$\Lambda_c(p, p') = -2\pi^2 i \int_0^1 \int_0^1 dx\, dy\, \frac{1}{x\left[\mu^2 + (y - y^2) q^2\right]}$$

$$\times \left\{\left[(1 - x)(1 - 2y + 2y^2) + x^2 y\right] q^2\slashed{\phi} - (x - x^2)i\mu\slashed{\phi}\slashed{q}\right\}$$

(600)

When $q^2 < -4\mu^2$, the external potential is able to create real pairs, and the denominator in (600) has poles in the range of the y-integration.

In this case the Feynman rule of adding a term $(-i\epsilon)$ to μ^2, where ϵ is an infinitesimal positive real number, will give an unambiguous determination of the integral. Just as in the case of the vacuum polarization formula (389), the integral will split into a real part and an imaginary part describing the effects of the real pairs created. We shall not discuss these effects of real pairs since they are not practically important. So we assume $q^2 > -4\mu^2$.

In (600) there are no longer any divergences arising from large k. But (584) has a logarithmic divergence at small k which appears in (600) as a divergence in the x-integration arising from the factor $(1/x)$. This last remaining divergence must now be examined in detail. It is the famous "Infra-Red Catastrophe".

To discover the physical meaning of the x-divergence, we consider what would be the effect on our calculations if the Maxwell field were somehow modified so that all field oscillations with wave-numbers satisfying

$$|\boldsymbol{k}| \geq r \tag{601}$$

were present as usual, while all oscillations not satisfying (601) were simply absent or incapable of being excited. We suppose r to be a constant small compared with m, p, p' and q. Thus photons will exist only if their energies exceed

$$\Delta E = \hbar c r \tag{602}$$

In the modified Maxwell theory the D_F function will still be given by the integral (431), the k_1, k_2, k_3 integrations being limited by (601), and the k_0 integration being taken as usual along the whole real axis from $-\infty$ to $+\infty$. Let $\Lambda^r(p, p')$ and $\Lambda_c^r(p, p')$ be the integrals which replace $\Lambda(p, p')$ and $\Lambda_c(p, p')$ when the Maxwell field is modified. We calculate the differences $(\Lambda - \Lambda^r)$ and $(\Lambda_c - \Lambda_c^r)$, considering these integrals only in the limit of small r, neglecting all terms which tend to zero with r. This means that we may neglect terms containing either k or x as a factor in the numerator of integrals such as (591) or (594).

In (583) there is just one factor D_F. Thus $\Lambda^r(p, p')$ is obtained from (584) simply by restricting the k_1, k_2, k_3 integration by (601). We can now follow the reduction of (584) to the form (594), except that we do not shift the origin of the k-integration through (xp_y) since this would disturb the condition (601). Dropping terms in the numerator having k or x as a factor, this gives the result

$$\Lambda(p, p') - \Lambda^r(p, p') = -4 \iint x \, dx \, dy \, (q^2 + 2\mu^2) \not{q} \int_F \frac{dk}{[k^2 + 2xk \cdot p_y]^3} \tag{603}$$

Hence by (597)

$$\Lambda_c(p, p') - \Lambda_c^r(p, p')$$

$$= -4 \iint x \, dx \, dy \, \displaystyle{\not}q \int_F dk$$

$$\times \left\{ \frac{q^2 + 2\mu^2}{[k^2 + 2xk \cdot p_y]^3} - \frac{\mu^2}{[k^2 + 2xk \cdot p]^3} - \frac{\mu^2}{[k^2 + 2xk \cdot p']^3} \right\} \qquad (604)$$

The integral (604), with the integration extended over the whole k-space, would give, using (385), (587), and the substitutions $k \to k - xp_y, k \to k - xp, k \to xp'$, respectively in the three integrals,

$$-2\pi i \iint dx \, dy \, \frac{(1 - 2y + 2y^2) \, q^2}{x(\mu^2 + (y - y^2) \, q^2)} \displaystyle{\not}q \qquad (605)$$

which is just the divergent part of (600). But the integration in (604) actually extends over k not satisfying (601). Therefore subtracting (604) from (600) and using (605), we find for $\Lambda_c^r(p, p')$ the final result

$$\Lambda_c^r(p, p') = -2\pi^2 i \int_0^1 dy \, \frac{1}{\mu^2 + (y - y^2) \, q^2} \left\{ \left(-1 + \frac{5}{2}y - 2y^2 \right) q^2 \displaystyle{\not}q - \frac{1}{2} i \mu \displaystyle{\not}q \displaystyle{\not}q \right\}$$

$$- 4 \iint x \, dx \, dy \, \displaystyle{\not}q \int_F dk$$

$$\times \left\{ \frac{q^2 + 2\mu^2}{[k^2 + 2xk \cdot p_y]^3} - \frac{\mu^2}{[k^2 + 2xk \cdot p]^3} - \frac{\mu^2}{[k^2 + 2xk \cdot p']^3} \right\} \qquad (606)$$

This integral is completely convergent for every finite r, the k-integration being restricted to k satisfying (601). The formula (606) is exact except for terms which tend to zero with r.

To evaluate the k integral in (606) for general p and p' is possible but tedious. So we shall do it only in the case of non-relativistic velocities, when

$$|\boldsymbol{p}| \ll \mu, \qquad |\boldsymbol{p}'| \ll \mu, \qquad |\boldsymbol{q}| \ll \mu \qquad (607)$$

where $|\boldsymbol{p}|$ means $\sqrt{p_1^2 + p_x^2 + p_3^2}$, the magnitude of the space-like part of the 4-vector p. In addition to (607) we still assume r small compared to q, p, p'.

We consider the integral

$$K = \int_0^1 \int_0^1 x \, dx \, dy \int_F dk \, \frac{1}{[k^2 + 2xk \cdot p_y]^3} \qquad (608)$$

integrated over k satisfying (601), and evaluate it including terms of order $|\boldsymbol{p}|^2$, q^2, $|\boldsymbol{p}'|^2$, but neglecting higher terms. Integrating over k_0 only, we have for any positive b

$$\int_F dk_0 \frac{1}{[k^2 + 2ak_0 + b]} = i \int_{-\infty}^{\infty} \frac{dk_0}{\left[|\boldsymbol{k}|^2 + k_0^2 + 2iak_0 + b\right]} = i\pi \frac{1}{\sqrt{|\boldsymbol{k}|^2 + a^2 + b}} \tag{609}$$

Differentiating (609) twice with respect to b,

$$\int_F \frac{dk_0}{[k^2 + 2ak_0 + b]^3} = \frac{3i\pi}{8} \left\{|\boldsymbol{k}|^2 + a^2 + b\right\}^{-5/2} \tag{610}$$

Hence[80]

$$
\begin{aligned}
K &= \iint x \, dx \, dy \, \frac{3i\pi}{8} \int_{|\boldsymbol{k}|>r} d^3\boldsymbol{k} \left\{|\boldsymbol{k} + x\boldsymbol{p}_y|^2 + x^2 \left(\mu^2 + (y - y^2) q^2\right)\right\}^{-5/2} \\
&= \iint x \, dx \, dy \, \frac{3i\pi}{8} \int_{|\boldsymbol{k}|>r} d^3\boldsymbol{k} \\
&\quad \times \left\{ \left(|\boldsymbol{k}|^2 + x^2\mu^2\right)^{-5/2} - \frac{5}{2} \left(2x\boldsymbol{k} \cdot \boldsymbol{p}_y + x^2|\boldsymbol{p}_y|^2 + x^2(y - y^2) q^2\right) \right. \\
&\quad \times \left(|\boldsymbol{k}|^2 + x^2\mu^2\right)^{-7/2} + \frac{35}{8} 4x^2 (\boldsymbol{k} \cdot \boldsymbol{p}_y)^2 \left(|\boldsymbol{k}|^2 + x^2\mu^2\right)^{-9/2} \right\} \\
&= \iint x \, dx \, dy \, \frac{3i\pi^2}{2} \int_r^{\infty} k^2 \, dk \\
&\quad \times \left\{ \left(k^2 + x^2\mu^2\right)^{-5/2} - \frac{5}{2} x^2 \left(|\boldsymbol{p}_y|^2 + (y - y^2) q^2\right) \left(k^2 + x^2\mu^2\right)^{-7/2} \right. \\
&\quad \left. + \frac{35}{6} x^2 k^2 |\boldsymbol{p}_y|^2 \left(k^2 + x^2\mu^2\right)^{-9/2} \right\}
\end{aligned}
\tag{611}
$$

We can now carry out the integrations over x and y, using

$$\int_0^1 |\boldsymbol{p}_y|^2 dy = \boldsymbol{p} \cdot \boldsymbol{p}' + \frac{1}{3} q^2 \tag{612}$$

This gives

$$K = \frac{3i\pi^2}{2} \int_r^\infty k^2\, dk \left\{ \frac{1}{3\mu^2} \left(\frac{1}{k^3} - \frac{1}{(k^2 + \mu^2)^{3/2}} \right) \right.$$

$$- \frac{\boldsymbol{p}\cdot\boldsymbol{p}' + \frac{1}{2}q^2}{3\mu^4} \left(\frac{1}{k^3} - \frac{1}{(k^2 + \mu^2)^{3/2}} - \frac{1}{2\mu^2(k^2 + \mu^2)^{5/2}} \right)$$

$$\left. + \frac{1}{6}k^2 \left[\boldsymbol{p}\cdot\boldsymbol{p}' + \frac{1}{3}q^2 \right] \frac{2}{\mu^4} \left(\frac{1}{k^5} - \frac{1}{(k^2 + \mu^2)^{5/2}} - \frac{5}{\mu^2}\frac{1}{(k^2 + \mu^2)^{7/2}} \right) \right\} \tag{613}$$

The k-integration is now elementary, and after dropping terms which tend to zero with r we have

$$K = \frac{3i\pi^2}{2\mu^2} \left\{ \frac{1}{3} \left(\log\frac{\mu}{2r} + 1 \right) - \frac{\boldsymbol{p}\cdot\boldsymbol{p}' + \frac{1}{2}q^2}{\mu^2} \left(\frac{1}{3}\log\frac{\mu}{2r} + \frac{1}{6} \right) \right.$$

$$\left. + \frac{1}{6}\frac{\boldsymbol{p}\cdot\boldsymbol{p}' + \frac{1}{3}q^2}{\mu^2} \left(2\log\frac{\mu}{2r} + \frac{5}{3} \right) \right\} \tag{614}$$

Putting $p = p'$, $q = 0$ in (614), we find

$$K_0 = \int_0^1 \int_0^1 x\, dx\, dy \int_F dk\, \frac{1}{[k^2 + 2xk\cdot p]^3}$$

$$= \frac{3i\pi^2}{2\mu^2} \left\{ \frac{1}{3} \left(\log\frac{\mu}{2r} + 1 \right) + \frac{1}{9}\frac{|\boldsymbol{p}|^2}{\mu^2} \right\} \tag{615}$$

