RELATIVISTIC QUANTUM THEORY

Physics IV Lecture Course, second semester 2004

Professor Don Melrose
School of Physics
University of Sydney

Contents

Preface iv

Lecture 1 : Nonrelativistic Quantum Mechanics 1

Lagrangian and Hamiltonian Mechanics, Poisson Brackets, The State Function and Operators, Representations, Pictures for the Time Evolution, The Schrödinger Equation, Inclusion of Spin.

Lecture 2 : 4-Tensor Notation 7

4-Tensor Equations, Important 4-Vectors, Lorentz Transformations, Covariant Form of Maxwell's Equations, Continuity Equations for Charge and Energy, Gauge Transformations.

Lecture 3 : The Klein Gordon and Dirac Equations 12

The Klein Gordon Equation, The Dirac Equation, The Dirac Hamiltonian, The Dirac Conjugate, Particles and Antiparticles.

Lecture 4 : Plane Wave Solutions 16

Momentum Space Representation, Free Spin 0 and Spin 1 Particles, Free Spin $\frac{1}{2}$ Particles, The Spinors $u_s(\boldsymbol{p})$ and $v_s(\boldsymbol{p})$, The Zitterbewegung, The Klein Paradox.

Lecture 5 : : Field Lagrangians 22

Euler-Lagrange Equations for a Field, The Klein Gordon Lagrangian, Dirac Lagrangian, The Lagrangian in Momentum Space, Particle Action and Occupation Numbers.

Lecture 6 : Motion in a Magnetostatic Field 26

Dirac's Equation in the Landau Gauge, Construction of the Wavefunctions, Spin Eigenfunctions, The Normalization and Orthogonality Relations, The Center of Gyration, The Magnetic Moment of the Electron, Magnetic Moment Tensor.

Lecture 7 : Dirac Theory for Hydrogen-like Atoms 32

Dirac Equation for a Coulomb Field, The Angular Momentum Eigenstates, The Radial Eigenvalues, The Fine Structure Formula, The Hydrogen-like Klein Gordon Atom.

Lecture 8 : Nonrelativistic Limit of the Dirac Equation 39

The Large and Small Components, The Foldy-Wouthuysen Transformation, Relativistic Corrections to the Pauli-Schrödinger Equation.

Lecture 9 : Formal Aspects of the Dirac Equation 43

Independent 4×4 Matrices, Traces of Products of γ-Matrices, Other Representations of the Dirac Matrices, Lorentz Transformation of the Dirac Equation, Explicit Forms for the Transformation Matrices $S(L)$, The Parity, Time Reversal and Charge Conjugation Operators.

Lecture 10 : Photons and Neutrinos 48

Neutral Bosons, Photons, Normalization to One Photon, Specific Wave Modes, Neutrinos.

Lecture 11 : Second Quantization 54

The Harmonic Oscillator, Anticommutation Relations for Fermion Fields, Quantization of Fields, Free Particle States, Normal Ordering.

Lecture 12 : Propagators 59

Solution of the Inhomogeneous Particle Equations, The Feynman Contour, The Chronological Operator, The Vacuum Expectation Value, The Propagator as a Vacuum Expectation Value, The Photon Propagator.

Lecture 13 : The S-Matrix 66

The Interaction Picture, Evolution of $\hat{S}$, Specific Interaction Terms, The Scattering Amplitudes T_{if} and M_{if}, The Density of Final States.

Lecture 14 : Wick's Theorem and Induced Effects 70

Identification of S_{fi} for Specific Processes, Contractions, First Order Processes, Equal Time Contractions, Second Order Processes, The Density Matrix, Statistical Averages, Induced Effects.

Lecture 15 : Introductions to Feynman Diagrams 79

Elements in Diagrams for QED, Initial and Final States, Propagators: Internal Lines, External Fields, Multiple Photon Vertices, Rules for Drawing Diagrams, Rules for Constructing S_{fi}.

Lecture 16 : Diagrams in Momentum Space 87

Conversion to Momentum Space, Loop Momenta, Static Fields, Sums over States of Polarization, Additional Rules for Diagrams in Momentum Space, Rules for Constructing iM_{fi} and $w_{i\to f}$.

Lecture 17: Cerenkov Emission 93

The Probability for Cerenkov Emission, Average over Polarization States, The Power Radiated in Transverse Waves, Kinetic Equations, Crossed Processes, Kinematic Restrictions.

Lecture 18 : Gyromagnetic Processes 100

The Probability of Gyromagnetic Emission, Transition Rates, Gyromagnetic Masers, One-Photon Pair Decay.

Lecture 19 : Mott Scattering and Bremsstrahlung 107

Scattering of an Electron by a Coulomb Field, The Mott Cross Section, The Bethe-Heitler Formula, Bremsstrahlung Emission of Soft Photons.

Lecture 20 : Compton Scattering and Relates Processes 115

Invariant Kinematics, Compton Scattering, The Cross Section, The Klein Nishina Formula, Pair Annihilation into Two Photons.

APPENDICES

Appendix A : Fourier Transforms 124

4-Dimensional Fourier Transform, The Reality Condition, The Reality Condition, The Convolution Theorem, Truncations and the Dirac δ-functions, Fourier Transforms of the Step and Sign Function, The Plemelj Formula, Confinement to the Forward Light Cone.

Appendix B : Specific Spin Eigenfunctions 127

Helicity States, Eigenstates of the Magnetic Moment Operator, Magnetic Moment Eigenstates in a Magnetic Field, Eigenstates in Cylindrical Coordinates in a Magnetic Field, Helicity Eigenstates in a Magnetic Field, Magnetic Moment Eigenstates in a Magnetic Field.

Appendix C : The Poincaré Group and Spin Operators 134

Generators of the Poincaré Group, Invariants of the Poincaré Group, Conserved Quantities and Constants of the Motion, Spin Operators that Commute with the Hamiltonian.

Appendix D : Wave Dispersion 139

The Linear Response Tensor, The Wave Equation, The Polarization 4-vector, Ratio of Electric to Total Energy, The Photon Propagator.

Appendix E : The Vertex Function in a Magnetic Field 148

The Vertex Function in the Landau Gauge, The Vertex Function in the Cylindrical Gauge, Average over the Position of the Gyrocenter, Properties of the Functions $J_\nu^n(x)$, Approximate Forms for the Vertex Function.

EXERCISES

Exercise Set 1 156 Exercise Set 1 158

Preface

In one sense, the general outline of this course has not been changed greatly since I first gave it in 1985. There have been numerous changes from year to year as some lectures have been deleted and other lectures have been added. This set of notes, which is being made available to all attending the course, includes most of the material given in the main part of the course in earlier years. It is unrealistic to attempt to cover all the material in these notes in detail in the time available, and some of the lectures will be used for reference purposes rather than as integral parts of the course.

In 1985 the course was nominally 27 lectures, and it is now nominally 20 lectures. The original course included sections on renormalization and radiative corrections in QED and on gauge field theories and QCD. The notes on these topics (Lectures 21–25) and two more detailed appendices (Appendices D and E) are available on request.

The present course may be separated into roughly three parts. The first part is an introduction to relativistic wave equations, specifically to the Dirac and Klein Gordon equations. The second part is an introduction to quantum field theory, with the main emphasis being on Quantum Electrodynamics. The third part covers several applications of Quantum Electrodynamics.

In preparing these notes I have not followed any textbook closely. The textbook to which I have referred most is

"Relativistic Quantum Theory, Part 1", by V.B. Berestetskii, E.M. Lifshitz and L.P. Pitaevskii, Pergamon Press (1971) (Volume 4, Part 1 in the Landau and Lifshitz series).

Other textbooks that I have used to a lesser extent include

"Advanced Quantum Theory", by M.D. Scadron, Springer (1979).

"Quantum Mechanics, Volume II", by A. Messiah, North Holland (1961).

"Quantum Field Theory", by C. Itzykson and J. Zuber, McGraw-Hill (1980).

Don Melrose
July 2004

Lecture 10

Photons and Neutrinos

So far we have been concerned only with charged bosons and fermions. Uncharged particles may be separated into two classes depending on their properties under under the formal operation that converts particles into antiparticles. One class consists of those uncharged particles that have some property other than charge that changes sign under this operation. An example is the Λ^0 hyperon that has strangeness $S = 1$; its antiparticle $\bar{\Lambda}^0$ has strangeness $S = -1$. Another example is the neutrino, of which there are several flavors; antineutrinos are different from neutrinos due to their helicities being opposite. The other class of uncharged particles have no such additional charge-like property that can distinguish the antiparticle from the particle. Such particles are their own antiparticles, and are called strictly neutral particles, or more simply *neutral* particles. Examples of neutral particles include the π^0 meson and the photon.

1. Neutral Bosons

Neutral bosons obey the Klein Gordon equation with a wavefunction $\Psi(x) = \Psi^*(x)$ that is real. Its Fourier transform

$$\phi(p) := \int d^4x \; \Psi(x) \; e^{ipx} \tag{10.1}$$

therefore satisfies

$$\phi(-p) = \phi^*(p). \tag{10.2}$$

The negative frequencies then contain the same information as the positive frequencies.

In the formal treatment of neutral bosons, the only notable change that needs to be made is the inclusion of a factor $\frac{1}{2}$ in the Lagrangian. The expression (5.5) for charged bosons is replaced by

$$\mathcal{L}(x) = \tfrac{1}{2}\left[(\partial_\mu \Psi)(\partial^\mu \Psi) - m^2\Psi^2\right]. \tag{10.3}$$

This change involves using the fact that Ψ is real, and noting that there is then only one independent function and not two, so that one needs to include the factor $\frac{1}{2}$ so that its effect is not counted twice.

2. Photons

Quantization of the electromagnetic field in vacuo presents difficulties. The field A^μ is gauge dependent, and for transverse waves in vacuo it must satisfy $\boldsymbol{k} \cdot \boldsymbol{A} = 0$. This implies that only two of the components of A^μ can describe photons. These two transverse components are to be quantized in order to obtain the two transverse states of polarization for photons, and this must involve some procedure for separating them from the other two components (the longitudinal and time components of A^μ) . One procedure for quantizing the field involves introducing fictitious longitudinal and time-like photons, quantizing all

four components, and then arranging so that the physical effects of the ficticious photons cancel each other exactly.

One of my research interests is in synthesizing quantum electrodynamic and the kinetic theory of plasmas, and this leads to quite a different attitude to the problem of quantizing the electromagnetic field. Even in the simplest plasma, longitudinal waves are allowed. Examples of longitudinal waves include Langmuir waves and ion sound waves. In less simple plasmas, specifically in anisotropic or magnetized plasmas, the polarization of waves may be intermediate between transverse and longitudinal. One preliminary problem before quantizing the wave field for waves in a plasma is to identify the properties of the waves, which properties include the polarization. One uses the theory of waves in plasmas, which is usually developed in a non-covariant form. A covariant version of the theory of waves in plasmas is summarized in Appendix D.

A formal problem arises: the wave field includes not only the electromagnetic field but also the response of the plasma, which involves induced motions of the particles. Thus, for example, the energy in the waves includes contributions from the electric field, from the magnetic field and from the induced particle motions. The formal problem is that one needs to define the wave subsystem before quantizing it. To overcome this problem I have developed a procedure for quantizing in momentum space, cf. (5.12) *et seq.* All the wave properties are functions of k and one may deduce the appropriate wave Lagrangian by requiring that the corresponding Euler-Lagrange equations reproduce the wave equation in k-space. This procedure is also outlined in Appendix D.

Quanta of Langmuir waves are often called *plasmons*, and other special names are given to quanta of other classes of waves, e.g., phonons, magnons, and so on. Here I shall not normally use such special names, and shall use *photon* in a generic sense to mean a quantum of any wave field. Different classes of waves are referred to as wave *modes*. Thus one has the Langmuir mode, the ion sound mode, the transverse mode, and so on.

There are three basic properties for each specific wave mode.