Substituting (614) and (615) into (606) and dropping terms of higher order than $|\boldsymbol{p}|^2$, $|\boldsymbol{p}'|^2$ and q^2,

$$\Lambda_c^r(p, p') = -2\pi^2 i \left\{ -\frac{5}{12}\frac{q^2}{\mu^2}\slashed{q} - \frac{1}{2}\frac{i}{\mu}\slashed{q}\slashed{q} \right\}$$

$$- 4\slashed{q}\frac{3i\pi^2}{2\mu^2} \left\{ \frac{1}{3}q^2 \left(\log\frac{\mu}{2r} + 1 \right) - \frac{1}{18}q^2 \left(2\log\frac{\mu}{2r} + \frac{5}{3} \right) \right\}$$

$$= -\frac{4}{3}\pi^2 i\frac{q^2}{\mu^2}\slashed{q} \left\{ \log\frac{\mu}{2r} + \frac{11}{24} \right\} - \frac{\pi^2}{\mu}\slashed{q}\slashed{q} \tag{616}$$

Using (596) and adding the contribution (554) from G_1, we find for the second-order terms in (541) the value

$$M_2 + M_2' = -\frac{\alpha}{3\pi} \left\{ \log\frac{\mu}{2r} + \frac{11}{24} - \frac{1}{5} \right\} \frac{q^2}{\mu^2} M_0 + \frac{\alpha}{4\pi}\frac{ie}{mc^2} \left(\overline{u}'\slashed{q}\slashed{q}u \right) \tag{617}$$

7.6 The Treatment of Low-Frequency Photons. The Infra-Red Catastrophe

The second-order correction (617) to the scattering matrix element M_0 has been made convergent by only taking into account the effects of photons with energy greater than ΔE according to (603). When $\Delta E \to 0$ the correction diverges logarithmically, and this divergence must now be interpreted.

In the approximation where $|\boldsymbol{p}|^2$ and $|\boldsymbol{p}'|^2$ are small compared with μ^2, (546) gives

$$M_1 = \frac{e}{\hbar c \mu |\boldsymbol{k}'|} (q \cdot e') M_0 \tag{618}$$

The total probability that an electron is scattered between the initial and final states (542), with the emission of a photon with potentials (440), summed over all photons with frequencies lying in the range[81]

$$r_1 < |\boldsymbol{k}'| < r_2 \tag{619}$$

is therefore

$$W_R(r_1, r_2) = \int d^3 \boldsymbol{k}' \sum_{e'} \frac{1}{(2\pi)^3} \left(\frac{\hbar c}{2|\boldsymbol{k}'|} \right) |M_1|^2$$

$$= \frac{e^2}{16\pi^3 \hbar c \mu^2} |M_0|^2 \int d^3 \boldsymbol{k}' \frac{1}{|\boldsymbol{k}'|^3} \sum_{e'} |q \cdot e'|^2$$

$$= \frac{\alpha}{\pi \mu^2} |M_0|^2 \int_{r_1}^{r_2} \frac{dk'}{k'} \frac{2}{3} q^2$$

$$= \frac{2\alpha}{3\pi} \left(\log \frac{r_2}{r_1} \right) \frac{q^2}{\mu^2} |M_0|^2 \tag{620}$$

In (620) it is assumed that both r_2 and r_1 are frequencies small compared with $|\boldsymbol{q}|$.

On the other hand, the probability that an electron is scattered between the states (542) without emitting a photon is given by

$$W_N = |M_0 + M_2 + M_2'|^2$$

$$= |M_0|^2 + M_0^* \left(M_2 + M_2' \right) + \left(M_2 + M_2' \right)^* M_0 \tag{621}$$

neglecting terms of fourth order in the radiation interaction. If we consider in (621) the contribution only from virtual photons with frequencies in the

range (619), then by (617) we have

$$M_2 + M_2' = -\frac{\alpha}{3\pi}\left(\log\frac{r_2}{r_1}\right)\frac{q^2}{\mu^2}\,|M_0|^2$$

$$W_N(r_1, r_2) = |M_0|^2 - \frac{2\alpha}{3\pi}\left(\log\frac{r_2}{r_1}\right)\frac{q^2}{\mu^2}\,|M_0|^2 \tag{622}$$

So the contributions to (617) from low-frequency virtual photons serve only to compensate exactly the probability for scattering with emission of low-frequency real photons. The non-radiative probability is decreased by the effects of low-frequency virtual photons, so that the total scattering probability, radiative plus non-radiative, is essentially independent of the presence of very low-frequency photons. The total scattering probability is thus a finite quantity free from any infra-red divergence.

To describe correctly the radiative corrections to scattering it is essential to specify the critical energy ΔE below which real photons will not be detected. Ideally we suppose that in each scattering event a photon of energy greater than ΔE is detected with 100% efficiency, a photon of energy less than ΔE with zero efficiency. Then the total observed probability for non-radiative scattering is given by (621), with $M_2 + M_2'$ given by (617) and

$$\log\frac{\mu}{2r} = \log\left(\frac{mc^2}{2\Delta E}\right) \tag{623}$$

This probability (621) will include the scatterings in which a photon is emitted with energy below the limit of detection. The formula (617) is valid so long as

$$r \ll |\boldsymbol{p}|, |\boldsymbol{p}'|, |\boldsymbol{q}| \ll \mu \tag{624}$$

The probability for radiative scattering (i.e. scattering with emission of a detectable photon) is given by (546).

It can be proved that this removal of the infra-red divergence by taking into account the existence of unobservable photons is quite general, and works equally well when q is not small. Only then the evaluation of (608) is considerably more unpleasant. Also, the same argument removes all infra-red divergences, also when higher-order radiative corrections are considered, in which case we are concerned with the effects of the emission of two or more soft photons during the scattering process. For a general discussion of this question see Bloch and Nordsieck, *Phys. Rev.* **52** (1937) 54.

Scattering by a Static Potential. Comparison with Experimental Results

Consider the scattering of an electron by a time-independent electrostatic potential

$$V(r) = \frac{1}{(2\pi)^3} \int d^3\boldsymbol{q}\, V(\boldsymbol{q})\, e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \tag{625}$$

Then (543) gives the matrix element for scattering without radiative corrections (compare (625) and (534))

$$M_0 = 2\pi i \frac{e}{\hbar c} (u'^* u) V(\boldsymbol{q})\, \delta(q_0) \tag{626}$$

Non-radiative scattering occurs only between states for which

$$q_0 = 0\,, \qquad |\boldsymbol{p}| = |\boldsymbol{p}'| \tag{627}$$

The cross-section for scattering between the states (542), per element of solid angle $d\Omega$ in the direction of p', is without radiative corrections

$$\sigma_0 = \left(\frac{em}{2\pi\hbar^2}\right)^2 |u'^* u|^2 \, |V(\boldsymbol{q})|^2 d\Omega \tag{628}$$

This follows immediately from (626) using (627), using the prescription derived on page 29, when we treated the Møller scattering the first time. The procedure is as follows:[82]

$$w_S = \frac{c|M_0|^2}{2\pi\,\delta(q_0)} = 2\pi \frac{e^2}{\hbar^2 c} |u'^* u|^2 \, |V(\boldsymbol{q})|^2 \, \delta(q_0)$$

$$\rho\,dE = \frac{mc^2}{E}\frac{d^3\boldsymbol{p}}{(2\pi)^3} \qquad E\,dE = \hbar^2 c^2 p\,dp \qquad d^3\boldsymbol{p} = p^2 dp\,d\Omega$$

$$\rho = \frac{mc^2}{E}\frac{p^2}{(2\pi)^3}\frac{dp}{dE}d\Omega = \frac{mp}{\hbar^2(2\pi)^3}\,d\Omega \qquad \delta(q_0) = \hbar c\,\delta(E)$$

$$w = \frac{2\pi e^2}{\hbar^2 c}\,\hbar c\,\frac{mp}{\hbar^2(2\pi)^3}\,d\Omega\,|u'^*u|^2\,|V(\boldsymbol{q})|^2 = \frac{e^2 mp}{(2\pi)^2\hbar^3}\,d\Omega\,|u'^*u|^2\,|V(\boldsymbol{q})|^2$$

$$\sigma = \frac{wV}{v} \qquad V = \frac{mc^2}{E} \qquad v = \frac{c^2\hbar p}{E}$$

$$\sigma = \frac{e^2 mp}{(2\pi)^2\hbar^3}\frac{mc^2}{E}\frac{E}{c^2\hbar p}\,d\Omega\,|u'^*u|^2|V(\boldsymbol{q})|^2 = \left(\frac{em}{2\pi\hbar^2}\right)^2\,d\Omega\,|u'^*u|^2\,|V(\boldsymbol{q})|^2$$

Summing over the final spin states, and averaging over the initial spin states, we have

$$\frac{1}{2}\sum_u\sum_{u'}|u'^*u|^2 = \frac{1}{2}\frac{1}{(2i\mu)^2}\,\mathrm{Tr}\left\{(\not{p}+i\mu)\gamma_4(\not{p}'+i\mu)\gamma_4\right\}$$

$$= \frac{1}{2\mu^2}\left\{\mu^2 + p_0 p_0' + \boldsymbol{p}\cdot\boldsymbol{p}'\right\} = \frac{1}{2\mu^2}\left\{2p_0^2 - \frac{1}{2}|\boldsymbol{q}|^2\right\}$$

$$= \frac{p_o^2}{\mu^2}\left\{1 - \frac{1}{4}\frac{|\boldsymbol{p}|^2 - 2\boldsymbol{p}\cdot\boldsymbol{p}' + |\boldsymbol{p}'|^2}{p_0^2}\right\} = \frac{p_o^2}{\mu^2}\left(1 - \beta^2\sin^2\frac{\theta}{2}\right)$$

$$\tag{629}$$

where θ is the angle between p and p', and

$$\beta = \frac{|\boldsymbol{p}|}{p_0} = \frac{v}{c} \tag{630}$$

where v is the velocity of the incident electron. Hence the cross-section for an unpolarized electron beam is

$$\overline{\sigma}_0 = \left(\frac{eE}{2\pi\hbar^2 c^2}\right)^2\left(1 - \beta^2\sin^2\frac{\theta}{2}\right)|V(\boldsymbol{q})|^2\,d\Omega \tag{631}$$

where E is the energy of the incident electron.

The second-order radiative correction to M_0 is given by (617), which in this case becomes

$$M_2 + M_2' = -\frac{\alpha}{3\pi}\left\{\log\frac{\mu}{2r} + \frac{11}{24} - \frac{1}{5}\right\}\frac{q^2}{\mu^2}\,M_0 - \frac{\alpha}{2}\frac{e}{\hbar c\mu}\,(u'^*\not{q}u)V(\boldsymbol{q})\,\delta(q_0) \tag{632}$$

This gives a second order correction to the cross-section σ_0 according to (621). The total non-radiative cross-section, for scattering without the emission of a photon of energy greater than ΔE, becomes

$$\sigma_N = \sigma_0 + \sigma_{2N} = \left(\frac{em}{2\pi\hbar^2}\right)^2 |V(\boldsymbol{q})|^2 \, d\Omega$$

$$\times \left| \left\{ 1 - \frac{\alpha}{3\pi} \left(\log \frac{\mu}{2r} + \frac{11}{24} - \frac{1}{5} \right) \frac{q^2}{\mu^2} \right\} (u'^*u) + \frac{i\alpha}{4\pi\mu} (u'^* \not q u) \right|^2 \quad (633)$$

Summing and averaging over the spin-states

$$\frac{1}{2} \sum_u \sum_{u'} (u'^* \not q u)(u^*u') = \frac{1}{2(2i\mu)^2} \, \mathrm{Tr}\left\{ (\not p + i\mu)\gamma_4(\not p' + i\mu)\gamma_4 \not q \right\}$$

$$= -\frac{1}{8\mu^2} \, \mathrm{Tr}\left\{ i\mu \left(\not p \gamma_4 \gamma_4 \not q + \gamma_4 \not p' \gamma_4 \not q \right) \right\}$$

$$= -\frac{1}{8\mu^2} \, \mathrm{Tr}\left\{ i\mu(\not p \not p' - \not p \not p - \not p' \not p' + \not p' \not p) \right\}$$

$$= -\frac{1}{8\mu^2} \, \mathrm{Tr}\left\{ i\mu(\not p' - \not p)(\not p' - \not p) \right\} = \frac{i}{2\mu} \, q^2 \quad (634)$$

Hence for an unpolarized electron beam the non-radiative cross-section is

$$\overline{\sigma}_N = \left(1 - \frac{2\alpha}{3\pi} \left(\log \frac{\mu}{2r} + \frac{11}{24} - \frac{1}{5} \right) \frac{q^2}{\mu^2} \right) \overline{\sigma}_0 - \left(\frac{em}{2\pi\hbar^2} \right)^2 |V(\boldsymbol{q})|^2 d\Omega \, \frac{\alpha}{4\pi} \frac{q^2}{\mu^2} \tag{635}$$

Since we are working only to order q^2 in the radiative corrections, the second term in (635) may be replaced by

$$-\frac{\alpha}{4\pi} \frac{q^2}{\mu^2} \overline{\sigma}_0 \tag{636}$$

and then we find

$$\sigma_N = \left(1 - \frac{2\alpha}{3\pi} \left(\log \frac{mc^2}{2\Delta E} + \frac{5}{6} - \frac{1}{5} \right) \frac{q^2}{\mu^2} \right) \overline{\sigma}_0 \tag{637}$$

Formulae (628) and (631) are exact for electrons of any energy; while (632) and (637) are valid only for slow electrons, terms of order higher than αq^2 being neglected.

To remove the dependence of (637) on r we must consider the cross-section for scattering with emission of a photon of frequency greater than

r. Since the electron is now considered to be slow, the maximum possible energy of the photon is

$$\hbar c k_{\max} = E - mc^2 \approx \frac{\hbar^2}{2m}|\boldsymbol{p}|^2 \tag{638}$$

and therefore for all possible photons the momentum $\hbar|\boldsymbol{k'}|$ will be very small compared with the electron momentum $\hbar|\boldsymbol{p}|$.[83] Thus the recoil of the electron, arising from the momentum carried off by the photon, can always be neglected. The matrix element and probability for radiative scattering will be given by (618) and (620), even when the photon takes away a large fraction of the kinetic energy of the electron.

We imagine a scattering experiment in which only the direction of the emerging electron is measured and not its energy. Then the radiative cross-section $\overline{\sigma}_R$ will measure the total probability of scattering the electron into a solid angle $d\Omega$ with emission of a photon having any frequency between the lower limit r and the upper limit $k_{\max}$ given by (638). The observed cross-section will be

$$\sigma_T = \overline{\sigma}_N + \overline{\sigma}_R \tag{639}$$

with the same low-frequency cut-off r in both $\overline{\sigma}_N$ and $\overline{\sigma}_R$. Thus σ_T gives the cross-section for scattering into a given solid angle $d\Omega$ with or without photon emission. Being directly observed, σ_T must be divergence-free and independent of r.

In the radiative scattering process we may take the final momentum of the electron to be $\lambda\hbar p'$ where $0 < \lambda < 1$ and p' satisfies (627). Instead of (627) the conservation of energy now gives by (638)

$$\hbar|\boldsymbol{p}|^2(1 - \lambda^2) = 2mc|\boldsymbol{k'}| \tag{640}$$

According to (620), the probability for scattering the electron into a state $\lambda\hbar p'$, with emission of a photon in any direction with frequency in the range $(k', k' + dk')$ is

$$w_R(k') = \frac{2\alpha}{3\pi} \frac{dk'}{k'} \frac{|\boldsymbol{p} - \lambda\boldsymbol{p'}|^2}{\mu^2} |M_0|^2 \tag{641}$$

This corresponds to a differential cross-section

$$\sigma_R(k') = \frac{2\alpha}{3\pi} \frac{dk'}{k'} \frac{|\boldsymbol{p} - \lambda\boldsymbol{p'}|^2}{\mu^2} \lambda \left(\frac{eE}{2\pi\hbar^2 c^2}\right)^2 |V(p - \lambda p')|^2 \, d\Omega \tag{642}$$

for scattering into the solid angle $d\Omega$, neglecting now the term in β which appeared in (631) since (642) is itself of order $\alpha\beta^2$ and higher terms are neglected. The factor λ comes in from $\frac{p_{\text{final}}}{p'}$. According to (640)

$$\frac{dk'}{k'} = -\frac{2\lambda\,d\lambda}{1-\lambda^2} \tag{643}$$

Hence the radiative cross-section integrated over the quantum frequency is

$$\overline{\sigma}_R = \frac{2\alpha}{3\pi} \int_0^{\lambda_m} \frac{2\lambda^2\,d\lambda}{1-\lambda^2} \frac{|\boldsymbol{p}-\lambda\boldsymbol{p}'|^2}{\mu^2} \left(\frac{eE}{2\pi\hbar^2 c^2}\right)^2 |V(\boldsymbol{p}-\lambda\boldsymbol{p}')|^2\,d\Omega \tag{644}$$

where by (640) and (638)

$$\lambda_m = \sqrt{1 - \frac{r}{k_{\max}}} = \sqrt{1 - \frac{\Delta E}{T}} \tag{645}$$

T is the initial kinetic energy of the electron given by (638).

Now we can combine (637) and (644) to give by (639) and (629)

$$\sigma_T = \left(1 - \frac{2\alpha}{3\pi}\left(\log\frac{mc^2}{2T} + \frac{5}{6} - \frac{1}{5}\right) 4\beta^2\sin^2\frac{\theta}{2}\right)\overline{\sigma}_0 + \frac{2\alpha}{3\pi}\left(\frac{e}{2\pi\hbar c}\right)^2 d\Omega$$

$$\times \int_0^1 \frac{2\lambda\,d\lambda}{1-\lambda^2}\left\{\lambda\,|\boldsymbol{p}-\lambda\boldsymbol{p}'|^2\,|V(\boldsymbol{p}-\lambda\boldsymbol{p}')|^2 - |\boldsymbol{p}-\boldsymbol{p}'|^2\,|V(\boldsymbol{p}-\boldsymbol{p}')|^2\right\} \tag{646}$$

Here we used the following trick: the integral over λ blows up at $\lambda = 1$. Therefore we subtract from the numerator its value at $\lambda = 1$, which makes the integral behave decently and permits us to change the upper limit from λ_m to 1, for small ΔE's. We also have to add then the integrand with the numerator having $\lambda = 1$; this gives then a logarithmic term which combines with (637) to give the first part of (646).

Formula (646) gives a result of the form

$$\sigma_T = \left(1 - \frac{8\alpha}{3\pi}\beta^2\sin^2\frac{\theta}{2}\left\{\log\frac{mc^2}{2T} + f(\theta)\right\}\right)\overline{\sigma}_0 \tag{647}$$

where for low velocities $f(\theta)$ is independent of T and of order 1 compared to the logarithm. For any special potential $f(\theta)$ can be calculated.

From (647) we see that the observable radiative correction is not of order α but of order

$$\alpha\left(\frac{v}{c}\right)^2\log\left(\frac{c}{v}\right) \tag{648}$$

which is much smaller if v is not relativistic. Thus the correction *cannot be observed at all* in a non-relativistic scattering experiment. In the relativistic region the effect is actually of order α as indicated by (647), but the correct formula is then enormously more complicated.

The exact formulae in both non-relativistic and relativistic cases have been published by J. Schwinger, *Phys. Rev.* **76** (1949) 790.

An experimental test in the relativistic range is just on the limits of the possible. See Lyman, Hanson and Scott, *Phys. Rev.* **84** (1951) 626. Scattering of 15 MeV electrons by nuclei were measured with very good energy-resolution, $\Delta E/E$ being 1–3%. In this case it was *only* the non-radiative cross-section $\overline{\sigma}_N$ that was observed, and so the radiative correction given by the relativistic form of (637) becomes quite large. In fact the radiative correction in $\overline{\sigma}_N$ in the relativistic range becomes of the order

$$\alpha \left\{ \log \frac{\Delta E}{E} \right\} \left\{ \log \frac{E}{mc^2} \right\} \tag{649}$$

according to Schwinger, *Phys. Rev.* **76**, 813, Eq. (2.105) (with misprint K for k, in my notation μ.) In the conditions of the experiment of Lyman-Hanson-Scott (649) becomes of the order 5% and was clearly observed, the experimental errors being $\sim 2\%$. However, (649) arises mainly from the low-energy virtual photons, with energies going down to ΔE. What is observed is just the decrease in the non-radiative cross-section due to the competition of radiative scattering with energy-loss in the range $[\Delta E, E]$. Thus the measurement of (649) by LHS is actually only a very inaccurate measurement of the cross-section for bremsstrahlung, which could be observed much more accurately by observing the photons which are really emitted.

The theoretically interesting part of the radiative corrections is the part which is not just the effect of real bremsstrahlung. This part is given by the terms in $\overline{\sigma}_N$ which are of order 1 compared with the $\log(\Delta E/E)$ appearing in (649). For example, we would have to observe with sufficient accuracy to see the terms $(\frac{5}{6} - \frac{1}{5})$ in (637) if we wished to verify the theoretical radiative corrections at low velocities. In the relativistic range the "genuine" radiative corrections are of order

$$\alpha \log \left(\frac{E}{mc^2} \right) \tag{650}$$

instead of (649), i.e. about 2% in the LHS experiment. To detect such effects is already possible, but to observe them accurately in a scattering experiment seems not very hopeful.