1) The condition for the wave equation to have a solution is an algebraic relation involving the components of k, called the *dispersion equation*. A specific solution of the dispersion equation is a *dispersion relation*. Each distinct dispersion relation defines a distinct wave mode (but when one solution is "distinct" from another is not defined, and is sometimes a matter of choice). Thus a dispersion relation is an algebraic relation amongst the components of k. It is usually convenient to express this as a relation for the frequency ω as a function of the components of the 3-vector $\boldsymbol{k}$. This involves solving the dispersion equation with the components of $\boldsymbol{k}$ being the independent variables, and ω being the dependent variable. Suppose a specific wave mode is labeled by a subscript M. Then the dispersion relation may be written in the form $\omega = \omega_M(\boldsymbol{k})$. An alternative formal way of writing a dispersion relation is in terms of the wave 4-vector, specifically $k = k_M$. The relation between these two forms is identified by writing $k^\mu_M = [\omega_M(\boldsymbol{k}), \boldsymbol{k}]$.
2) Once a dispersion relation has been found, the wave equation may be solved for $A^\mu(k_M)$. The phase, gauge and normalization of $A^\mu(k_M)$ are arbitrary. After making appropriate choices of these, the solution is written as the polarization 4-vector $e^\mu_M(\boldsymbol{k})$. One needs to choose a specific gauge in order to impose a normalization on $e^\mu_M(\boldsymbol{k})$. The most convenient choice of gauge is the temporal gauge; then one has

$e^{\mu}_{M}(\boldsymbol{k}) = [0, \boldsymbol{e}_M(\boldsymbol{k})]$, and the normalization $\boldsymbol{e}_M(\boldsymbol{k}) \cdot \boldsymbol{e}^{*}_{M}(\boldsymbol{k}) = 1$ implies the normalization $e^{\mu}_{M}(\boldsymbol{k}) e_{M\mu}(\boldsymbol{k}) = -1$.

3.) The remaining property is the ratio $R_M(\boldsymbol{k})$ of electric to total energy in the waves. Physically this property is important because the coupling between a charged particle and a wave field is due only to the electric field in the wave. Hence the rate for any process is proportional to $R_M(\boldsymbol{k})$.

3. Normalization to One Photon

We wish to normalize our wave field so that it corresponds to one photon in the normalization volume. This is achieved as follows.

The 4-potential $A^{\mu}_{M}(x)$ for waves in a mode M may be written in the form

$$A^{\mu}_{M}(x) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, a_M(\boldsymbol{k}) \big[e^{\mu}_{M}(\boldsymbol{k}) e^{-ik_M x} + e^{*\mu}_{M}(\boldsymbol{k}) e^{ik_M x} \big], \tag{10.4}$$

where the positive and negative solutions are included separately so that $\omega_M(\boldsymbol{k}) > 0$ is implicit. Suppose we use the temporal gauge so that the scalar potential A^0 is zero. Then the electric field is given by $\boldsymbol{E} = -\partial \boldsymbol{A}/\partial t$, which may be evaluated using (10.4) simply by noting that $\partial/\partial t$ operating on $e^{\pm ik_M x}$ gives $\pm\omega_M(\boldsymbol{k}) e^{\pm ik_M x}$. Then integrating the electric energy density $\frac{1}{2}\varepsilon_0 |\boldsymbol{E}|^2$ over all space (and averaging over time) gives the electrical energy in the wave field. This is

$$\int d^3\boldsymbol{x}\, \tfrac{1}{2}\, \varepsilon_0 |\boldsymbol{E}(x)|^2 = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, \varepsilon_0 |\omega_M(\boldsymbol{k}) a_M(\boldsymbol{k})|^2, \tag{10.5}$$

where the normalization condition for the polarization vectors have been used. The electrical energy in the range $V\, d^3\boldsymbol{k}/(2\pi)^3$ may then be identified as $\varepsilon_0 |\omega_M(\boldsymbol{k}) a_M(\boldsymbol{k})|^2$. The electrical energy may be expressed in terms of the total energy $W_M(\boldsymbol{k})$ in this range by writing the electrical energy as $R_M(\boldsymbol{k}) W_M(\boldsymbol{k})$. Thus one has

$$R_M(\boldsymbol{k}) W_M(\boldsymbol{k}) = \varepsilon_0 |\omega_M(\boldsymbol{k}) a_M(\boldsymbol{k})|^2. \tag{10.6}$$

Normalization to one photon then involves identifying the energy $W_M(\boldsymbol{k})$ as being equal to the energy $\hbar\omega_M(\boldsymbol{k})$, where $\hbar$ is included momentarily for clarity, and is now set equal to unity again. Thus normalization to one photon corresponds to

$$|a_M(\boldsymbol{k})| = \left[\frac{\mu_0 R_M(\boldsymbol{k})}{\omega_M(\boldsymbol{k}) V} \right]^{1/2}, \tag{10.7}$$

where $\mu_0 \varepsilon_0 = 1$ (in natural units) has been used.

4. Specific Wave Modes

The only waves that considered here are transverse waves in vacuo and in isotopic media and Langmuir waves in an isotropic plasma. The properties of these waves are summarized here. These properties apply only in the rest frame of the medium (except of course for transverse waves in vacuo). The wave properties in any other frame may be obtained by applying a Lorentz transformation.

Transverse Waves in an Isotropic Medium

In an isotropic medium, such as air, water or glass, electromagnetic waves are transverse with a refractive index $\mu = |\boldsymbol{k}|/\omega$ that is different from unity. The dispersion relation may be written in terms of the refractive index $\mu(\omega)$ as a function of ω. The wave properties are

$$\mu(\omega) = |\boldsymbol{k}|/\omega, \quad \boldsymbol{e}\cdot\boldsymbol{\kappa} = 0, \quad R(\omega) = \frac{1}{2\mu(\omega)(\partial/\partial\omega)\big(\omega\mu(\omega)\big)}. \tag{10.8}$$

In this case the dispersion equation has been solved with ω and $\boldsymbol{\kappa}$ as the independent variables, and $|\boldsymbol{k}|$ or μ as the dependent variable. Transverse waves in vacuo have $\mu(\omega) = 1$.

Transverse waves have two degenerate states of polarization, and in general it is not possible to describe the polarization in terms of a single polarization vector. In practice one often sums over final states of polarization and averages over initial states of polarization, with the average being half the sum. The sum over states of polarization may be performed as follows. First note that if the two transverse polarization vectors are chosen to correspond to linear polarizations, then the corresponding two polarization vectors form an orthonormal set of basis vector with $\boldsymbol{\kappa}$. By representing the unit tensor δ_{ij} in terms of this basis, one infers the relation

$$\sum e_i e_j^* = \delta_{ij} - \boldsymbol{\kappa}_i\boldsymbol{\kappa}_j, \tag{10.9}$$

where the sum is over the two linear polarization vectors. One then notes that the sum on the left applies to any two orthonormal vectors that span the 2-dimensional space orthogonal to $\boldsymbol{\kappa}$, so that (10.9) is completely general.

Waves in an Isotropic Plasma

A simple model for a plasma is that of an isotropic thermal plasma. Assuming the plasma be neutral and to consist of electrons and singly charged ions of one species, there are three relevant plasma parameters in this model: the electron number density n_e, the electron temperature T_e and the ion mass m_i. The electron plasma frequency, usually called simply the *plasma frequency* ω_p, and the ion plasma frequency ω_{pi} are defined by

$$\omega_p = \left(\frac{e^2 n_e}{\varepsilon_0 m_e}\right)^{1/2}, \quad \omega_{pi} = \left(\frac{e^2 n_i}{\varepsilon_0 m_i}\right)^{1/2}, \tag{10.10}$$

with $n_i = n_e$. Other relevant plasma parameters are *electron thermal speed* V_e, the *Debye length* λ_{De}, and the *ion sound speed* v_s:

$$V_e = \left(\frac{T_e}{m_e}\right)^{1/2}, \quad \lambda_{De} = \frac{V_e}{\omega_p}, \quad v_s = \omega_{pi}\lambda_{De} = \left(\frac{m_e}{m_i}\right)^{1/2} V_e, \tag{10.11}$$

where we use units in which Boltzmann's constant is set equal to unity.

There are three wave modes in such a plasma. The transverse mode has the properties summarized in (10.8) with

$$\mu(\omega) = \left(1 - \frac{\omega_p^2}{\omega^2}\right)^{1/2}. \tag{10.12}$$

Note that in this case one has $R(\omega) = \frac{1}{2}$, and that transverse wave exist only for $\omega > \omega_p$. The dispersion relation (10.12) can be written in the standard form $\omega = \omega_T(\boldsymbol{k})$ with

$$\omega_T(\boldsymbol{k}) = [\omega_p^2 + \boldsymbol{k}^2 c^2]^{1/2}, \tag{10.13}$$

where $M = T$ denotes the transverse mode.

The other two modes are longitudinal and hence have polarization vector

$$\boldsymbol{e} = \boldsymbol{\kappa}. \tag{10.14}$$

One is the Langmuir mode ($M = L$) with dispersion relation $\omega = \omega_L(\boldsymbol{k})$, with

$$\omega_L(\boldsymbol{k}) = \omega_p + \frac{3\boldsymbol{k}^2 V_e^2}{2\omega_p}, \qquad R_L(\boldsymbol{k}) = \frac{\omega_p^2}{2\omega_L^2(\boldsymbol{k})}. \tag{10.15}$$

The Langmuir mode exists only for $|\boldsymbol{k}|\lambda_{De} \ll 1$. The other mode is the ion sound mode $M = S$ with

$$\omega_S(\boldsymbol{k}) = \frac{|\boldsymbol{k}|v_s}{\sqrt{1 + \boldsymbol{k}^2\lambda_{De}^2}}, \qquad R_S(\boldsymbol{k}) = \frac{\omega_S^2(\boldsymbol{k})}{2\omega_{pi}^2}. \tag{10.16}$$

The ion sound mode exists only at $\omega < \omega_{pi}$. Although these properties are independent of the temperature of the ions, in fact the mode exists only if the ions are much cooler than the electrons.

4. Neutrinos

Neutrinos are massless spin-$\frac{1}{2}$ particles. For $m = 0$ the covariant form of Dirac's equation (3.5) reduces to

$$i\partial\!\!\!/\Psi(x) = 0, \quad \text{or} \quad \hat{p}\!\!\!/\Psi(x) = 0. \tag{10.17}$$

A plane wave solution

$$\Psi(x) = u(p)\, e^{-ipx} \tag{10.18}$$

satisfies

$$\hat{p}^\mu \Psi(x) = p^\mu \Psi(x). \tag{10.19}$$

The plane wave solutions (10.18) are eigenfunctions of the 4-momentum operator by construction. To construct spin eigenfunctions we need to introduce an appropriate spin

operator. A suitable operator is the *helicity* operator $\hat{W}^\mu$ which is introduced in (C.13) of Appendix C:

$$\hat{W}^\mu := -\tfrac{1}{2}\epsilon^{\mu\alpha\beta\gamma}\hat{S}_{\alpha\beta}\hat{p}_\gamma, \tag{10.20}$$

where

$$\hat{S}^{\alpha\beta} = \frac{i}{4}[\gamma^\alpha, \gamma^\beta]. \tag{10.21}$$

is the spin 4-tensor in the Dirac theory. The helicity operator $\hat{W}^\mu$ may be shown to satisfy the identities

$$\hat{W}^\mu = \frac{1}{4}[\gamma^\mu, \hat{\not{p}}]\gamma^5, \quad \hat{W}^\mu\gamma^5 = \frac{1}{4}[\gamma^\mu, \hat{\not{p}}], \tag{10.22}$$

which enable us to derive eigenfunctions of $\hat{p}^\mu$ and $\hat{W}^\mu$ as follows.