This is all we can say at present about radiative corrections to scattering by an electrostatic potential.

8.1   The Magnetic Moment of the Electron

The scattering by an electrostatic potential, the two terms in (617) were lumped together. Both gave contributions of the same order of magnitude, $\alpha(q^2/\mu^2)$ in the cross section. What then is the meaning of the special form of the second term in (617)? This term has no infra-red divergence and therefore should be particularly simple to interpret experimentally.

Consider scattering of a slow electron by a slowly-varying[84] magnetic field. The potentials (534) can then be taken to be a pure vector potential, so that

$$e_4(q) = 0 \tag{651}$$

The matrix elements of $\gamma_1, \gamma_2, \gamma_3$ between positive energy electron states are of the order (v/c). Hence M_0 given by (543) is of order (v/c). The first term in (617) is thus of the order $\alpha(v/c)^3$ while the second is of order $\alpha(v/c)$. Therefore the second term in (617) is the main term in considering magnetic effects, and the first term can be ignored. The meaning of the second term must be a change in the *magnetic properties* of a non-relativistic electron.

As we saw in discussing the Dirac equation, (Eqs. (99) and (100)), an electron by virtue of its charge $(-e)$ behaves in non-relativistic approximation as if it had a magnetic moment

$$M = -\frac{e\hbar}{2mc} \tag{652}$$

This moment has an energy of interaction with an external Maxwell field $(\boldsymbol{E}, \boldsymbol{H})$ given by

$$H_M = -M(\boldsymbol{\sigma} \cdot \boldsymbol{H} - i\boldsymbol{\alpha} \cdot \boldsymbol{E}) \tag{653}$$

the term which appears in the non-relativistic Schrödinger equation (100).

Now suppose that the electron possesses an additional magnetic moment δM which does not arise from its charge. Such an additional moment is called "anomalous". To give the electron an anomalous moment, we only need to add arbitrarily a term proportional to (653) to the Hamiltonian. Comparing (654) with (97) and (98), we see that (653) is a relativistic invariant and can be written

$$H_M = \frac{1}{2}iM \sum_\mu \sum_\nu \sigma_{\mu\nu} F_{\mu\nu} \tag{654}$$

Hence an anomalous magnetic moment δM will be given to the electron if the term

$$L_M = -\frac{1}{2}i\delta M \sum_\mu \sum_\nu \sigma_{\mu\nu} F_{\mu\nu} \tag{655}$$

is added to the Lagrangian. This refers still to the one-electron Dirac equation. In the theory of the quantized Dirac field, the corresponding addition to the Lagrangian density (410) is

$$\mathscr{L}_M = -\frac{1}{2}i\delta M \overline{\psi} \sum_\mu \sum_\nu \sigma_{\mu\nu} \psi F_{\mu\nu}^e \tag{656}$$

where it is supposed that the anomalous moment is interacting with the external Maxwell field. The addition (656) to the Lagrangian[85] gives a relativistically invariant description of an anomalous moment.

Consider the effect of (656) on the scattering of an electron by the potentials (534). Treating the scattering in Born approximation and using (420), the contribution of (656) to the scattering matrix element is

$$U_M = \sum_{\mu,\nu} \frac{\delta M}{2\hbar c} \int \overline{\psi}(x)\sigma_{\mu\nu}\psi(x)F_{\mu\nu}^e(x)\,dx \tag{657}$$

the integral being over all space-time. Using (542) for the initial and final electron wave-functions, and defining q, e by (544) and (545), this matrix element becomes

$$U_M = i\frac{\delta M}{2\hbar c} \sum_{\mu,\nu} (\overline{u}'\sigma_{\mu\nu}u)(q_\mu e_\nu - q_\nu e_\mu) = i\frac{\delta M}{2\hbar c}\left[\overline{u}'(\not{q}\not{e} - \not{e}\not{q})u\right] \tag{658}$$

where we used $\gamma_k\gamma_\ell = i\sigma_m$, $k,\ell,m = (1,2,3)$ cyclically permuted. Since we have also assumed (588) we may write simply

$$U_M = -i\frac{\delta M}{\hbar c}(\overline{u}'\not{e}\not{q}u) \tag{659}$$

Now comparing the matrix element (659) with (617), we see that the magnetic effect of the second-order radiative correction to scattering is exactly described by saying that the electron has an anomalous magnetic moment δM given by

$$\delta M = -\frac{\alpha}{4\pi}\frac{e\hbar}{mc} = +\frac{\alpha}{2\pi}M \tag{660}$$

This is the famous Schwinger correction to the electron magnetic moment, which we have now calculated. Not only for scattering but for all phenomena

in the non-relativistic range, the magnetic part of the second-order radiative correction to the motion of an electron is equivalent simply to the anomalous magnetic moment (660).

This anomalous moment has been extremely accurately confirmed experimentally by[86] Kusch, Prodell and Koenig (*Phys. Rev.* **83** (1951) 687) who find

$$\frac{\delta M}{M} = 0.001145 \pm 0.000013$$

The calculated value including a fourth-order α^2 correction found by Karplus and Kroll (*Phys. Rev.* **77** (1950) 536) is

$$\frac{\delta M}{M} = \frac{\alpha}{2\pi} - 2.973 \left(\frac{\alpha^2}{\pi^2} \right) = 0.0011454$$

8.2   Relativistic Calculation of the Lamb Shift

To make a correct relativistic calculation of the Lamb shift we have to repeat the treatment of line-shifts and line-widths which we gave earlier, only now using the relativistic theory for the atom. So we should set up the equation of motion of atom plus radiation field in the *Bound Interaction Representation.* The equation of motion is then given by (245), (247), only with the j_μ operator now describing the system of a relativistic atom. The solution of (245) can then be found as in the non-relativistic case, using the known wave-functions of the stationary states of the atom. In this way the Lamb shift was actually calculated by Lamb and Kroll[87], *Phys. Rev.* **75** (1949) 388. However, in their calculations Lamb and Kroll had troubles with the subtraction of the divergent mass-renormalization effect. Because the calculation was all in terms of the atomic wave-functions, it was not possible to keep using relativistically invariant notations all the way through. Hence the mass term could not be clearly separated from the remaining finite terms by its dependence on the particle momentum p, as it was separated for example in Eq. (566) during the calculation of radiative corrections to scattering. The final result of Lamb and Kroll was uncertain because of this difficulty in the mass separation. They obtained the correct answer 1052 Mc but only by making use of the experimentally *measured* value $\alpha/2\pi$ of the electron anomalous magnetic moment.[88]

Learning from the calculation of the radiative corrections to scattering, we see that to make a clear separation of the mass renormalization from observable effects, we must arrange the calculations so that the separation

is done for a particle of momentum p in a *variable* Lorentz system. Then by varying the Lorentz system we can vary p and identify the mass term unambiguously as the expression which has the correct dependence on p. To work in a variable Lorentz system, we must work in a representation which is independent of the Lorentz system, so that the calculations are formally invariant. The only convenient invariant representation is the *Free Interaction Representation*.

Hence our program is to set up the equation of motion first in the *Free Interaction Representation*, then carry through a relativistically invariant calculation to identify and cancel the divergent renormalization effects unambiguously. After this we must transform to the *Bound Interaction Representation* for the final calculation of the line-shift. This two-stage calculation is absolutely necessary in order to get the right answers. It was Schwinger's invention.

In the FIR the equation of motion is

$$i\hbar\frac{\partial\Psi}{\partial t} = \left(H^e(t) + H^I(t)\right)\Psi \tag{661}$$

with H^I given by (532) including the mass-renormalization term H_S. The first stage in the solution of (661) is to write

$$\Psi(t) = \Omega_1 \Phi(t) \tag{662}$$

where $\Omega_1(t)$ is defined as in (492), only with $e\overline{\psi}A\psi$ replaced by $[e\overline{\psi}A\psi + i\,\delta m\,c^2\overline{\psi}\psi]$, and the function $g_A(t)$ is as before supposed to tend to the limiting form $g_A(t) \to 1$ at the end of the calculation. The operator $\Omega_1(t)$ satisfies

$$i\hbar\frac{\partial\,\Omega_1(t)}{\partial t} = H^I(t)\,\Omega_1(t) \tag{663}$$

for any value of t not in the remote past, so that we may take $g_A(t) = 1$. Hence the equation of motion for $\Phi(t)$ derived from (661) is just

$$i\hbar\frac{\partial\Phi}{\partial t} = H_T(t)\Phi \tag{664}$$

$$H_T(t) = (\Omega_1(t))^{-1}\,H^e(t)\,\Omega_1(t) \tag{665}$$

The covariant part of the calculation, which is done in the FIR, is just the evaluation of this transformed Hamiltonian $H_T(t)$.

8.2.1   *Covariant Part of the Calculation*

Let us write

$$H_F(t) = \Omega_2(t)\, H^e(t)\, \Omega_1(t) \tag{666}$$

with Ω_2 given by (490). Then

$$H_T(t) = (\Omega_2(t)\Omega_1(t))^{-1}\, H_F(t) = S^{-1} H_F(t) \tag{667}$$

where S, given by (495), is the scattering matrix defined in the absence of the external potential A^e_μ. We shall now restrict our attention to systems in which only one electron is actually present. We saw by Eq. (533) that S applied to a one-electron state is equivalent to the unit operator, i. e.g. S produces no real scattering or shift of phase in one-particle states. Therefore in discussing the hydrogen atom we may simply omit the S^{-1} in (667) and write

$$H_T(t) = H_F(t) \tag{668}$$

Now $H_F(t)$ is just the term involving $H^e(t)$ in the series expansion (421). In fact we have, by direct multiplication of the series for Ω_1 and Ω_2,

$$H_F(t) = \sum_{n=0}^{\infty} \left(\frac{-i}{\hbar}\right)^n \frac{1}{n!} \int \ldots \int dt_1\, dt_2\, \ldots\, dt_n$$

$$\times P\left\{ H^e(t), H^I(t_1), \ldots, H^I(t_n) \right\} \tag{669}$$

the damping functions $g(t_i)$ being always understood when they are not written explicitly. Now (667) is directly related to the operator U given by (535), namely

$$U = -\frac{i}{\hbar} \int_{-\infty}^{\infty} H_F(t)\, dt \tag{670}$$

We write

$$H_T(t) = H^e(t) + H_{T1}(t) + H_{T2}(t) \tag{671}$$

expanding H_T in powers of the radiation interaction just as U was expanded in (536)–(540). But the matrix elements of $\left(U_2 + U_2^1\right)$ referring to one-electron transition have already been calculated and are given by (617) assuming non-relativistic velocities for the electron. This enables us to write down immediately a formula for the operator H_{T2}, valid for one-electron

transitions at non-relativistic velocities. In (617) each factor q_λ may be replaced by $(-i\partial/\partial x_\lambda)$ operating on the potentials (534). Then (617) becomes

$$U_2 + U_2' = \frac{1}{\mu^2}\frac{\alpha}{3\pi}\left\{\log\frac{\mu}{2r} + \frac{11}{24} - \frac{1}{5}\right\}\left(\frac{e}{\hbar c}\right)\int \overline{\psi}(\square^2 A\!\!\!/)\psi(x)\,dx$$

$$+ \frac{\alpha}{4\pi}\frac{e}{mc^2}\int \overline{\psi}\sum_\lambda \frac{\partial A\!\!\!/}{\partial x_\lambda}\gamma_\lambda\psi(x)\,dx \tag{672}$$

And using (670) and (668) we deduce

$$H_{T2} = \frac{ie}{\mu^2}\frac{\alpha}{3\pi}\left\{\log\frac{\mu}{2r} + \frac{11}{24} - \frac{1}{5}\right\}\int \overline{\psi}(\square^2 A\!\!\!/)\psi(x)\,d^3\boldsymbol{x}$$

$$+ \frac{\alpha}{4\pi}\frac{ie\hbar}{mc}\int \overline{\psi}\sum_\lambda \frac{\partial A\!\!\!/}{\partial x_\lambda}\gamma_\lambda\psi(x)\,d^3\boldsymbol{x} \tag{673}$$

Specializing now to the case of a time-independent electrostatic potential given by

$$A_4 = i\varphi(r)\,, \qquad V = -e\varphi, \tag{674}$$

$$H^e(t) = \int V(r)(\psi^*\psi)\,d^3\boldsymbol{r}, \tag{675}$$

we find[89]

$$H_{T2} = \frac{\alpha}{3\pi\mu^2}\left\{\log\frac{\mu}{2r} + \frac{11}{24} - \frac{1}{5}\right\}\int (\nabla^2 V)(\psi^*\psi)d^3\boldsymbol{r}$$

$$- \frac{i\alpha}{4\pi\mu}\int \psi^*\gamma_4(\boldsymbol{\alpha}\cdot\nabla V)\psi\,d^3\boldsymbol{r} \tag{676}$$

The calculation of H_{T2}, which is the main part of the Lamb shift calculation, can thus be taken over directly from the scattering calculation. In particular, the elimination of divergent renormalization effects does not have to be considered afresh. Once (676) has been derived, everything is finite and we are permitted to carry out the rest of the calculation in a non-covariant way. Note however that the low-energy photon frequency cut-off r still appears in (676). We expect that this dependence on r will finally disappear when the effects of H_{T1} are considered, the same cut-off being used in both H_{T1} and H_{T2}.

We next turn to the evaluation of H_{T1}, which is related to U_1 in the same way as H_{T2} to $(U_2 + U_2^1)$. According to (484) the matrix element of

U_1 for a one-electron transition between the states (542), with emission of the photon (440), is [90]

$$M_1 = -\frac{e^2}{\hbar^2 c^2}\overline{u}' \left\{ \cancel{e}\frac{1}{\cancel{p} - \cancel{k}' - i\mu}\cancel{e}' + \cancel{e}'\frac{1}{\cancel{p}' + \cancel{k}' - i\mu}\cancel{e} \right\} u \qquad (676a)$$

where now

$$e_\mu = e_\mu(p' + k' - p) \qquad (677)$$

is given by the Fourier expansion (534). Note that we cannot yet use the simple form (546) for M_1, because we do not know that $k' \ll p, q$ for the photons which will be important in this problem. Using the Dirac equation satisfied by u and u', we may write without approximations using (587) and rule 4 on page 87,

$$M_1 = -\frac{e^2}{2\hbar^2 c^2}\overline{u}' \left\{ \frac{\cancel{e}\cancel{k}'\cancel{e}' - 2(p \cdot e')\cancel{e}}{p \cdot k'} + \frac{\cancel{e}'\cancel{k}'\cancel{e} + 2(p' \cdot e')\cancel{e}}{p' \cdot k'} \right\} u \qquad (677a)$$

Since p and p' are assumed non-relativistic we may write

$$p \cdot k' = p' \cdot k' = -\mu k'_0$$

and then

$$M_1 = \frac{e^2}{2\hbar^2 c^2 \mu k'_0}\overline{u}' \left\{ 2\left((p'-p)\cdot e'\right)\cancel{e} + \cancel{e}'\cancel{k}'\cancel{e} + \cancel{e}\cancel{k}'\cancel{e}' \right\} u \qquad (678)$$

Now because we shall be considering only an electrostatic potential (674), $\cancel{e}$ is a multiple of γ_4 simply. Then if k'_3 is the space-like part of the vector k', we have

$$\cancel{e}'\cancel{k}'_3\cancel{e} + \cancel{e}\cancel{k}'_3\cancel{e}' = \cancel{e}\left\{\cancel{e}'\cancel{k}'_3 + \cancel{k}'_3\cancel{e}'\right\} = 0$$

Therefore

$$\cancel{e}'\cancel{k}'_3\cancel{e} + \cancel{e}\cancel{k}'_3\cancel{e}' = ik'_0\cancel{e}'\,(2\gamma_4\cancel{e}) \qquad (679)$$

Now this term (679) is small compared to the other term in (678), because $\cancel{e}'$ involves the matrices $\gamma_1, \gamma_2, \gamma_3$ while $\cancel{e}$ involves γ_4, and the matrix elements of $\gamma_1, \gamma_2, \gamma_3$ for non-relativistic transitions are small, of the order of (v/c). The term (679) in fact describes magnetic radiation, whereas the other term in (678) gives electric. The electric term gives an effect of the order of the Lamb shift. Hence in our approximation we may neglect the magnetic term and write

$$M_1 = \frac{e^2}{\hbar^2 c^2 \mu k'_0}\left((p'-p)\cdot e'\right)(\overline{u}'\cancel{e}u) \qquad (680)$$

which is the same result as we should have obtained from (546).

Let $Z_A(x)$ be the Hertzian vector corresponding to the radiation field potentials $A_\lambda(x)$, defined by

$$A_\lambda(x) = \frac{d}{dt} Z_\lambda(x) \tag{681}$$

Then the matrix element of $Z_\lambda(x)$ for emitting the photon with potentials (440) is

$$Z_\lambda(x) = \frac{1}{ick_0'} e_\lambda' e^{-ik' \cdot x} \tag{682}$$

Compare with (422).[91]

Thus the operator U_1 which has the matrix element (680) may be written[92]

$$U_1 = \frac{e^2}{\hbar^2 c\mu} \int dx \, \overline{\psi}(Z \cdot \partial) A^e \psi(x) \tag{683}$$

Using (670) and specializing by means of (674), this gives

$$H_{T1} = \frac{e}{\hbar\mu} \int \psi^*(Z \cdot \partial V)\psi \, d^3 \boldsymbol{r} \tag{684}$$

This completes the evaluation of H_T.

8.2.2   *Discussion and the Nature of the Φ-Representation*

To understand the effect of the transformation (662), we observe that if $\Psi(t)$ is the state of one real electron in the absence of an external field, then $\Phi(t)$ will be independent of t and will represent one "bare" electron with the same momentum as the real electron. In an actual hydrogen atom, we may consider the state $\Psi(t)$ to a very good approximation as a superposition of states of a single real free electron; then $\Phi(t)$ is a superposition of states of a single bare electron with the same distribution of momenta. Thus by the transformation from Ψ to Φ we have eliminated the radiation field surrounding the electron, all remaining effects of this radiation field being contained in the operator H_T.

It is essential at this point to make sure that in the Φ representation the field-operators are still free-particle operators, with the correct equations of motion for operators in the FIR. Thus the transformation (662) is only a transformation from one set of variables to another within the FIR, and does not take us out of the FIR. This point was never explained properly by Schwinger in his papers, although he no doubt understood it himself.

Let then $Q(x)$ be a field-operator of the Ψ representation. Being a FIR operator, $Q(x)$ satisfies

$$i\hbar\frac{dQ}{dt} = [Q, H_0] \qquad (685)$$

where H_0 is the Hamiltonian of the Dirac and Maxwell fields without interaction. In the Φ representation the corresponding field-operator is

$$Q'(x) = (\Omega_1(t))^{-1}\, Q(x)\, \Omega_1(t) \qquad (686)$$

Now $\Omega_1(t)$ is given by (492) where we now take the limit $g_A = 1$. The operators appearing in (492) are all FIR operators satisfying equations of motion of the form (685). When the integrations in (492) are carried out, the integrated terms will still have the same time variation given by (685), except for those terms which correspond to transitions in which energy is conserved. The energy-conserving matrix elements will have an explicit linear dependence on t after integration, which is not in accordance with (685). Therefore we conclude that the equation of motion

$$i\hbar\frac{d\Omega_1}{dt} = [\Omega_1, H_0] \qquad (687)$$

is valid for all matrix elements of Ω_1 which are not diagonal in the FIR. The same equation of motion is satisfied by $(\Omega_1(t))^{-1}$ with the same condition. Now we have seen that $\Omega_1(t)$ has no matrix elements diagonal in H_0 which give transitions either from or into one-particle states. Hence (687) holds for all matrix elements in which either initial or final state is a one-particle state.

Combining (687) and the corresponding equation for $(\Omega_1)^{-1}$ with (685) and (686), we have

$$i\hbar\frac{dQ'}{dt} = [Q', H_0] \qquad (688)$$

This equation (688) is valid for all matrix elements between one-particle states. Thus we can conclude that so long as we confine attention to a one-electron system, (688) is satisfied by all field operators of the Φ representation, and so the Φ representation is still within the FIR.

When we consider systems containing more than one particle, then $\Omega_1(t)$ will have an explicit dependence on time in addition to (687). Then the Φ representation would no longer belong to the FIR. And this is physically reasonable, because in many-electron systems it is not in general possible to transform away the radiation interaction completely, the radiation interaction by itself giving rise to real effects such as Møller scattering which we should not wish to transform away.

8.2.3   *Concluding Non-Covariant Part of the Calculation*

Having established that the Φ-representation in which (664) holds is the
FIR, we now proceed at once to transform to the BIR in which we shall
finish the calculation of the Lamb shift. To transform to the BIR, we write

$$\Phi(t) = e^{iH_0t/\hbar} e^{-i\{H_0+H^e\}t/\hbar} \Phi'(t) \tag{689}$$

The new wave-function $\Phi'(t)$ thus satisfies

$$i\hbar \frac{\partial \Phi'}{\partial t} = \{H_{T1}(t) + H_{T2}(t)\} \Phi' \tag{690}$$

where H_{T1} and H_{T2} are given by (684) and (676), only now the ψ^* and ψ
operators have the time-variation of the Dirac field in the external potential
V.