Using (10.17), (10.19) and (10.22) one finds

$$\begin{aligned} \hat{W}^\mu\Psi(x) &= -\tfrac{1}{2}p^\mu\gamma^5\Psi(x), \\ \hat{W}^\mu\gamma^5\Psi(x) &= -\tfrac{1}{2}p^\mu\Psi(x). \end{aligned} \tag{10.23}$$

It follows that

$$\Psi_L(x) := \tfrac{1}{2}(1+\gamma^5)\Psi(x), \qquad \Psi_R(x) := \tfrac{1}{2}(1-\gamma^5)\Psi(x) \tag{10.24}$$

are simultaneous eigenfunctions of $\hat{p}^\mu$ and $\hat{W}^\mu$ with eigenvalues $-\frac{1}{2}p^\mu$ and $\frac{1}{2}p^\mu$ of $\hat{W}^\mu$, respectively. These are the helicity eigenfunctions. The operators $\frac{1}{2}(1\pm\gamma^5)$ may be regarded as projection operators. They satisfy

$$[\tfrac{1}{2}(1\pm\gamma^5)]^2 = \tfrac{1}{2}(1\pm\gamma^5), \tag{10.25}$$

and project onto the two helicity states. Note that the Dirac conjugates of (10.24) are

$$\overline{\Psi}_L(x) := \tfrac{1}{2}\overline{\Psi}(x)(1-\gamma^5), \qquad \overline{\Psi}_R(x) := \tfrac{1}{2}\overline{\Psi}(x)(1+\gamma^5). \tag{10.26}$$

It is not obvious how antineutrinos are related to neutrinos. It is conceivable that there are neutrinos with both helicities and antineutrinos with both helicities, although it is not clear how one would then distinguish neutrinos from antineutrinos. It is also conceivable that the neutrino has one helicity and the antineutrino has the other, but this would imply that parity is not conserved in interactions involving neutrinos. Put another way, the latter possibility implies a definite handedness is associated with matter and the opposite handedness with antimatter. This question is answered by appealing to experiment. It is found that neutrinos are left handed, corresponding to $\Psi_L(x)$, and that antineutrinos are right handed, corresponding to $\Psi_R(x)$.

Lecture 11

Second Quantization

A classical field may be regarded as a collection of oscillators. This is most easily visualized for a finite system with a discrete set of eigenmodes. For a box with sides of length L, each eigenmode is characterized by wavenumbers k_x, k_y, k_z each equal to an integral multiple of $2\pi/L$. Each eigenmode may be regarded as a separate harmonic oscillator. *Second quantization*, as quantization of the field is called, corresponds to describing these oscillators as quantum mechanical systems. In the continuum limit a quantized field is regarded as a continuous set of quantized harmonic oscillators.

1. The Harmonic Oscillator

The Hamiltonian for an harmonic oscillator is

$$\hat{H} = \frac{1}{2m}\,\hat{p}^2 + \frac{k}{2}\,\hat{q}^2, \tag{11.1}$$

where m and k are constants. The natural frequency of the oscillator is $\omega = \sqrt{k/m}$.

As is well known, the states of the oscillator may be labeled by a quantum number $n = 0, 1, \ldots$, with the energy E_n of the nth state given by (in natural units)

$$E_n = (n + \tfrac{1}{2})\,\omega. \tag{11.2}$$

The nth state is described by the ket $|n\rangle$. The operators

$$\hat{a} = \frac{1}{\sqrt{2\omega}}\,(\omega\hat{q} + i\hat{p}) \quad \text{and} \quad \hat{a}^\dagger = \frac{1}{\sqrt{2\omega}}\,(\omega\hat{q} - i\hat{p}) \tag{11.3}$$

act as lowering and raising operators, respectively. They satisfy the commutation relations

$$[\hat{a}, \hat{a}] = 0, \quad [\hat{a}^\dagger, \hat{a}^\dagger] = 0, \quad [\hat{a}, \hat{a}^\dagger] = 1. \tag{11.4}$$

One has

$$\hat{a}\,|\,n\rangle = \sqrt{n}\,|\,n-1\rangle, \qquad \hat{a}^\dagger\,|\,n\rangle = \sqrt{n+1}\,|\,n+1\rangle. \tag{11.5}$$

The ket $|\,n\rangle$ may be constructed from the ground state $|\,\rangle$ by operating n times with the raising operator:

$$|\,n\,\rangle = \frac{1}{\sqrt{n!}}\,(\hat{a}^\dagger)^n|\,\rangle. \tag{11.6}$$

The number operator is identified as

$$\hat{n} = \hat{a}^\dagger\hat{a}, \qquad \hat{n}\,|\,n\,\rangle = n\,|\,n\,\rangle, \tag{11.7}$$

and the Hamiltonian may be written

$$\hat{H} = \omega(\hat{n} + \tfrac{1}{2}). \tag{11.8}$$

2. Anticommutation Relations for Fermion Fields

In the application to quantized fields one reinterprets the raising and lowering operators as creation and annihilation operators for the field quanta. Let the quantum numbers of a state be described collectively by q, with ε_q the energy of the state. The creation and annihilation operators for the state are written $\hat{a}_q^\dagger$ and $\hat{a}_q$ respectively. These satisfy the commutation relations

$$[\hat{a}_q, \hat{a}_{q'}] = 0, \quad [\hat{a}_q^\dagger, \hat{a}_{q'}^\dagger] = 0, \quad [\hat{a}_q, \hat{a}_{q'}^\dagger] = \delta_{qq'}. \tag{11.9}$$

The commutation relations (11.9) are satisfactory for boson fields, but a problem arises with fermion fields: the Pauli exclusion principle implies that there can be no more than one fermion in a given state. This condition may be written formally as

$$\hat{n}_q^2 = \hat{n}_q, \tag{11.10}$$

so that $\hat{n}_q$ has only the eigenvalues $n_q = 0, 1$. In order to satisfy this condition it is postulated that fermion operators anticommute rather than commute. The anticommutator of two operators $\hat{A}$ and $\hat{B}$ is defined by

$$[\hat{A}, \hat{B}]_+ := \hat{A}\hat{B} + \hat{B}\hat{A}. \tag{11.11}$$

Thus it is postulated that the relations (11.9) are replaced by

$$[\hat{a}_q, \hat{a}_{q'}]_+ = 0, \quad [\hat{a}_q^\dagger, \hat{a}_{q'}^\dagger]_+ = 0, \quad [\hat{a}_q, \hat{a}_{q'}^\dagger]_+ = \delta_{qq'}. \tag{11.12}$$

for fermion fields. One then has

$$(\hat{a}_q)^2 = 0, \qquad (\hat{a}_q^\dagger)^2 = 0, \tag{11.13}$$

and hence

$$(\hat{n}_q)^2 = \hat{a}_q^\dagger \hat{a}_q \hat{a}_q^\dagger \hat{a}_q = -(\hat{a}_q^\dagger)^2 (\hat{a}_q)^2 + \hat{a}_q^\dagger \hat{a}_q = \hat{n}_q, \tag{11.14}$$

as required. The relations (11.5) to (11.8) are unaffected by the change from commutation to anticommutation relations.

3. Quantization of Fields

Quantizing a field involves interpreting the wave function as a sum of creation and annihilation operators. Formally this involves introducing a new Hilbert space in which these operators act. The kets in this Hilbert space correspond to oscillator states for each eigenmode of oscillation. (There is a countably infinite number of eigenmodes for a finite system; in the continuum limit assumed here the number is uncountable.) Each eigenmode for a particle system is labeled by its 3-momentum $\boldsymbol{p}$ and its spin states (for $S \neq 0$). Antiparticles are regarded as distinct from particles (except for π^0 mesons, photons, etc., that are their own antiparticles). The label ϵ for the sign of the energy distinguishes two

pairs of creation and annihilation operators. Each eigenmode for a wave system is labeled by its mode M and its wave vector $\boldsymbol{k}$.

It is conventional to use $\hat{a}^\dagger, \hat{a}$ for particles $\hat{b}^\dagger, \hat{b}$ for antiparticles and $\hat{c}^\dagger, \hat{c}$ for photons. Here *photon* is used in the generic sense to mean a quantum of any wave field. Three specific types of field are of particular relevance: these are for spin 0 particles, for spin $\frac{1}{2}$ particles and for photons in an arbitrary mode M. While the spin 0 case is of little physical interest in itself, it is a useful supplement to the spin $\frac{1}{2}$ case. Quantum electrodynamics with spin 0 particles, called *scalar electrodynamics*, is often considerably simpler than the spin $\frac{1}{2}$ case, and comparison of the spin 0 and the spin $\frac{1}{2}$ cases is of interest in that it allows one to identify the role that the spin plays in any specific process.

The positive energy terms in the wavefunction are associated with the annihilation operator for a particle. These are the terms proportional to $\exp(-i\varepsilon_q t)$. The negative frequency terms in the wavefunction are associated with the creation operator for an antiparticle. Both annihilation of a particle and creation of an antiparticle change the quantum numbers for the particle/antiparticle system in the same way. The positive and negative frequency terms in the adjoint wavefunction are associated with annihilation of an antiparticle and creation of a particle respectively.

The normalization of the wavefunctions is fixed to correspond to one particle (or antiparticle) in the normalization volume. The reason for this is that the actual level of excitation is now included in the oscillator states. The solutions $\varphi(\epsilon\boldsymbol{p})$ introduced in Lecture 4 for particle/antiparticle states are already normalized appropriately. For photon states the appropriate normalization is (10.7).

The second quantization procedure for an arbitrary particle/antiparticle field is as follows. Let the quantum numbers be ϵ and the set q. The eigenstates $\Psi_q^\epsilon(\boldsymbol{x})$ are assumed to be normalized to one particle or antiparticle in the normalization volume, that is to an energy ε_q in the normalization volume (cf. Lecture 4). The wavefunction is then written as the operator (in the Hilbert space of the oscillator states)

$$\hat{\Psi}(x) = \sum_q \Big[\hat{a}_q\, \Psi_q^+(\boldsymbol{x})\, e^{-i\varepsilon_q t} + \hat{b}_q^\dagger\, \Psi_q^-(\boldsymbol{x}) e^{i\varepsilon_q t}\Big]. \tag{11.15}$$

The adjoint wavefunction, denoted by $\dagger$, is written

$$\hat{\Psi}^\dagger(x) = \sum_q \Big[\hat{a}_q^\dagger\, \Psi_q^{+\dagger}(\boldsymbol{x})\, e^{i\varepsilon_q t} + \hat{b}_q\, \Psi_q^{-\dagger}(\boldsymbol{x}) e^{-i\varepsilon_q t}\Big]. \tag{11.16}$$

4. Free Particle States

As mentioned above, for plane wavefunctions the quantum numbers can be made discrete by choosing a box normalization and constructing the box eigenfunctions. In the continuum limit the sum over states and the Kronecker delta in (11.9) are replaced according to

$$\begin{gathered}
\sum_q \to V \sum_s \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}, \qquad \delta_{qq'} \to \delta_{ss'} \frac{(2\pi)^3}{V}\, \delta^3(\boldsymbol{p}-\boldsymbol{p}'). \\
\Psi_q^\epsilon(\boldsymbol{x})\, \exp[-i\epsilon\varepsilon_q t] \to \varphi(\epsilon\boldsymbol{p})\, \exp[-i\epsilon(\varepsilon t - \boldsymbol{p}\cdot\boldsymbol{x})] \\
\Psi_q^{\epsilon\dagger}(\boldsymbol{x})\, \exp[i\epsilon\varepsilon_q t] \to \varphi^\dagger(\epsilon\boldsymbol{p})\, \exp[i\epsilon(\varepsilon t - \boldsymbol{p}\cdot\boldsymbol{x})]
\end{gathered} \tag{11.17}$$

For particles with nonzero spin, a label s for the spin is needed on the wavefunctions $\varphi(\epsilon\boldsymbol{p}) \to \varphi_s(\epsilon\boldsymbol{p})$.