To solve (690) we may now use precisely the method that we used for
solving (245) in the non-relativistic treatment. There are only two differ-
ences, (i) we have now the extra term H_{T2}, and (ii) the form of H_{T1} is
different form (247).

Since we are working only to second order in the radiation interaction
and H_{T2} is already of that order, H_{T2} is to be treated only as a first-order
perturbation. Then H_{T2} will have no effect on the line-width Γ, and will
contribute to the line-shift ΔE just the expectation-value of (676) in the
state ψ_0 of the atom, namely

$$\Delta E_2 = \frac{\alpha}{3\pi\mu^2} \left\{ \log \frac{\mu}{2r} + \frac{11}{24} - \frac{1}{5} \right\} \int (\nabla^2 V)|\psi_0|^2 \, d^3\boldsymbol{r}$$

$$- \frac{i\alpha}{4\pi\mu} \int \psi_0^* \gamma_4 (\boldsymbol{\alpha} \cdot \nabla V)\psi_0 \, d^3\boldsymbol{r} \tag{691}$$

The effect of the change from (247) to (684) is that the matrix element
$j_\mu^k(n\ m)$ given by (256) now becomes everywhere replaced by

$$J_\mu^k(n\ m) = \frac{ie}{\hbar\mu|\boldsymbol{k}|} \int \psi_n^* \frac{\partial V}{\partial x_\mu} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \psi_m \, d^3\boldsymbol{r} \tag{692}$$

We see this by comparing (247) and (684), noting that

$$\boldsymbol{j}(r,t) \leftrightarrow \frac{ec}{\hbar\mu}\psi^*\psi\nabla V \int dt$$

or

$$\int \boldsymbol{j}^S(r)\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} d^3\boldsymbol{r} \leftrightarrow \frac{iec}{\hbar\mu|\boldsymbol{k}|c} \int \psi^*\psi\nabla V\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\, d^3\boldsymbol{r}$$

In the previous calculation we used a non-relativistic dipole approximation for j_μ which gave according to (272)

$$j_\mu^k(n\ m) = +\frac{ie\hbar}{m} \int \psi_n^* \frac{\partial\psi}{\partial x_\mu} d^3\boldsymbol{r} \tag{693}$$

We shall again use a dipole approximation and drop the exponential factor in (692). Then taking for the atom the non-relativistic Hamiltonian

$$H = \frac{p^2}{2m} + V \tag{694}$$

the difference between (692) and (693) becomes

$$J_\mu^k(n\ m) - j_\mu^k(n\ m) = \frac{ie}{\hbar\mu|\boldsymbol{k}|} \left\{ \int \psi_n^* \frac{\partial\psi}{\partial x_\mu} d^3\boldsymbol{r} \right\} (E_m - E_n - hc|\boldsymbol{k}|) \tag{695}$$

where we used $[p^2, p_\mu] = 0$ and $\int p_\mu\psi_n^* p^2\psi_m\, d^3\boldsymbol{r} = -\int \psi_n^*\, p_\mu p^2\psi_m\, d^3\boldsymbol{r}$. This difference vanishes for transitions in which energy is conserved. Hence the value of Γ given by (262) is unaffected by the change from j to J. The value Γ calculated previously is still valid in the relativistic theory, except for very small effects from magnetic radiation which we have neglected.

Using (695) we have the simple relation between j_μ and J_μ,

$$J_\mu^k(n\ m) = j_\mu^k(n\ m)\, \frac{E_m - E_n}{hc|\boldsymbol{k}|} \tag{696}$$

Using formula (261) with J substituted for j, the contribution from H_{T1} to the line-shift becomes instead of (273)

$$\Delta E_1 = -\frac{e^2}{6\pi^2 m^2\hbar c^3} \int_r^\infty \frac{dk}{k} \sum_n \frac{(E_n - E_0)^2 |p_{n0}|^2}{E_n - E_0 + hc|\boldsymbol{k}|} \tag{697}$$

The integral is now convergent at high frequencies and only divergent at low frequencies where the cut-off r is now required to make it finite. The shift (697) would be zero for a free particle, and so there is no question of subtracting away a mass-renormalization term as we did from (273); in the relativistic treatment the mass-subtraction was already done long before this stage of the calculation was reached.

Integrating (697) directly over k, and taking r to be small compared with $(E_n - E_0)$, we find

$$\Delta E_1 = -\frac{e^2}{6\pi^2 m^2 \hbar c^3} \sum_n (E_n - E_0)^2 \, |p_{n0}|^2 \log \frac{|E_n - E_0|}{hcr} \qquad (698)$$

This is precisely the non-relativistic line-shift (278) with r substituted for K. Defining $(E - E_0)_{\text{av}}$ by (279) and using (281), we have

$$\Delta E_1 = \frac{\alpha}{3\pi\mu^2} \left\{ \log \frac{hcr}{(E - E_0)_{\text{av}}} \right\} \int (\nabla^2 V) \, |\psi_0|^2 \, d^3 \boldsymbol{r} \qquad (699)$$

This combines with (691) to give for the total level shift

$$\Delta E = \frac{\alpha}{3\pi\mu^2} \left\{ \log \frac{mc^2}{2(E - E_0)_{\text{av}}} + \frac{11}{24} - \frac{1}{5} \right\} \int (\nabla^2 V) \, |\psi_0|^2 \, d^3 \boldsymbol{r}$$

$$- \frac{i\alpha}{4\pi\mu} \int \psi_0^* \gamma_4 (\boldsymbol{\alpha} \cdot \nabla V) \psi_0 \, d^3 \boldsymbol{r} \qquad (700)$$

a result which is completely divergence-free and independent of r.

The second term of (700) represents the effect of the anomalous magnetic moment of the electron on the energy levels. It therefore gives a spin-dependent shift which modifies slightly the fine-structure which arises from the Dirac magnetic moment. To evaluate this term we use the Dirac equations (see (38)).

$$mc^2 \gamma_4 \psi_0 = (E_0 - V)\psi_0 + i\hbar c (\boldsymbol{\alpha} \cdot \nabla) \psi_0 \qquad (701)$$

$$mc^2 \psi_0^* \gamma_4 = \psi_0^* (E_0 - V) - i\hbar c (\nabla \psi_0^* \cdot \boldsymbol{\alpha}) \qquad (702)$$

Using both (701) and (702) in turn in the second term of (700) and adding the results using $\alpha^i \gamma_4 + \gamma_4 \alpha^i = 0$ the terms in $(E_0 - V)$ cancel and we find[93]

$$2mc^2 \int \psi_0^* \gamma_4 (\boldsymbol{\alpha} \cdot \nabla V) \psi_0 \, d^3 \boldsymbol{r}$$

$$= -i\hbar c \int \left\{ (\nabla \psi_0^* \cdot \boldsymbol{\alpha})(\boldsymbol{\alpha} \cdot \nabla V)\psi_0 + \psi_0^* (\boldsymbol{\alpha} \cdot \nabla V)(\boldsymbol{\alpha} \cdot \nabla \psi_0) \right\} d^3 \boldsymbol{r}$$

$$= -i\hbar c \int \left\{ (\nabla \psi_0^* \cdot \boldsymbol{\sigma})(\boldsymbol{\sigma} \cdot \nabla V)\psi_0 + \psi_0^* (\boldsymbol{\sigma} \cdot \nabla V)(\boldsymbol{\sigma} \cdot \nabla \psi_0) \right\} d^3 \boldsymbol{r}$$

$$= -i\hbar c \int \left\{ \psi_0 \left[\nabla \psi_0^* \cdot \nabla V + i\boldsymbol{\sigma} \cdot (\nabla \psi_0^* \times \nabla V) \right] \right.$$

$$+ \psi_0^* \left[\nabla \psi_0 \cdot \nabla V + i \boldsymbol{\sigma} \cdot (\nabla \psi_0 \times \nabla V) \right] \} \, d^3 \boldsymbol{r}$$

$$= i \hbar c \int \left\{ + (\nabla^2 V) \psi_0^* \psi_0 - 2 i \psi^* [\boldsymbol{\sigma} \cdot (\nabla V \times \nabla)] \psi_0 \right\} \, d^3 \boldsymbol{r} \quad (703)$$

Here we used $\alpha^i = \epsilon \sigma_i$ (see p. 20), $\epsilon^2 = \mathbb{I}$, and the formula

$$(\boldsymbol{\sigma} \cdot \boldsymbol{B})(\boldsymbol{\sigma} \cdot \boldsymbol{C}) = (\boldsymbol{B} \cdot \boldsymbol{C}) + i(\boldsymbol{\sigma} \cdot \boldsymbol{B} \times \boldsymbol{C})$$

See Dirac, *The Principles of Quantum Mechanics*, third edition, p. 263.

Now suppose V is a central potential, a function of r only. Then

$$\nabla V \times \nabla = \frac{1}{r} \frac{dV}{dr} (\boldsymbol{r} \times \nabla) = \frac{1}{r} \frac{dV}{dr} \left(\frac{i}{\hbar} \boldsymbol{L} \right) \quad (704)$$

with $\boldsymbol{L}$, the orbital angular momentum, is given by (39). In this case (700) becomes

$$\Delta E = \frac{\alpha}{3\pi\mu^2} \left\{ \log \frac{mc^2}{2(E - E_0)_{\text{avg}}} + \frac{5}{6} - \frac{1}{5} \right\} \int (\nabla^2 V) \, |\psi_0|^2 \, d^3 \boldsymbol{r}$$

$$+ \frac{\alpha}{4\pi\mu^2\hbar} \int \psi_0^* \left(\frac{1}{r} \frac{dV}{dr} \right) (\boldsymbol{\sigma} \cdot \boldsymbol{L}) \, \psi_0 \, d^3 \boldsymbol{r} \quad (705)$$

In the non-relativistic theory of the hydrogen atom, the quantum number j given by (72) is related to the operator $(\boldsymbol{\sigma} \cdot \boldsymbol{L})$ by

$$\frac{1}{\hbar}(\boldsymbol{\sigma} \cdot \boldsymbol{L}) = \begin{cases} \ell, & j = \ell + \frac{1}{2} \\ -\ell - 1, & j = \ell - \frac{1}{2} \end{cases} \quad (706)$$

Therefore for the hydrogen atom

$$\Delta E = \frac{\alpha e^2}{3\pi\mu^2} \left\{ \log \frac{mc^2}{2(E - E_0)_{\text{avg}}} + \frac{5}{6} - \frac{1}{5} \right\} |\psi_0(0)|^2$$

$$+ \frac{\alpha e^2}{16\pi^2\mu^2} q \int \frac{1}{r^3} |\psi_0|^2 d^3 \boldsymbol{r} \quad (707)$$

where q is the coefficient (706).

For s-states we have $q = 0$ and so the shift reduces to (see (284))

$$\Delta E = \frac{8\alpha^3}{3\pi} \frac{1}{n^3} \, \mathrm{Ry} \left\{ \log \frac{mc^2}{2(E - E_0)_{\text{avg}}} + \frac{5}{6} - \frac{1}{5} \right\} \quad (708)$$

for the state with principal quantum number n. For all other states the term in $|\psi_0(0)|^2$ is zero, and the shift depends only on the integral

$$\overline{\left(\frac{1}{r^3}\right)} = \int \frac{1}{r^3} |\psi_0|^2 d^3\boldsymbol{r} \tag{709}$$

The value of (709) is given by Bethe, *Handbuch der Physik*, Vol. 24/1, p. 286, Eq. (3.26) [20]. It is

$$\overline{\left(\frac{1}{r^3}\right)} = \frac{1}{\ell(\ell + \frac{1}{2})(\ell + 1)n^3 a_o^3} \tag{710}$$

where a_o is the Bohr radius of the hydrogen atom. Hence the shift for states with $\ell \neq 0$ becomes

$$\Delta E = \frac{\alpha^3}{2\pi} \frac{1}{n^3} \operatorname{Ry} \frac{1}{(\ell + \frac{1}{2})(\ell + 1)} \quad \text{for } j = \ell + \frac{1}{2} \tag{711}$$

$$\Delta E = -\frac{\alpha^3}{2\pi} \frac{1}{n^3} \operatorname{Ry} \frac{1}{\ell(\ell + \frac{1}{2})} \quad \text{for } j = \ell - \frac{1}{2} \tag{712}$$

For the relative displacement of the $2s$ and $2p_{1/2}$ levels, which in the Dirac theory were degenerate, we have finally by subtracting (712) from (708)

$$\Delta E = \frac{\alpha^3}{3\pi} \operatorname{Ry} \left\{ \log \frac{mc^2}{2(E - E_0)_{\text{avg}}} + \frac{5}{6} - \frac{1}{5} + \frac{1}{8} \right\} = 1051 \, \text{Megacycles} \tag{713}$$

8.2.4   *Accuracy of the Lamb Shift Calculation*

With the relativistic calculation of the Lamb shift which we have done, this course comes to an end. In this calculation we have met and seen how to overcome all the problems of mass and charge renormalization. We can say we now have a workable quantum electrodynamics which will give finite and unambiguous values for all observable quantities.

This calculation of the Lamb shift was of course not exact. The two most important errors were

(i) using non-relativistic wave-functions and the dipole radiation approximation in evaluating the effects of H_{T1};

(ii) neglecting the finite mass of the proton.

To correct these errors, very long calculations have been done. In connection with (i), Baranger [21] has calculated the effect of using relativistic theory in the treatment of H_{T1} and he finds the observed shift increases by 7 megacycles. The effects of (ii) are being looked at by Salpeter [22] but are not greater than 1 - 2 megacycles at most. In addition we have not considered

(iii) effects of fourth order in the radiation interaction. These are being looked at by Kroll and others [23]; they are certainly less than 1 megacycle. [24]

Therefore the theoretical value of the Lamb shift now stands at 1058 ± 2 megacycles. There is no clear discrepancy between this and the experimental value 1062 ± 5 though a discrepancy may be found when the experiments and the theory are further cleaned up.

This page intentionally left blank

Notes

[1] "is" substituted for "in"

[2] Both the first and the second editions use Δ for the Laplacian differential operator; as there are many quantities with Δ in them to denote a change or shift, the use of ∇^2 for the Laplacian seems a safer choice.

[3] The text symbols "div", "grad" and "curl" have been replaced throughout by "$\nabla\cdot$", "∇", and "$\nabla\times$", respectively. For example, the equation (3) originally read

$$\frac{\partial \rho}{\partial t} + \operatorname{div} j = 0$$

[4] "differential" replaces "diff."

[5] In the second edition, this equation lacked a label.

[6] In the first edition, the coefficient of $\boldsymbol{\alpha} \cdot \nabla$ is $+i\hbar c$

[7] E lacked the exponent 2 in Eq. (47).

[8] e_s lacked a subscript s

[9] The quantity $\left(\left(mc^2 - E\right)^2 - \left(m^2 c^4 - E^2\right) \right)$ lacked an exponent 2.

[10] The *spur* of a matrix A is the sum of its diagonal elements, and denoted Spur A or Sp A. This term was originally used. This term is also, and more commonly, known as the *trace* of a matrix, denoted Tr A.

[11] "ever" was written for "even"

[12] This equation differs from the form given in the first edition by a factor of $-\frac{6}{\pi}$, but this form also seems incorrect; perhaps the correct relation is

$$V(r) = -4\pi \int_0^r \frac{1}{r_1^2}\, dr_1 \int_0^{r_1} r_2{}^2 \rho(r_2)\, dr_2$$

[13] In the second edition, the factor "dE_1" was missing

[14]In the literature, the gauge condition $\nabla \cdot \boldsymbol{A} = 0$ is now called "Coulomb gauge"; the choice of the gauge condition $\partial_\mu A^\mu = 0$ (using the Einstein summation convention) is called "Lorentz gauge". (See also Eq. (588).) In the first edition, Dyson uses Einstein's convention; in the second edition, Moravcsik does not. See also the parenthetical remark following Eq. (234a).

[15]Rewritten. In v.1, Dyson writes "The factor $\boldsymbol{p}_1 E_2 - \boldsymbol{p}_2 E_1$ is invariant for a Lorentz transformation parallel to the 3 axis." In v.2, Moravcsik writes "It is worth noting that the factor $\boldsymbol{p}_1 E_2 - \boldsymbol{p}_2 E_1$ is invariant under Lorentz transformations leaving the x_1 and x_2 components unchanged."

[16]These three articles may be found in Schwinger, *Selected Papers on Quantum Electrodynamics*.

[17]Deleted "for"; the original statement read "condition that for its $\delta I(\Omega) = 0$"

[18]Eq. (185) lacked a label in v.2. The discussion beginning at Eq. (182) and continuing to Eq. (186) is unusually different between the editions. What is here follows Moravcsik's v.2 with the addition of the phrase "the matrix element" at Eq. (185).

[19]A δ was missing: the equation read $i\hbar\,\mathcal{O} = [\,\delta I(\Omega), \mathcal{O}(\sigma)\,]$.

[20]The notation originally used for anticommutators was $AB + BA = [A, B]_+$. The more familiar $\{A, B\}$ has been used instead.

[21]Here, ψ was substituted for the original ϕ (the variable in Eq. (187)) for clarity.

[22]In the last commutator, the operator $\tilde{a}_{k'\lambda}$ lacked the tilde.

[23]"Bleuler" was written "Bleuber".

[24]Originally, the phrase read $A_\mu = \partial\Lambda/\partial x_\mu$; this seemed confusing as the original potential is itself A_μ.

[25]Eq. (222) lacked the lower limit $k_0 > 0$ on the integral.

[26]$D^{(1)}(x)$ was added on the left-hand side for clarity

[27]Eq. (230) lacked the sum over μ; *cf.* Eq. (170).

[28]"obtained" was inserted; statement read "per unit time is using . . ."

[29]The subscript A was missing in the term $j_{\lambda A}(x')$.

[30]The exponential in the first integral lacked i; it had the argument $-\boldsymbol{k}\cdot\boldsymbol{r}$.

[31]Bethe & Salpeter, Ref. [20], p. 249, Eq. 59.7

[32]There was no subscript "o" on the vacuum expectation value.

[33]"if" substuted for "it"; and "gives" added to the previous sentence.

[34]The limits on the integral were -1 and $+1$, and the value of the integral was given as $\frac{2}{3}$.

[35] The expression for a_o is not here in the original, but it appears before Eq. (240).

[36] The expression for ρ is not in the original.

[37] $a_o = 0.529177 \times 10^{-8}$ cm; Ry = 13.6056 eV.

[38] "time-independent" substituted for "time-dependent", which describes the Heisenberg representation.

[39] Sum over μ inserted.

[40] Following "given by (263)", the second edition has the phrase "with o suffices". The first edition lacks this phrase. As the sentence makes more sense without it, it has been deleted.

[41] "s states" replaces "x states".

[42] The coefficient of the first integral in Eq. (283) was $\hbar$; it has been replaced by $\hbar^2$. Also, Eq. (283) lacked a label in the second edition.

[43] Here, the Bohr radius was denoted a; it seemed reasonable to use a_o instead.

[44] "is" replaces "being"

[45] "by" replaces "be"

[46] In v.2, the Hermitian conjugate b_m^* lacked the asterisk.

[47] The argument of the exp function originally had the factor $(t - t_o)$; this has been replaced by the factor $(t - t')$.

[48] The function $F_\nu(k)$ was written as a scalar, $F(k)$. This is misleading; the right-hand side is a vector function, because it is linear in γ_ν. So $F(k)$ was promoted to $F_\nu(k)$.

[49] The phrase "integral in the" was inserted.

[50] "definite" replaces "indefinite"

[51] In v.2, this reads "Tr $\sum_\mu \gamma_\mu \gamma_\nu = 4$"

[52] The bottom limit of 0 was added to the last integral sign. Note that the change of variable is easier to follow by first observing

$$\int_0^1 dz \, (z - z^2) f(z - z^2) = 2 \int_0^{1/2} dz \, (z - z^2) f(z - z^2)$$

because the expression $(z - z^2)$ is symmetric about $z = \frac{1}{2}$.

[53] The original read "Then by (390) in (396) the second term is zero".

[54] "proton" replaces "photon"

[55] The potentials had a subscript ν and the gamma matrices a subscript λ.

[56] "from" replaces a second "in"

[57] The factor x_2 had a superfluous superscript "'".

[58] *A*, *B* added for clarity

[59] The words after "because", "using 4) … (451)" were added.

[60] In the second edition, the intermediate calculation was wrong; however, the conclusion was correct. It was rewritten up to Eq. (464). See the trace theorems Eq. (585) *et seq.* Also, note the identity

$$\text{Tr}\,(\not{a}_1 \not{a}_2 \not{a}_3 \not{a}_4) = (a_1 \cdot a_2)(a_3 \cdot a_4) - (a_1 \cdot a_3)(a_2 \cdot a_4) + (a_1 \cdot a_4)(a_2 \cdot a_3)$$

[61] The third equation lacked a subscript "0" on the variable k_0'; the fourth equation lacked a superscript "′" on the variable k_0'.

[62] Eq. (468) lacked a label. The word "simply" was inserted.

[63] "Thomson" replaces "Thompson".

[64] The first spinor u lacked a bar; $\overline{u}$ replaces u.

[65] The expression for r_o was added.

[66] In the 3$^{\text{rd}}$ edition of Heitler's book, see §25.