The specific forms for free fields of interest here are as follows. For spin 0 particle/antiparticle states, (4.1) with (4.5) is replaced by

$$\hat{\Psi}(x) = V\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\,\frac{1}{\sqrt{2\varepsilon V}}\,\Big[\hat{a}(\boldsymbol{p})\,e^{-ipx} + \hat{b}^\dagger(\boldsymbol{p})\,e^{ipx}\Big], \tag{11.18}$$

and the adjoint (the complex conjugate in this case) is replaced by

$$\hat{\Psi}^*(x) = V\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\,\frac{1}{\sqrt{2\varepsilon V}}\,\Big[\hat{a}^\dagger(\boldsymbol{p})\,e^{ipx} + \hat{b}(\boldsymbol{p})\,e^{-ipx}\Big]. \tag{11.19}$$

For spin $\frac{1}{2}$ particles, (4.1) with (4.11) and (4.17) is replaced by

$$\hat{\Psi}(x) = V\sum_s \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\,\frac{1}{\sqrt{2\varepsilon V}}\,\Big[\hat{a}_s(\boldsymbol{p})\,u_s(\boldsymbol{p})\,e^{-ipx} + \hat{b}_s^\dagger(\boldsymbol{p})\,v_s(\boldsymbol{p})\,e^{ipx}\Big], \tag{11.20}$$

and the Dirac adjoint is replaced by

$$\hat{\overline{\Psi}}(x) = V\sum_s \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\,\frac{1}{\sqrt{2\varepsilon V}}\,\Big[\hat{a}_s^\dagger(\boldsymbol{p})\,\bar{u}_s(\boldsymbol{p})\,e^{ipx} + \hat{b}_s(\boldsymbol{p})\,\bar{v}_s(\boldsymbol{p})\,e^{-ipx}\Big]. \tag{11.21}$$

For a wave field, the field (10.4) is replaced by

$$\hat{A}^\mu_M(x) = V\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\,\sqrt{\frac{\mu_0 R_M(\boldsymbol{k})}{\omega_M(\boldsymbol{k})V}}\,\Big[\hat{c}_M(\boldsymbol{k})\,e^\mu_M(\boldsymbol{k})\,e^{-ik_M x} + \hat{c}^\dagger_M(\boldsymbol{k})\,e^*_M{}^\mu\,e^{ik_M x}\Big], \tag{11.22}$$

where (10.7) is used. Note that in last lecture the energy in the wave field is defined per elemental range $Vd^3\boldsymbol{k}/(2\pi)^3$ so that the sum over photon states includes a power of V and the normalization factor (10.7) for photon states includes a factor $\sqrt{V}$ in the denominator. This ensures that the expression (11.22) for the photon field operator has the powers of V appearing in the same way as for the particle field operators.

5. Normal Ordering

There is ambiguity in the order that the creation and annihilation operators appear in quantities that are bilinear in the field and its adjoint. Consider the action, or more specifically the occupation numbers (5.20) for spin 0 and (5.21) for spin $\frac{1}{2}$. After taking the normalizations into account, these imply occupation number operators

$$n^+(\boldsymbol{p}) = \hat{a}^\dagger(\boldsymbol{p})\hat{a}(\boldsymbol{p}), \qquad n^-(\boldsymbol{p}) = \hat{b}(\boldsymbol{p})\hat{b}^\dagger(\boldsymbol{p}),$$

for spin 0, and

$$n_s^+(\boldsymbol{p}) = \hat{a}_s^\dagger(\boldsymbol{p})\hat{a}_s(\boldsymbol{p}), \qquad n_s^-(\boldsymbol{p}) = -\hat{b}_s(\boldsymbol{p})\hat{b}_s^\dagger(\boldsymbol{p}),$$

for spin $\frac{1}{2}$, where a label s for the spin states is now included. These operators are written in the order that is implied by the order in which the wavefunctions appear in (5.20) and (5.21). However this order is arbitrary because the wavefunctions can be freely interchanged before the second quantization is performed. It is desirable to have an unambiguous prescription for how the operators are to be ordered when second quantizing a bilinear form.

A related problem concerns the fact that the action operator is to be interpreted as a number operator, but the order of the operators for antiparticles in the second quantized version of (5.20) or (5.21) is incorrect; specifically, it is $\hat{b}\hat{b}^\dagger$ rather than $\hat{b}^\dagger\hat{b}$. One can reverse the order by using the commutation or anticommutation relations, but doing so adds a unit term. This unit term gives an infinite contribution when one sums over the continuum of states.

These problems are overcome by defining the vacuum to be the state $|\,0\,\rangle$ that gives zero when operated on by any annihilation operator, and by requiring that any operator whose vacuum expectation value should be zero be in *normal order*. Normal order (denote by : :) corresponds to all creation operators to the left of all annihilation operators. The anticommutation relations must be taken into account in the reordering for fermion operators. One has

$$: \hat{A}\hat{B}^\dagger := \begin{cases} \hat{B}^\dagger\hat{A} & \text{for bosons,} \\ -\hat{B}^\dagger\hat{A} & \text{for fermions.} \end{cases} \tag{11.23}$$

The number operators implied by (11.19) for spin 0 are

$$n^+(\boldsymbol{p}) = : \hat{a}^\dagger(\boldsymbol{p})\hat{a}(\boldsymbol{p}) : = \hat{a}^\dagger(\boldsymbol{p})\hat{a}(\boldsymbol{p}), \qquad n^-(\boldsymbol{p}) = : \hat{b}(\boldsymbol{p})\hat{b}^\dagger(\boldsymbol{p}) : = \hat{b}^\dagger(\boldsymbol{p})\hat{b}(\boldsymbol{p}). \tag{11.24}$$

The number operators implied by (11.20) for spin $\frac{1}{2}$ are

$$n_s^+(\boldsymbol{p}) = : \hat{a}_s^\dagger(\boldsymbol{p})\hat{a}_s(\boldsymbol{p}) : = \hat{a}_s^\dagger(\boldsymbol{p})\hat{a}_s(\boldsymbol{p}), \qquad n_s^-(\boldsymbol{p}) = - : \hat{b}_s(\boldsymbol{p})\hat{b}_s^\dagger(\boldsymbol{p}) : = \hat{b}_s^\dagger(\boldsymbol{p})\hat{b}_s(\boldsymbol{p}). \tag{11.25}$$

The particle current involves a bilinear form, and is to be interpreted in terms of normal products. Thus for spin 0 (5.7) leads to

$$\hat{J}^\mu(x) = q\ : \hat{\Psi}^*(x)\hat{p}^\mu\hat{\Psi}(x) :\ . \tag{11.26}$$

For spin $\frac{1}{2}$ (5.11) leads to

$$\hat{J}^\mu(x) = q\ : \hat{\overline{\Psi}}(x)\gamma^\mu\hat{\Psi}(x) :\ . \tag{11.27}$$

Lecture 12

Propagators

A *propagator* is a Greens function, which is a solution of an inhomogeneous equation with an idealized source term (a δ-function in coordinate space or a unit source term in momentum space).

1. Solution of the Inhomogeneous Particle Equations

The particle propagators are derived by solving the inhomogeneous Klein Gordon equation for spin zero, and the inhomogeneous Dirac equation for spin $\frac{1}{2}$. The photon propagator is found by solving the inhomogeneous wave equation. Alternatively, propagators may be derived from vacuum expectation values. This alternative method is important because it is in terms of vacuum expectation values that the propagators appear in the expansion of the S-matrix.

The particle propagator in coordinate space $G(x, x')$ satisfies the inhomogeneous Klein Gordon equation for spin 0 or the inhomogeneous Dirac equation for spin $\frac{1}{2}$:

$$(\partial^\mu \partial_\mu + m^2)G(x, x') = \delta^4(x - x') \quad \text{or} \quad (i\partial\!\!\!/ - m)G(x, x') = \delta^4(x - x'). \tag{12.1}$$

One is free to separate $G(x, x')$ into a function $G(x - x')$ plus an arbitrary solution of the homogeneous equation. The simplest way to solve (12.1) is to do this and then to Fourier transform. The Fourier transform $G(p)$ of $G(x-x')$ is the propagator in momentum space. The δ-function gives unity on Fourier transforming.

The propagator enables one to solve the inhomogeneous equation with an arbitrary source term. Let the source term be $S(x)$. One can write

$$S(x) = \int d^4x'\, \delta^4(x - x')\, S(x'), \tag{12.2}$$

The combination $\int d^4x'\, S(x')$ may be regarded as an operator that converts $\delta^4(x-x')$ into $S(x)$. It then follows that the solution sought may found by applying this operator to the Greens function. That is, the solution sought is

$$\Psi(x) = \int d^4x'\, G(x - x')\, S(x'). \tag{12.3}$$

The convolution theorem for Fourier transforms (Appendix A) implies that the Fourier transform of (12.3) is

$$\Phi(p) = G(p)\, S(p), \tag{12.4}$$

where $\Phi(p)$ is the Fourier transform of $\Psi(x)$, cf. (4.1), and $S(p)$ is the Fourier transform of $S(x)$.

For spin 0 and spin $\frac{1}{2}$ $G(p)$ satisfies

$$(p^2 - m^2)G(p) = -1 \quad \text{and} \quad (p\!\!\!/ - m)G(p) = 1, \tag{12.5}$$

respectively. The propagators are

$$G(p) = \frac{-1}{p^2 - m^2} \quad \text{and} \quad G(p) = \frac{\not{p} + m}{p^2 - m^2}, \tag{12.6}$$

for spin 0 and spin $\frac{1}{2}$ respectively.

2. The Feynman Contour

The denominators of the propagators (12.6) have poles at $p^2 - m^2 = 0$. The two solutions of $p^2 - m^2 = 0$ give poles at $p^0 = \pm\varepsilon$ which correspond to particles and antiparticles, respectively. The question arises as to how one is to integrate around these poles when inverting the Fourier transform. The path along which one integrates over p^0 in inverting the Fourier transform, called the *contour* of the p^0-integration, is along the real p^0-axis, and one must decide what to do at the poles.

Note that the Cauchy integral theorem implies that the actual shape of the contour is irrelevant. The integral around any contour that does not enclose a pole is zero. Hence in evaluating the integral one can arbitrarily deform the contour, provided that in deforming the contour one never makes it cross a pole (or, more generally, a singularity of the integrand).

Thus there are only three possibilities at any given pole:

(a) One can make the contour of integration deviate above the pole;

(b) one can make makes the contour of integration deviate below the pole; or

(c) one can make the contour of integration approach the pole arbitrarily closely on either side but exclude the pole itself.

Possibility (c) is called the *Cauchy principal value* of the integral.

The causal condition, discussed in Appendix A, implies that if the pole corresponds to a particle propagating forward in time then one is to give p^0 an infinitesimal positive imaginary part $i0$. That is, one is to deform the contour above the pole. (Formally, a causal function has no poles in the upper half p^0-plane, so that one may close the contour by a large semi-circle in the upper half p^0-plane.) The argument that leads to this conclusion is expressed in Appendix A in terms of the Fourier transform of the step function. The argument is due originally to Landau who derived it using the theory of Laplace transforms. Laplace transforms and Fourier transforms are related by a rotation through $\pi/2$ in the complex space. Other than this the only difference is that the Laplace transform of a function involve the function only for $t > 0$. Hence the causal condition, that is that a response vanish at times earlier that its cause, is built into the Laplace transform in a natural way. There is a unique prescription for inverting Laplace transform, and when translated into a prescription for Fourier transforms this implies that one is to add $i0$ to p^0 to impose the causal condition.

However, although it would appear logical to choose the causal contour for all poles, this is not actually the case. The pole at negative energies describes a negative energy particle propagating forward in time if one adds $i0$ to p^0 at the pole $p^0 = \varepsilon$. Feynman suggested that it is much more convenient to regard the negative energy solutions as

positive energy particles propagating backwards in time. The reason for this will become apparent when we consider Feynman diagrams.

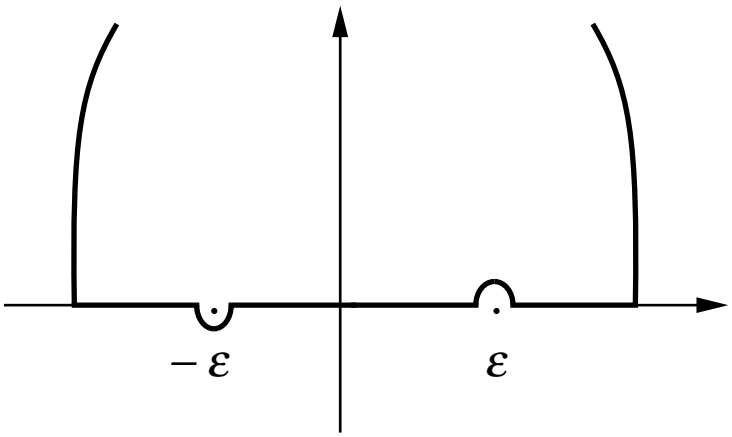

Figure 12.1 The Feynman contour in the complex p^0-plane is along the p^0-axis below the pole at $p^0 = -\varepsilon$, above the pole at $p^0 = \varepsilon$ and closing in the upper half plane.