[67] In the 3$^{\text{rd}}$ edition of Heitler's book, see §26.

[68] The first potential lacked a slash; $\not{A}$ replaces A.

[69] The phrase "in which all emission operators stand to the left of all absorption operators" was lost in the transition from the first edition to the second.

[70] Nowadays called "normal order", this ordering arises in connection with *Wick's Theorem*:

(time ordered operators) = (normal ordered operators) + (all contractions)

the contractions being equal to the propagators S_F, D_F, and so on.

[71] In Schwinger's QED anthology.

[72] In Schwinger's anthology.

[73] Both the time ordering brackets lacked a right bracket. These were added.

[74] Unlike Dyson, Moravcsik cited Eq. (585) as well as Eq. (376). In Eq. (585) are Dirac matrix identities which establish the equality between the first two integrals in Eq. (522). Logically these identities should have been introduced before Chapter 6, but nothing prevents a reader making use of a "forward" reference.

[75] The subscript on the second $\mathscr{L}$ was originally "$_O$". It has been replaced with a subscript "$_D$".

[76] Again, a subscript μ has been appended to the function $F(k)$ to make it a Lorentz vector. See note 31 at Eq. (371).

[77] The last curly bracket was missing; it has been added.

[78] The times symbol $\times$ was inserted.

[79] This sentence read formerly "...the factor $\epsilon(q - \epsilon - \epsilon')$ is a continuous function of $(\epsilon + \epsilon')$ and tends to $\not{\epsilon}(q) \ldots$"

[80] A sentence, "Here, $(k \cdot p_y)_3$ denotes the scalar product of the space-like parts of the vectors k and p_y.", was deleted, because the expression $\boldsymbol{k} \cdot \boldsymbol{p}_y$ is self-explanatory. In both the first and second editions, scant attention was paid to three-vectors; sometimes an overhead arrow was used, but these were very few. In this typed version care has been taken to represent three-vectors with bold type, thus: $(A_x, A_y, A_z) = \boldsymbol{A}$.

[81] "in" inserted; the original read "lying the range"

[82] In the first line, twice in the fourth line, and in the first appearance in the sixth line, the expression $|u'^* u|$ lacked the exponent 2. These have been supplied.

[83] The modulus bars around p were absent.

[84] "slowly-varying" replaces "slowly-carrying"

[85] For "Lagrangian" read here "Lagrangian density". Field theorists, by an abuse of language, often say the first and mean the second.

[86] The original citation lacked Prodell's name.

[87] In Schwinger's anthology.

[88] The notation "Mc" is outmoded; usually this is written "MHz."

[89] "α" is the fine-structure constant, $\frac{1}{137.036}$; "$\boldsymbol{\alpha}$" is the Dirac matrix. Originally "$\boldsymbol{\alpha} \cdot \nabla$" was rendered as "$\alpha \cdot \mathrm{grad}$".

[90] The equation labeling was faulty. In the second edition, what is here labeled (676a) was a *second* (676), and what is here labeled (677a) had no label at all. It might have been all right to leave Eq. (677a) unlabeled, except that in the first edition, *both* Eq. (677) *and* Eq. (677a) are labeled (677)! This is a compromise.

[91] The comparison with Eq. (422) is not obvious. Perhaps Eq. (438) was meant?

[92] The integrand was originally written

$$\overline{\psi} \left(Z \cdot \mathrm{grad} \right) \not{A}^e \psi(x)$$

Since Z_λ is a Lorentz vector, the gradient must likewise be. So "grad" here must be ∂_λ, not ∇. One goes from (682) to (680) by an integration by parts; since there are in (680) apparently the dot product of two 4-vectors, this supports the identification here of grad $= \partial_\lambda$.

[93] In the second to last line, the term $i\boldsymbol{\sigma} \cdot (\nabla \psi_0 \times \nabla V)$ was written with a dot product between the two gradients, rather than a cross product.

This page intentionally left blank

References

[1] Wolfgang Pauli, *General Principles of Quantum Mechanics*, trans. P. Achuthan and K. Venkatesan, Springer-Verlag, Berlin, 1980. This is an English translation of "Principien der Quantentheorie I" in *Handbuch der Physik*, v. 5, 1958, which is a revised edition of the original 1933 work reprinted by Edwards in 1947. The 1933 chapter on quantum electrodynamics is reprinted as Chapter X in the revised English edition.

[2] W. Heitler, *The Quantum Theory of Radiation*, 3$^{\rm rd}$ ed., Oxford U. P., Oxford, 1954. Reissued in 1984 by Dover Publications.

[3] G. Wentzel, *Introduction to the Quantum Theory of Wave Fields*, Interscience, NY, 1949. Reissued in 2003 by Dover Publications as *Quantum Theory of Fields*.

[4] J. Schwinger, ed., *Selected Papers on Quantum Electrodynamics*, Dover Publications, New York, 1958. Many of the most important Feynman, Schwinger and Dyson papers, together with those of other authors, are gathered in this anthology edited by Schwinger.

[5] Arthur I. Miller, *Early Quantum Electrodynamics: a source book*, Cambridge U. P., Cambridge UK, 1994. Miller's book includes a valuable historical essay and English translations of three articles cited by Dyson: Heisenberg's on the Dirac theory of the positron (*Zeits. f. Phys.* **90** (1934) 209), Kramers's suggestion of mass renormalization (*Nuovo Cim.* NS **15** (1938) 108), and the Pauli-Weisskopf discussion of the relativistic many-particle (scalar) theory (*Helv. Phys. Acta* **7** (1934) 709).

[6] Silvan S. Schweber, _An Introduction to Relativistic Quantum Field Theory_, Row, Peterson & Co., Evanston, IL, 1961. This magisterial textbook has been reissued by Dover Publications (2005) in paperback. Contains a very complete set of references to the QED work done from 1926-1960.

[7] Silvan S. Schweber, _QED and the Men Who Made It: Dyson, Feynman, Schwinger and Tomonaga_, Princeton U. P., Princeton NJ, 1994. A very readable, technical history of QED.

[8] David Kaiser, _Drawing Theories Apart: the dispersion of Feynman diagrams in postwar physics_, U of Chicago Press, Chicago, 2005. The sociology of the transmission of Feynman's graphical techniques.

[9] P. A. M. Dirac, "The quantum theory of the electron", _Proc. Roy. Soc. A_ **117** (1928) 610.

[10] H. Yukawa, "On the interaction of elementary particles", _Prog. Theo. Phys._ **17** (1935) 48. In Henry A. Boorse and Lloyd Motz, _The World of the Atom_, vol. II, Basic Books, Inc., New York, 1966, pp 1419–1422.

[11] W. Pauli, "The Connection Between Spin and Statistics", _Phys. Rev._ **58** (1940) 716. In Schwinger, _Selected Papers in Quantum Electrodynamics_, pp 372–378.

[12] W. Pauli and V. Weisskopf, "The quantization of the scalar relativistic wave equation", _Helv. Phys. Acta_ **7** (1934) 709. In Miller, _Early Quantum Electrodynamics_, pp 188–205 (English).

[13] R. E. Peierls, "The commutation laws of relativistic field theory", _Proc. Roy. Soc. A_ **214** (1952) 143. Note the year of publication is 1952.

[14] Theodore A. Welton, "Some Observable Effects of the Quantum-Mechanical Fluctuations of the Electromagnetic Field", _Phys. Rev._ **74** (1948) 1157. A modern and very instructive discussion of Welton's work may be found in Barry R. Holstein's _Topics in Advanced Quantum Mechanics_, Addison-Wesley Publishing Co., Redwood City, CA, 1992, pp 181-184. Ted Welton was a friend and undergraduate classmate of Feynman's at MIT. See Schweber, _QED and the Men Who Made It_, pp 375-387.

[15] R. R. Wilson, "Scattering of 1.3^3 MeV Gamma Rays by an Electric Field", *Phys. Rev.* **90** (1953) 720. Wilson was at Cornell at the time; also he notes "The measurements here reported were all made in 1951. Publication has been held up until now in the hope that the Rayleigh scattering could be calculated more accurately."

[16] H. A. Kramers, "The interaction between charged particles and the radiation field", *Nuovo Cim.* NS **15** (1938) 108. English translation in Miller, *Early Quantum Electrodynamics*, pp 254–258.

[17] E. A. Uehling, "Polarization Effects in the Positron Theory", *Phys. Rev.* **48** (1935) 55.

[18] E. C. G. Stueckelberg, "Une propriété de l'operateur S en mécanique asymptotique", *Helv. Phys. Acta* **19** (1946) 242. See also D. Rivier & E. C. G. Stuecklberg (*sic*), "A convergent expression for the magnetic moment of the muon", *Phys. Rev.* **74** (1948) 218.

[19] F. J. Dyson, "Heisenberg operators in quantum electrodynamics", *Phys. Rev.* **82** (1951) 428. Dyson introduces the term "normal product" on pp 429–430.

[20] Hans A. Bethe & Edwin E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms*, Springer-Verlag, Berlin, 1957. Reissued by Plenum Publishing Co., New York, 1977, (paperback edition). This is a revised and updated version of the article Dyson cites, H. A. Bethe, "Quantenmechanik der Ein- und Zwei-Elektronenprobleme", *Handbuch der Physik*, Bd. 24/1, Springer, Berlin, 1933. The relevant formula for Eq. (710) is to be found in the new work with exactly the same label, (3.26), on p. 17. Note that Bethe & Salpeter use Hartree's "atomic units", so that distances are measured in terms of a_o.

[21] M. Baranger, H. A. Bethe & R. P. Feynman, "Relativistic Corrections to the Lamb Shift", *Phys. Rev.* **92** (1953) 482.

[22] E. E. Salpeter, "Mass Corrections to the Fine Structure of Hydrogen-Like Atoms", *Phys. Rev.* **87** (1952) 328.

[23] M. Baranger, F. J. Dyson & E. E. Salpeter, "Fourth-Order Vacuum Polarization", *Phys. Rev.* **88** (1952) 680.

[24] E. E. Salpeter, "The Lamb Shift for Hydrogen and Deuterium", *Phys. Rev.* **89** (1953) 93.

[25] J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons*, Addison-Wesley Publishing Co., Cambridge, MA, 1955.

[26] E. T. Jaynes, "Disturbing the Memory", http://bayes.wustl.edu/etj/ node2.html; link #18, 1984.

[27] F. J. Dyson, "The Radiation Theories of Tomonaga, Schwinger and Feynman", *Phys. Rev.* **75** (1949) 486.

[28] F. J. Dyson, "The S-Matrix in Quantum Electrodynamics", *Phys. Rev.* **75** (1949) 1736.

[29] F. J. Dyson, "Advanced Quantum Mechanics", http://hrst.mit.edu/ hrs/renormalization/dyson51-intro/index.html, 1951.

Index