Thus the Feynman propagator describes particles propagating forward in time and antiparticles propagating backward in time and this is achieved by adding $i0$ to p^0 at the pole $p^0 = \varepsilon$ and by adding $-i0$ to p^0 at the pole $p^0 = -\varepsilon$. Thus one writes

$$\begin{aligned}\frac{1}{p^2 - m^2} =& \frac{1}{2\varepsilon}\left(\frac{1}{p^0 - \varepsilon} - \frac{1}{p^0 + \varepsilon}\right) \\ \rightarrow& \frac{1}{2\varepsilon}\left(\frac{1}{p^0 - \varepsilon + i0} - \frac{1}{p^0 + \varepsilon - i0}\right) = \frac{1}{p^2 - m^2 + i0}.\end{aligned} \tag{12.7}$$

This contour is illustrated in Figure 12.1.

The singular terms, implied by the terms $\pm i0$ and the Plemelj formula (A.17) of Appendix A, in the denominators in (12.7) give the resonant part of the propagator. The resonant part is associated with the creation and annihilation of a real particle.

4. The Chronological Operator

The propagator $G(x, x')$ may be interpreted in terms of creation at x' of a particle that propagates to x where it is annihilated. This interpretation applies provided that t is later than t'. For t earlier than t' the Feynman propagator is interpreted in terms of creation at x of an antiparticle that propagates to x' where it is annihilated.

The following argument suggests that the propagator can be expressed as a vacuum expectation value. The argument is given here for spin $\frac{1}{2}$; the corresponding argument for spin 0 is similar. Consider the product of second quantized Dirac wavefunction $\hat{\Psi}(x)\hat{\overline{\Psi}}(x')$. This contains four terms each of which involves

the product of two creation or annihilation operators. One of these terms involve two creation operators $\hat{a}\hat{b}$ and another contains two annihilation operators $\hat{b}^\dagger\hat{a}^\dagger$. These two terms give zero when operating on the vacuum. The other terms contain operators $\hat{a}\hat{a}^\dagger$ and $\hat{b}^\dagger\hat{b}$. Provided that the operators $\hat{a}\hat{a}^\dagger$ act only for t later than t', they correspond to the required creation of an electron at one point and annihilation of the electron at another point at the later time. However the operators $\hat{b}^\dagger\hat{b}$, that act for t' later than t, give zero when operating on the vacuum. This term could be interpreted in terms of creation and annihilation of a positron provided that the operators were in the opposite order.

To achieve this result, one introduces the chronological operator $\hat{\mathcal{T}}$, which is defined by

$$\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\Psi}^*(x')\} = \begin{cases} \hat{\Psi}(x)\hat{\Psi}^*(x') & \text{for } t > t', \\ \hat{\Psi}^*(x')\hat{\Psi}(x) & \text{for } t < t', \end{cases} \tag{12.8}$$

for spin 0, and

$$\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\overline{\Psi}}(x')\} = \begin{cases} \hat{\Psi}(x)\hat{\overline{\Psi}}(x') & \text{for } t > t', \\ -\hat{\overline{\Psi}}(x')\hat{\Psi}(x) & \text{for } t < t', \end{cases} \tag{12.9}$$

for spin $\frac{1}{2}$.

5. The Vacuum Expectation Value

The chronologically ordered products (12.8) or (12.9) may be rewritten using the step functions $H(t-t')$ and $H(t'-t)$ to isolate the terms that apply for $t > t'$ and $t < t'$ respectively. For wavefunctions of the general form (11.15) and (11.16), the vacuum expectation values of these products become

$$\begin{aligned}\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\Psi}^*(x')\}|\, 0\,\rangle = \sum_{qq'} \Big\{ H(t-t')\Big[\langle\, 0\,|\hat{a}_q\hat{a}^\dagger_{q'}|\, 0\,\rangle\, \Psi^+_q(\boldsymbol{x})\Psi^{+*}_{q'}(\boldsymbol{x}')\, e^{-i\varepsilon_q t + i\varepsilon_{q'}t'}\Big] \\ + H(t'-t)\Big[\langle\, 0\,|\hat{b}_{q'}\hat{b}^\dagger_q|\, 0\,\rangle \Psi^-_q(\boldsymbol{x})\Psi^{-*}_{q'}(\boldsymbol{x}')\, e^{i\varepsilon_q t - i\varepsilon_{q'}t'}\Big]\Big\},\end{aligned} \tag{12.10}$$

for spin 0, and

$$\begin{aligned}\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\overline{\Psi}}(x')\}|\, 0\,\rangle = \sum_{qq'} \Big\{ H(t-t')\Big[\langle\, 0\,|\hat{a}_q\hat{a}^\dagger_{q'}|\, 0\,\rangle\, \Psi^+_q(\boldsymbol{x})\overline{\Psi}^{\,+}_{q'}(\boldsymbol{x}')\, e^{-i\varepsilon_q t + i\varepsilon_{q'}t'}\Big] \\ - H(t'-t)\Big[\langle\, 0\,|\hat{b}_{q'}\hat{b}^\dagger_q|\, 0\,\rangle \Psi^-_q(\boldsymbol{x})\overline{\Psi}^{\,-}_{q'}(\boldsymbol{x}')\, e^{i\varepsilon_q t - i\varepsilon_{q'}t'}\Big]\Big\},\end{aligned} \tag{12.11}$$

for spin $\frac{1}{2}$. Using the relations

$$\langle\, 0\,|\hat{a}_q\hat{a}^\dagger_{q'}|\, 0\,\rangle = \delta_{qq'} = \langle\, 0\,|\hat{b}_{q'}\hat{b}^\dagger_q|\, 0\,\rangle, \tag{12.12}$$

these reduce to

$$\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\Psi}^*(x')\}|\,0\,\rangle = \sum_q \Big\{ H(t-t')\Psi_q^+(\boldsymbol{x})\Psi_q^{+*}(\boldsymbol{x}')\, e^{-i\varepsilon_q(t-t')} + H(t'-t)\Psi_q^-(\boldsymbol{x})\Psi_q^{-*}(\boldsymbol{x}')\, e^{i\varepsilon_q(t-t')} \Big\}, \tag{12.13}$$

for spin 0, and

$$\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\overline{\Psi}}(x')\}|\,0\,\rangle = \sum_q \Big\{ H(t-t')\Psi_q^+(\boldsymbol{x})\overline{\Psi}_q^{\,+}(\boldsymbol{x}')\, e^{-i\varepsilon_q(t-t')} - H(t'-t)\Psi_q^-(\boldsymbol{x})\overline{\Psi}_q^{\,-}(\boldsymbol{x}')\, e^{i\varepsilon_q(t-t')} \Big\}, \tag{12.14}$$

for spin $\frac{1}{2}$.

6. The Propagator as a Vacuum Expectation Value

The vacuum expectation values (12.13) and (12.14) for free particles may be related to the propagators as follows. First one expresses the step functions in terms of their Fourier transforms using (A.12) of Appendix A. Specifically one writes

$$H(t-t') = \int \frac{d\omega}{2\pi}\, \frac{i}{\omega+i0}\, e^{-i\omega(t-t')}, \qquad H(t'-t) = \int \frac{d\omega}{2\pi}\, \frac{-i}{\omega-i0}\, e^{-i\omega(t-t')}. \tag{12.15}$$

Then (12.13) and (12.14) give

$$\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\Psi}^*(x')\}|\,0\,\rangle = \sum_q \int \frac{d\omega}{2\pi}\, e^{-i\omega(t-t')} \Big[\frac{i}{\omega+i0}\, \Psi_q^+(\boldsymbol{x})\Psi_q^{+*}(\boldsymbol{x}')\, e^{-i\varepsilon_q(t-t')} + \frac{-i}{\omega-i0}\, \Psi_q^-(\boldsymbol{x})\Psi_q^{-*}(\boldsymbol{x}')\, e^{i\varepsilon_q(t-t')} \Big], \tag{12.16}$$

for spin 0, and

$$\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\overline{\Psi}}(x')\}|\,0\,\rangle = \sum_q \int \frac{d\omega}{2\pi}\, e^{-i\omega(t-t')} \Big[\frac{i}{\omega+i0}\, \Psi_q^+(\boldsymbol{x})\overline{\Psi}_q^{\,+}(\boldsymbol{x}')\, e^{-i\varepsilon_q(t-t')} - \frac{-i}{\omega-i0}\, \Psi_q^-(\boldsymbol{x})\overline{\Psi}_q^{\,-}(\boldsymbol{x}')\, e^{i\varepsilon_q(t-t')} \Big], \tag{12.17}$$

for spin $\frac{1}{2}$. The integrals over ω are then rewritten as integrals over p^0 such that all the time dependences are of the form $\exp[-ip^0(t-t')]$. This involves replacing

ω by $p^0 \mp \varepsilon_q$. Next one assumes plane wave solutions by making the replacements (11.17) and with ε_q replaced by $\varepsilon = \sqrt{m^2 + \boldsymbol{p}^2}$. Then (12.16) and (12.17) become

$$\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\Psi}^*(x')\}|\, 0\,\rangle = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3} \int \frac{dp^0}{2\pi}\, \frac{i}{2\varepsilon}\, e^{-ip^0(t-t')} \left\{ \frac{1}{p^0 - \varepsilon + i0}\, \exp[i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}')] + \frac{1}{p^0 + \varepsilon - i0}\, \exp[-i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}')] \right\}, \tag{12.18}$$

for spin 0, where (4.1), (4.2) and (4.5) are used, and

$$\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\overline{\Psi}}(x')\}|\, 0\,\rangle = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3} \int \frac{dp^0}{2\pi}\, \frac{i}{2\varepsilon}\, e^{-ip^0(t-t')} \left\{ \frac{\tilde{\not{p}} + m}{p^0 - \varepsilon + i0}\, \exp[i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}')] + \frac{\tilde{\not{p}} - m}{p^0 + \varepsilon - i0}\, \exp[-i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}')] \right\}, \tag{12.19}$$

for spin $\frac{1}{2}$, with $\tilde{p} = [\varepsilon, \boldsymbol{p}]$. In deriving (12.19), the following results are used: (4.1), (4.8), (4.17), (4.22) and (4.23). The final step is to write the integral over $d^3\boldsymbol{p}$ in terms of one over $d^3\boldsymbol{P}$, with $\boldsymbol{P} = \epsilon\boldsymbol{p}$.

Comparison of the resulting expressions from (12.18) and (12.19) show that they are in the form of Fourier integrals of the propagators (12.6). Thus one finds

$$G(x-x') = i\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\Psi}^*(x')\}|\, 0\,\rangle, \quad \text{or} \quad G(x-x') = -i\langle\, 0\,|\hat{\mathcal{T}}\{\hat{\Psi}(x)\hat{\overline{\Psi}}(x')\}|\, 0\,\rangle \tag{12.20}$$

for spin 0 or spin $\frac{1}{2}$ respectively.

7. The Photon Propagator

The photon propagator is defined only to within a gauge-like transformation that allows a variety of different choices. A general form of the propagator in a medium is derived in Appendix D. Here only the case of a vacuum is considered. Then the real photons have $k^2 = 0$ or, equivalently, $\omega = |\boldsymbol{k}|$.

Maxwell's equation (2.1) with (2.31) give, after Fourier transforming,

$$(k^2\, g^{\mu\nu} - k^\mu k^\nu)\, A_\nu(k) = -\mu_0\, J^\mu(k). \tag{12.21}$$

Hence the inhomogeneous equation that we need to solve for the photon Greens function $D^{\mu\nu}(k)$ is

$$(k^2\, g^{\mu\nu} - k^\mu k^\nu)\, D_{\nu\rho}(k) = -\mu_0\, g^\mu{}_\rho. \tag{12.22}$$

Although it is not difficult to solve (12.22) it is more convenient to argue as follows. The form of $D(k)$ is not unique. In view of the charge continuity relation $k_\mu J^\mu(k) = 0$, one may add to $J^\mu(k)$ an arbitrary function times k^μ without affecting the value of $A^\mu(k)$. Also, $A^\nu(k)$ in (12.21) is defined only to within a

gauge transformation, and in a gauge independent theory one may also add an arbitrary function times k^ν to $D^{\mu\nu}(k)$, as this only causes a gauge transformation of $A^\nu(k)$. Thus $D^{\mu\nu}(k)$ is defined only to within a transformation of the form

$$D^{'\mu\nu}(k) = D^{\mu\nu}(k) + \chi^\mu(k)k^\nu + k^\mu\zeta^\nu(k), \tag{12.23}$$

where $\chi(k)$ and $\zeta(k)$ are arbitrary.

We use this freedom to replace (12.22) by

$$(k^2\, g^{\mu\nu} - k^\mu k^\nu)\, D_{\nu\rho}(k) = -\mu_0 \left(g^\mu{}_\rho - \frac{k^\mu k_\rho}{k^2}\right). \tag{12.24}$$

The additional term gives zero because the condition $k_\mu J^\mu(k) = 0$ applies to every 4-current. It is simple to recognize that

$$D^{\mu\nu}(k) = \frac{\mu_0}{k^2 + i0} \left(g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\right) \tag{12.25}$$

is a solution of (12.24). All other solutions may be obtained from (12.25) by appropriate choices of the arbitrary functions in (12.23). The specific choice (12.25) is referred to as the *Landau gauge*.

Other choices for the photon propagator include

$$D^{\mu\nu}(k) = \frac{\mu_0\, g^{\mu\nu}}{k^2 + i0} \tag{12.26}$$

and

$$D^{00}(k) = -\frac{\mu_0}{\boldsymbol{k}^2}, \quad D^{0i}(k) = D^{i0}(k) = 0, \quad D^{ij}(k) = \frac{\mu_0}{k^2 + i0} \left(g^{ij} - \frac{k^i k^j}{\boldsymbol{k}^2}\right). \tag{12.27}$$

Lecture 13

The S-Matrix

So far only free fields have been considered. The interaction between fields is described by the effect of the interaction terms. In QED there is only one interaction term, which is the $J^\mu(x)A_\mu(x)$ term that is the source for the photon field in the wave equation and is the source for the Dirac field in the Dirac equation. In scalar electrodynamics (and nonrelativistic quantum electrodynamics) the detailed form of the $J^\mu(x)A_\mu(x)$ term is different and there is a further interaction term proportional to $A^2(x)$; this term arises when (3.3) is substituted into (3.2).

1. The Interaction Picture

The interaction terms may be isolated by writing the Hamiltonian in the form

$$\hat{H}(t) = \hat{H}_0 + \hat{H}_1(t), \tag{13.1}$$

where $\hat{H}_0$ describes the background system and $\hat{H}_1(t)$ describes the interaction. In the *interaction picture* the states are assumed to evolve due to $\hat{H}_1(t)$ and the operators are assumed to evolve due to $\hat{H}_0$. Thus in the absence of the interaction terms this picture is the same as the Heisenberg picture in which the states do not change with time. The evolution of the system due to the presence of the interaction terms is described by the time evolution of the states.

Let the kets, bras and operators in the interaction picture be denoted by subscript I. The equations that describe the time evolution in this picture are

$$i\frac{d}{dt}\,|\,t\,\rangle_I = \hat{H}_{1I}(t)|\,t\,\rangle_I \tag{13.2}$$

for an arbitrary state $|\,t\,\rangle_I$ and

$$i\frac{d}{dt}\,\hat{K}_I(t) = [\hat{K}_I(t), \hat{H}_0] \tag{13.3}$$

for an arbitrary operator $\hat{K}(t)$.

The S-matrix is defined as the matrix elements of the operator $\hat{S}(t, t_0)$ that transforms a state a time t_0 into a state a time t. Thus this operator satisfies

$$|\,t\,\rangle_I = \hat{S}(t, t_0)\,|t_0\rangle_I. \tag{13.4}$$

In practice one sets $t_0 = -\infty$ and $t = \infty$, and regards these as initial and final states. (Formally one should turn the interaction term on and off adiabatically, but this is of no consequence here.) The matrix element between an initial (i) and a final (f) state is

$$S_{fi} = \langle\, f\,|\,\hat{S}\,|\,i\,\rangle \tag{13.5}$$

2. Evolution of $\hat{S}$

On substituting (13.4) into (13.2) one finds that the operator $\hat{S}$ evolves according to

$$i\frac{d}{dt}\,\hat{S}(t,t_0) = \hat{H}_I(t)\,\hat{S}(t,t_0), \tag{13.6}$$

where the superscript 1 on the interaction Hamiltonian is now omitted for simplicity in writing. One may replace this differential equation by the equivalent integral equation

$$\hat{S}(t,t_0) = \hat{1} - i\int_{t_0}^{t} dt'\;\hat{H}_I(t')\hat{S}(t',t_0), \tag{13.7}$$

where the boundary condition $\hat{S}(t_0,t_0) = \hat{1}$ has been incorporated.

In practice one assumes that the interaction terms may be treated using perturbation theory, and expands in powers of $\hat{H}_I$. This gives

$$\hat{S}(t,t_0) = \hat{1} - i\int_{t_0}^{t} dt'\,\hat{H}_I(t') + (-i)^2\int_{t_0}^{t} dt'\int_{t_0}^{t'} dt''\,\hat{H}_I(t')\hat{H}_I(t'') + \ldots. \tag{13.8}$$

The actual expansion used is in powers of $\hat{A}$, and to achieve this a further expansion of $\hat{H}_I$ in power of $\hat{A}$ is required for scalar electrodynamics. For the present this complication is ignored. On setting $t_0 = -\infty$ and $t = \infty$, one has the expansion

$$\hat{S}(\infty,-\infty) = \sum_{n=0}^{\infty}\hat{S}^{(n)}, \tag{13.9}$$

with $\hat{S}^{(0)} = \hat{1}$ corresponding to no interaction, and with

$$\begin{aligned}
\hat{S}^{(1)} &= -i\int_{-\infty}^{\infty} dt\;\hat{H}_I(t),\\
\hat{S}^{(2)} &= \frac{(-i)^2}{2}\int_{-\infty}^{\infty} dt_1\int_{-\infty}^{\infty} dt_2\;\hat{\mathcal{T}}\{\hat{H}_I(t_1)\hat{H}_I(t_2)\},\\
&\;\;\vdots\\
\hat{S}^{(n)} &= \frac{(-i)^n}{n!}\int_{-\infty}^{\infty} dt_1\int_{-\infty}^{\infty} dt_2\ldots\int_{-\infty}^{\infty} dt_n\;\hat{\mathcal{T}}\{\hat{H}_I(t_1)\hat{H}_I(t_2)\ldots\hat{H}_I(t_n)\}.
\end{aligned} \tag{13.10}$$

The chronological operator has been introduced to allow all the integrals to be extended to infinity; additional numerical factors appear because, for example the integral over $-\infty < t_1 < \infty, -\infty < t_2 < t_1$ of the time ordered integrand is half the integral over $-\infty < t_1, t_2 < \infty$.

In the cases where $\hat{H}_I$ is not linear in $\hat{A}$ (or whatever other expansion parameter is relevant) one makes the additional expansion

$$\hat{H}_I = \sum_{\ell=1}^{\infty} \hat{H}_I^{(\ell)}. \tag{13.11}$$

On inserting this expansion in (13.9) one collects together the terms of each order. Thus for $n = 1, 2$ (13.9) is replaced by

$$\begin{aligned} \hat{S}^{(1)} &= -i \int_{-\infty}^{\infty} dt\ \hat{H}_I^{(1)}(t), \\ \hat{S}^{(2)} &= -i \int_{-\infty}^{\infty} dt\ \hat{H}_I^{(2)}(t) + \frac{(-i)^2}{2} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2\ \hat{\mathcal{T}}\{\hat{H}_I^{(1)}(t_1)\hat{H}_I^{(1)}(t_2)\}. \end{aligned} \tag{13.12}$$

3. Specific Interaction Terms

In QED there is only the one interaction term. It arises from the interaction Lagrangian density when the electromagnetic field is included in the Dirac Lagrangian (5.10). One makes the minimal coupling replacement (3.3), viz.

$$\partial^\mu \to D^\mu = \partial^\mu + ieA^\mu(x), \tag{13.13}$$

where the charge on the electron is $q = -e$. The interaction term in the Lagrangian implied by (5.10) is then

$$\mathcal{L}^{(1)}(x) = -e\overline{\Psi}(x)\not{A}(x)\Psi(x). \tag{13.14}$$

The interaction Hamiltonian density is equal to minus the interaction Lagrangian density, and hence after second quantization the interaction Hamiltonian is identified as

$$\hat{H}_I(t) = \int d^3\boldsymbol{x}\, \hat{\mathcal{H}}_I(x), \tag{13.15}$$

with

$$\hat{\mathcal{H}}_I(x) = -e\ :\hat{\overline{\Psi}}(x)\not{\hat{A}}(x)\hat{\Psi}(x):\ . \tag{13.15}$$

Thus in QED the nth order term in (13.9) becomes

$$\begin{aligned} \hat{S}^{(n)} = \frac{(ie)^n}{n!} \int d^4x_1 \int d^4x_2 \ldots \int d^4x_n \\ \hat{\mathcal{T}}\{:\hat{\overline{\Psi}}(x_1)\not{\hat{A}}(x_1)\hat{\Psi}(x_1):\hat{\overline{\Psi}}(x_2)\not{\hat{A}}(x_2)\hat{\Psi}(x_2):\ \ldots\ :\hat{\overline{\Psi}}(x_n)\not{\hat{A}}(x_n)\hat{\Psi}(x)_n:\}. \end{aligned} \tag{13.17}$$

In scalar electrodynamics, the corresponding results may be derived starting from the Lagrangian (5.5). In this case one finds a first and a second order term

$$\hat{\mathcal{H}}_I(x) = \hat{\mathcal{H}}_I^{(1)}(x) + \hat{\mathcal{H}}_I^{(2)}(x) \tag{13.18}$$

with

$$\hat{\mathcal{H}}_I^{(1)}(x) = -q\,\{:\hat{\Psi}^*(x)\hat{A}_\mu(x)(\hat{p}^\mu\hat{\Psi}(x)):+:(\hat{p}^\mu\hat{\Psi}^*(x))\hat{A}_\mu(x)\hat{\Psi}(x):\}, \tag{13.19}$$

and a second order term

$$\hat{\mathcal{H}}_I^{(2)}(x) = -q^2\ :\hat{\Psi}^*(x)\hat{A}^2(x)\hat{\Psi}(x):\ . \tag{13.20}$$

4. The Scattering Amplitudes T_{if} and M_{if}

In the absence of any static field the 4-momentum is conserved, so that the final 4-momentum p_f is equal to the initial 4-momentum p_i. This may be built into the scattering matrix by writing

$$S_{fi} = \delta_{fi} + i(2\pi)^4\,\delta^4(p_f - p_i)T_{fi}, \tag{13.21}$$

which defines the scattering amplitude T_{fi}.

The probability of a transition from the initial to the final state is given by the square of the S-matrix element. For a nontrivial transition (13.21) implies a probability

$$p_{i\to f} = VT\,(2\pi)^4\,\delta^4(p_f - p_i)|T_{fi}|^2, \tag{13.22}$$

where the square of the δ-function has been rewritten using (A.9) of Appendix A. Thus the probability per unit time of a transition is

$$w_{i\to f} = V\,(2\pi)^4\,\delta^4(p_f - p_i)|T_{fi}|^2. \tag{13.23}$$

It is often convenient to define another scattering amplitude M_{fi} that includes all the normalization factors for the initial and final photon and particle states. This may be written schematically as

$$T_{fi} = \prod_i a_M(\boldsymbol{k})\,\frac{1}{\sqrt{2\varepsilon V}}\prod_f a_M(\boldsymbol{k})\,\frac{1}{\sqrt{2\varepsilon V}}\,M_{fi}, \tag{13.24}$$

with, cf. (10.7),

$$a_M(\boldsymbol{k}) = \left[\frac{\mu_0 R_M(\boldsymbol{k})}{\omega_M(\boldsymbol{k})V}\right]^{1/2}, \tag{13.25}$$

where the products are over all initial and final particles, antiparticles and photons.

5. The Density of Final States

The probability per unit time (13.23) is for a transition from a given initial state to a given final state. In practice one is concerned with transitions to a continuum of final states. One then needs to multiply by the density of final states factor

$$D_f = \prod_f\left(\frac{Vd^3\boldsymbol{p}}{(2\pi)^3}\right)\prod_f\left(\frac{Vd^3\boldsymbol{k}}{(2\pi)^3}\right), \tag{13.26}$$

where the products are over all final particles (including antiparticles) and photons respectively.

Lecture 14

Wick's Theorem and Induced Effects

Two formal problems are discussed in this Lecture. The first concerns the construction of S_{fi}, which involves identifying the relevant term in the operator $\hat{S}$ and then writing the creation and annihilation operators in this term in normal order. The way one identifies the S-matrix for specific processes then involves introducing contractions and relating them to propagators using Wick's Theorem. The second formal problem concerns induced effects, resulting from the statistical probability of the presence of particles and photons in the initial and final states.

1. Identification of S_{fi} for Specific Processes

Any given initial state $|\, i \,\rangle$ and given final state $\langle\, f \,|$ may be constructed from the vacuum by using the relevant creation operators. Thus one has

$$|\, i \,\rangle = \Big(\prod_i \hat{a}^\dagger \hat{b}^\dagger \hat{c}^\dagger \Big) |\, 0 \,\rangle, \qquad \langle\, f \,| = \langle\, 0 \,| \Big(\prod_f \hat{a}\, \hat{b}\, \hat{c} \Big) \tag{14.1}$$

where the products are over all particles antiparticles and photons in the initial and final states. The S-matrix element (13.5), viz.,

$$S_{fi} = \langle\, f \,|\, \hat{S} \,|\, i \,\rangle \tag{14.2}$$

gives a nonzero result only if $\hat{S}$ includes annihilation operators $\hat{a}$, $\hat{b}$, $\hat{c}$ that correspond one-to-one to the creation operators in $|\, i \,\rangle$, and creation operators $\hat{a}^\dagger$, $\hat{b}^\dagger$, $\hat{c}^\dagger$ that correspond one-to-one to the annihilation operators in $\langle\, f \,|$.

Consider the terms in $\hat{S}$ that do contain the relevant operators. These operators may be arranged in normal order. For each annihilation operator in (14.3) there is a creation operator in the construction (14.1) of $|\, i \,\rangle$. These pairs of operators $\hat{a}\hat{a}^\dagger$ may be rewritten using the commutation or anticommutation relation, giving

$$\hat{a}\, \hat{a}^\dagger = 1 \pm \hat{a}^\dagger \hat{a}.$$

The operator $\hat{a}$ operates on $|\, 0 \,\rangle$ and so gives zero. It follows that one has

$$\hat{S} = S_{fi} \Big(\prod_f \hat{a}^\dagger \hat{b}^\dagger \hat{c}^\dagger \Big) \Big(\prod_i \hat{a}\, \hat{b}\, \hat{c} \Big) + \ldots, \tag{14.3}$$

where $\ldots$ implies all the other terms in $\hat{S}$ that do not contain the correct number and kind of creation and annihilation operators for the given initial and final states. Thus once a process is specified in terms of the number and kind of particles and photons in the initial and final states, then S_{fi} is identified as the coefficient of the term is $\hat{S}$ with the appropriate number and kind of annihilation (for initial particles and photons) and creation (for final particles and photons) operators written in normal order.

2. Contractions

The operator $\hat{S}$ involves an integral over a chronologically ordered products of interaction Hamiltonian densities, cf. (13.9). Each Hamiltonian density is in normal order. Thus the first order term in $\hat{S}$ are already in the required form (14.3). However the second and higher order terms in $\hat{S}$ are not in this form.

To rewrite any of the higher order term in $\hat{S}$ in the form (14.3) requires that all operators be rewritten in normal order. A pair of operators corresponding to one type of particle may be commuted with a pair of operators corresponding to any other type of particle. In this way it is trivial to rearrange any higher order term in $\hat{S}$ so that it is in the form of a product of terms each of which is a chronologically ordered product of operators corresponding to one type of particle. Thus the only nontrivial part of the rearrangement of the order of the operators in $\hat{S}$ involves the rearrangement into normal order of a chronologically ordered product of operators corresponding to one type of particle.

Consider a second order term involving the chronologically ordered product of two interaction Hamiltonians. The particle operators appear in pairs, $\hat{\Psi}(x)$ and $\hat{\Psi}^\dagger(x)$, in each interaction Hamiltonian. The operators $\hat{\Psi}(x)$ and $\hat{\Psi}(x')$ at two different space-time points commute or anticommute depending on whether they describe bosons or fermions, and the same applies to $\hat{\Psi}^\dagger(x)$ and $\hat{\Psi}^\dagger(x')$. However the operator and its adjoint at x and x', i.e., $\hat{\Psi}(x)$ and $\hat{\Psi}^\dagger(x')$ or $\hat{\Psi}(x')$ and $\hat{\Psi}^\dagger(x)$, do not commute or anticommute. Thus one needs to rearrange a product of an operator $\hat{A}(x)$ and its adjoint $\hat{A}^\dagger(x')$ from chronological to normal order, where $\hat{A}(x)$ corresponds either to a $\hat{\Psi}(x)$ or $\hat{\Psi}^\dagger(x)$, or to a photon field operator.

For any such $\hat{A}(x)$ one has

$$\hat{\mathcal{T}}\{\hat{A}(x)\hat{A}^\dagger(x')\} =: \hat{A}(x)\hat{A}^\dagger(x') : + \underbrace{\hat{A}(x)\hat{A}^\dagger}(x'), \tag{14.4}$$

where $\underbrace{\hat{A}(x)\hat{A}^\dagger}(x')$ is a c-number (i.e., a non-operator quantity). The argument leading to (14.4) is as follows. The rearrangement of a pair of operators from chronological to normal order involves using the commutation or anticommutation relations, so that the two differ by at most a c-number. The value of this c-number may be determined by taking the vacuum expectation value, and noting that this is zero for any normally ordered set of operators. Thus one finds

$$\underbrace{\hat{A}(x)\hat{A}^\dagger}(x') = \langle\, 0\,|\hat{\mathcal{T}}\{\hat{A}(x)\hat{A}^\dagger(x')\}|\,0\,\rangle. \tag{14.5}$$

The extension of (14.4) with (14.5) to an arbitrary number of operators is Wick's theorem. The proof is by induction. The result is stated here only for three and four operators. The extension of Wick's theorem to an arbitrary number of operators should be self evident.

The only relevant case where three operators are involved is when they are all photon operators, and one has

$$\begin{aligned}\hat{\mathcal{T}}\{\hat{A}(x_1)\hat{B}(x_2)\hat{C}(x_3)\} =: \hat{A}(x_1)\hat{B}(x_2)\hat{C}(x_3) : &+ \underbrace{\hat{A}(x_1)\hat{B}}(x_2)\hat{C}(x_3) \\ &+ \underbrace{\hat{A}(x_1)\hat{B}(x_2)\hat{C}}(x_3) + \hat{A}(x_1)\underbrace{\hat{B}(x_2)\hat{C}}(x_3).\end{aligned} \tag{14.6}$$

For four operators one has

$$\begin{aligned}\hat{\mathcal{T}}\{\hat{A}(x_1)\hat{B}(x_2)\hat{C}(x_3)\hat{D}(x_4)\} =& :\hat{A}(x_1)\hat{B}(x_2)\hat{C}(x_3)\hat{D}(x_4): \\ &+ \hat{A}(x_1)\hat{B}(x_2)\,:\hat{C}(x_3)\hat{D}(x_4): + \,:\hat{A}(x_1)\hat{B}(x_2)\hat{C}(x_3)\hat{D}(x_4): + \ldots \\ &+ \hat{A}(x_1)\hat{B}(x_2)\hat{C}(x_3)\hat{D}(x_4) + \hat{A}(x_1)\hat{B}(x_2)\hat{C}(x_3)\hat{D}(x_4) \\ &+ \hat{A}(x_1)\hat{D}(x_4)\hat{B}(x_2)\hat{C}(x_3).\end{aligned} \tag{14.7}$$

In (14.7) ... denotes the four other choices of the pair of operators over which the one contraction is performed. In the second last term it is important to preserve the order of the operators; one has

$$\hat{A}(x_1)\hat{B}(x_2)\hat{C}(x_3)\hat{D}(x_4) = \pm\hat{A}(x_1)\hat{C}(x_3)\hat{B}(x_2)\hat{D}(x_4),$$

with the + sign for bosons and the − sign for fermions. This distinction is unnecessary in the last term in (14.7) because the change of order from $\hat{A}\hat{B}\hat{C}\hat{D}$ to $\hat{A}\hat{D}\hat{B}\hat{C}$ involves an even number of interchanges and hence has the same sign whether the operators commute or anticommute.

3. First Order Processes

A specific process corresponds to specified sets of particles and photons in the initial and final states. All terms in the S-matrix that include the relevant creation and annihilation operators are to be included in S_{fi} in (14.3). In general there are one or more lowest order terms in this S_{fi} plus a hierarchy of higher order terms. The probability involves the square of the scattering amplitude, and so the various terms interfere.

In QED the interaction Hamiltonian is strictly first order. The first order term may be written in the form (14.3) by noting that there are eight terms. Using (11.15), (11.16) and (11.21), and introducing a schematic notation one has

$$\hat{\overline{\Psi}}\hat{\not{A}}\hat{\Psi} = \sum_{i=1}^{8}\hat{f}^{(i)}, \tag{14.8}$$

with

$$\begin{aligned}\hat{f}^{(1)} &\propto \overline{\Psi}\,{}^{+}\not{A}^{+}\Psi^{+}\hat{a}^{\dagger}\hat{c}^{\dagger}\hat{a} & \hat{f}^{(2)} &\propto \overline{\Psi}\,{}^{+}\not{A}^{-}\Psi^{+}\hat{a}^{\dagger}\hat{c}\,\hat{a} \\ \hat{f}^{(3)} &\propto \overline{\Psi}\,{}^{-}\not{A}^{+}\Psi^{-}\hat{b}\,\hat{c}^{\dagger}\hat{b}^{\dagger} & \hat{f}^{(4)} &\propto \overline{\Psi}\,{}^{-}\not{A}^{-}\Psi^{-}\hat{b}\,\hat{c}\,\hat{b}^{\dagger} \\ \hat{f}^{(5)} &\propto \overline{\Psi}\,{}^{+}\not{A}^{+}\Psi^{-}\hat{a}^{\dagger}\hat{c}\,\hat{b}^{\dagger} & \hat{f}^{(6)} &\propto \overline{\Psi}\,{}^{+}\not{A}^{-}\Psi^{-}\hat{a}^{\dagger}\hat{c}^{\dagger}\hat{b}^{\dagger} \\ \hat{f}^{(7)} &\propto \overline{\Psi}\,{}^{-}\not{A}^{+}\Psi^{+}\hat{b}\,\hat{c}\,\hat{a} & \hat{f}^{(8)} &\propto \overline{\Psi}\,{}^{-}\not{A}^{-}\Psi^{+}\hat{b}\,\hat{c}^{\dagger}\hat{a}\end{aligned}. \tag{14.9}$$

In vacuo all of these correspond to kinematically forbidden processes, but in a medium all are allowed in principle. (They are forbidden in vacuo because the momentum conservation condition $p \pm p' \pm k = 0$ is incompatible with $p^2 = m^2, p'^2 = m^2, k^2 = 0$.)

Process (1) corresponds to *Cerenkov emission* by an electron, and process (2) corresponds to the inverse absorption process that is referred to as *Landau damping*. Processes (3) and (4) correspond to the same processes for positrons. Process (5) corresponds to *one-photon pair creation*, and process (8) to the inverse annihilation of a pair into one photon. Process (6) corresponds to simultaneous creation of a photon and a pair, and process (7) to the simultaneous annihilation of a photon and a pair. These process are forbidden except if the photon has negative energy. Negative energy waves are possible in principle in an unstable plasma where the waves are intrinsically growing. However the conditions for such waves to be associated with pair creation are extreme.

There are higher order terms that contribute to the foregoing specific processes, and these are necessarily of odd order. The reason for this is that the processes (14.9) involve one and only one photon, and in a second (or higher even order process) there is an even number of photon operators. This is obvious by writing out the product of two terms of the form (14.8). A contraction involving two photon fields reduces this number by two, so that the even order terms can never produce a term with only one photon operator. Thus the lowest order corrections are of third order. To have the required number and kind of operators, there must be two contractions over particle fields and one contraction over photon fields. More generally the S-matrix for the processes (14.9) is of the form

$$S_{fi} = S_{fi}^{(1)} + S_{fi}^{(3)} + S_{fi}^{(5)} + \dots .$$

The probability involves the square of the S-matrix, and the lowest order correction arise from $S_{fi}^{(1)}(S_{fi}^{(3)})^*$ and its complex conjugate, the next order corrections arise from $S^{(1)}(S^{(5)})^* + (S^{(1)})^* S^{(5)}$ and $S^{(3)}(S^{(3)})^*$, and so on.

There is an ambiguity in the use of the term *order* in QED. In the foregoing discussion the *order* refers to the order of the S-matrix element in an expansion in powers of A. The other sense in which *order* is used is to refer to processes. Thus Cerenkov emission by an electron and the other seven processes implied by (14.9) are said to be first order processes. In these cases the first order process is described by a probability (or cross section) that is proportional to the square of a first order term in the S-matrix. In this case there is no ambiguity. Similarly each of the second order processes discussed below is described by a probability that is proportional to the square of a second order term in the S-matrix, and again no ambiguity arises. However the lowest order corrections to Cerenkov emission involve the product of first and third order terms in the S-matrix, and such a product is of the same order as the square of a second order term in the S-matrix. To avoid confusion, here an nth order process is defined to mean only those processes described by a probability that involves the square of an nth order term in the S-matrix. Other processes that are technically of order n in that the are described by a probability that involves the product of an mth times and an $(n-m)$th order term in the S-matrix are referred to as corrections to relevant lower order processes.

In quantum plasmadynamics (by which mean QED applied to a medium, and more specifically to a plasma) there are intrinsically new first order processes due to the nonlinear responses. The first order interaction Hamiltonian (13.20) involves three photon field operators, and each photon field operator involves a creation and an annihilation operator. Thus there are four conceivable processes, corresponding to sets of operators (in normal

order)

$$\hat{a}\,\hat{a}\,\hat{a}, \quad \hat{a}^\dagger\hat{a}\,\hat{a}, \quad \hat{a}^\dagger\hat{a}^\dagger\hat{a}, \quad \hat{a}^\dagger\hat{a}^\dagger\hat{a}^\dagger.$$

The third set corresponds to one photon in the initial state and two photons in the final state. This process is called *photon splitting* in QED and *wave decay* in plasma physics. The second set corresponds to the inverse process of coalescence of two photons into one. The first and fourth sets correspond to simultaneous annihilation and creation, respectively of three photons and these processes are only possible if one of the waves has negative energy.

4. Equal Time Contractions

Contractions are not relevant for the first order terms. It might appear that this is obvious because the interaction Hamiltonian is already assumed to be in normal form, cf. (13.15) for example. It is usually assumed that there are to be no equal time contractions (contractions of the two particle fields at the same x). However the reason for this requires further comment.

The postulate that the operators be written in normal order is so that the interaction term gives zero when operating on the vacuum. This presupposes that there is no physical effect on the vacuum that can be attributed to the interaction Hamiltonian. This is indeed the case for the interaction Hamiltonian in QED. Suppose on the contrary that it were not the case. Then one would need to consider the effect of the term involving the contraction over $\hat{\overline{\Psi}}(x)$ and $\hat{\Psi}(x)$. This term has one photon creation or annihilation operator, and so it corresponds to creation or annihilation of a single photon. However, conservation of 4-momentum requires that this photon have $k^\mu = 0$, and it is meaningless to consider photons with $k^\mu = 0$. Thus in this case the interaction term resulting from the equal time contraction describes a null effect. The neglect of equal time contractions is therefore justified in QED.

A similar conclusion is reached when on considers the processes described by the interaction Hamiltonian (13.20) which involves three photon field operators. Again, after a contraction only one photon operator remains, and the processes correspond to creation or annihilation of a photon with $k^\mu = 0$.

However, the equal time contractions do need to be considered when discussing the effect of the second order term in the interaction Hamiltonian in scalar electrodynamics, and also when considering the effect of the cubic response of the medium in quantum plasmadynamics, cf. (13.21). In both cases an equal time contraction (over two particle fields in the former case) leaves a term with two photon creation or annihilation operators. Such terms lead to a modification of the properties of photons, and need to be included in any complete discussion.

5. Second Order Processes

In QED the second order processes involve a product of two interaction Hamiltonians, and this is of the form

$$:\hat{\overline{\Psi}}(x)\not{\hat{A}}(x)\hat{\Psi}(x):\,:\hat{\overline{\Psi}}(y)\not{\hat{A}}(y)\hat{\Psi}(y):$$

If there are no contraction then such a term corresponds to unrelated first order processes occurring at x and y.

Consider the cases where one contraction is performed. If the contraction is over a pair of particle field operators, over $\hat{\Psi}(x)$ and $\hat{\overline{\Psi}}(y)$ or equivalently over $\hat{\overline{\Psi}}(x)$ and $\hat{\Psi}(y)$, then there are two remaining particle creation or annihilation operators, and two photon creation or annihilation operators. The corresponding processes are *Compton scattering* and crossed processes, i.e., processes related to Compton scattering by crossing symmetries. These processes will be discussed in a later Lecture. If the one contraction is performed over the photon fields then there are no photon creation or annihilation operators, and there remain four particle creation or annihilation operators. The corresponding processes are *Møller scattering* and crossed processes related to it.

If two contractions are performed then there are only two remaining creation and annihilation operators. If both contractions are over particle fields, then the two remaining operators are photon operators. The corresponding process corresponds to a modification of the properties of photons. This is described in terms of the *vacuum polarization*. One important point to note is that the two contracted terms can be arranged in the appropriate order only after an odd number of interchanges. In accord with the discussion following (14.7), this introduces a minus sign when the contracted terms are replaced by propagators. In the language of Feynman diagrams, the two contracted terms form a *closed fermion loop*. A minus sign is to be included in the amplitude for each closed fermion loop.

If three contractions are performed then there are no remaining operators. In the language of Feynman diagrams such terms are called *vacuum diagrams*. They are of no physical interest and, except for some formal purposes, may simply be neglected.

In scalar electrodynamics there is a second order interaction Hamiltonian given by (13.20). It is of the same form as the Compton scattering terms discussed above. It needs to be included when treating Compton scattering and the related crossed processes in scalar electrodynamics. In the language of Feynman diagrams this term is described by a *seagull diagrams*. The seagull diagram with a contraction over the particle fields needs to be included in treating the polarization of the vacuum in scalar electrodynamics. This is an example where the equal time commutator cannot be neglected.

6. The Density Matrix

Suppose we wish to include the presence of photons and electrons in the initial and final states. This is equivalent to including the effects of a medium. The medium may be described in terms of the average occupation numbers of the states.

The formal statistical description of the medium is based on a density matrix $\hat{w}$, which may be separated into factors for each species of particle and for each wave mode present in the medium. Let the particle states be denoted by the set of quantum numbers $\{\epsilon q\}$, where ϵ is the sign of the energy and q denotes the remaining quantum numbers. For simplicity only one species of particle, electron/positron species in most cases, is included explicitly in the following. Let the waves be described by their mode M and their wave 3-vector $\boldsymbol{k}$. The density matrix may then be written in the form

$$\hat{w} = \hat{w}_{\mathrm{P}}\, \hat{w}_{\mathrm{W}}, \tag{14.10}$$

where

$$\hat{w}_{\rm P} = \sum_q \prod_\epsilon w_{\epsilon q} \, |\epsilon q\rangle \, \langle \epsilon q| \tag{14.11}$$

describes the particle states, and

$$\hat{w}_{\rm W} = V \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \prod_M w_M(\boldsymbol{k}) \, |M\boldsymbol{k}\rangle \, \langle M\boldsymbol{k}| \tag{14.12}$$

describes the photon states.

This description presupposes that the density matrix is diagonal. More generally, consider an arbitrary ket $|\ \rangle$. It may be expanded in the eigenstates. Considering only the particle eigenstates for the present. The expansion is of the form

$$|\ \rangle = \sum_q \prod_\epsilon c_q^\epsilon \, |\epsilon q\rangle, \tag{14.13}$$

where the c_q^ϵ are complex numbers. The density matrix for the arbitrary state is

$$|\ \rangle \, \langle\ | = \sum_{qq'} \prod_{\epsilon\epsilon'} c_q^\epsilon c_{q'}^{*\epsilon'} \, |\epsilon q\rangle \, \langle \epsilon' q'|. \tag{14.14}$$

This general form is replaced here by the form (14.11) after assuming that the statistical average gives

$$\overline{c_q^\epsilon c_{q'}^{*\epsilon'}} = w_q^\epsilon \, \delta^{\epsilon\epsilon'} \, \delta_{qq'}, \tag{14.15}$$

where the statistical average is denoted by a bar. The treatment of the photon states is similar.

The assumption of a diagonal density matrix involves neglecting coherence effects. These are more familiar for photon states than for particle states. The neglect of coherence effects for photon states is referred to as the *random phase approximation*. The random phase approximation does not requires that the phases actually be random; it requires that they be irrelevant. An example is in a laser. The laser light is partially coherent because of the reflecting boundaries at either end of the laser cavity. The amplification process itself is random phase; it amplifies the radiation without regard to its coherence properties.

7. Statistical Averages

The statistical average over any operator $\hat{K}$ may be performed using the density matrix. If $\hat{K}$ involves particle operators, then one has

$$\overline{K} = {\rm Tr}\,(\hat{K}\hat{w}) \ := \sum_{q\epsilon} w_q^\epsilon \, \langle \epsilon q|\hat{K}|\epsilon q\rangle, \tag{14.16}$$

and if $\hat{K}$ involves operators for photons in the mode M, then one has

$$\overline{K}_M = \mathrm{Tr}\,(\hat{w}\hat{K}) := V\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, w_M(\boldsymbol{k})\,\langle M\boldsymbol{k}|\hat{K}|M\boldsymbol{k}\rangle. \tag{14.17}$$

Of particular importance are the statistical averages of the number operators. It is convenient to write the particle and antiparticle annihilation operators in terms of

$$\hat{a}_q^\epsilon = \begin{cases} \hat{a}_q, & \text{for } \epsilon = 1, \\ \hat{b}_q, & \text{for } \epsilon = -1, \end{cases} \tag{14.18}$$

with the corresponding creation operators written in terms of $\hat{a}_q^{\epsilon\dagger}$. The number operators are then

$$\hat{n}_q^\epsilon = \hat{a}_q^{\epsilon\dagger}\hat{a}_q^\epsilon \quad \text{and} \quad \hat{N}_M(\boldsymbol{k}) = \hat{c}_M^\dagger(\boldsymbol{k})\hat{c}_M(\boldsymbol{k}), \tag{14.19}$$

for particles and photons respectively. The mean occupation numbers are given by

$$n_q^\epsilon = \mathrm{Tr}\,(\hat{w}\hat{n}_q^\epsilon) \quad \text{and} \quad N_M(\boldsymbol{k}) = \mathrm{Tr}\,(\hat{w}\hat{N}_M(\boldsymbol{k})) \tag{14.20}$$

respectively. Note also that one has

$$\mathrm{Tr}\,(\hat{a}_q^\epsilon\hat{a}_q^{\epsilon\dagger}\hat{w}) = 1 \pm n_q^\epsilon \quad \text{and} \quad \mathrm{Tr}\,(\hat{c}_M(\boldsymbol{k})\hat{c}_M^\dagger(\boldsymbol{k})\hat{w}) = 1 + N_M(\boldsymbol{k}), \tag{14.21}$$

where the $+$ is for bosons and the $-$ sign is for fermions. For fermions the mean occupation number must be in the range $0 \le n_q^\epsilon \le 1$; the only restriction on the mean occupation numbers for bosons (and photons) is that they be non-negative.

8. Induced Effects

It is implicit in the forgoing treatment of specific processes, cf. (14.1) to (14.3), that the initial and final states are assumed to contain only the particles and photons specifically involved. In the presence of a medium there are statistical distributions of electrons and of photons.

The transition probability (13.23) is proportional to $|S_{fi}|^2$. The statistical average of this transition rate is of the form

$$\overline{w}_{i\to f} \propto \mathrm{Tr}(\hat{S}\,|\,i\,\rangle\,\langle\,i\,|\,\hat{S}^\dagger\,|f\rangle\,\langle\,f\,|). \tag{14.22}$$

It then follows from (14.20) applied to the initial state and (14.21) applied to the final state that the statistical average leads to additional factors

$$d_i = \prod_i n_q^\epsilon N_M(\boldsymbol{k}) \quad \text{and} \quad d_f = \prod_f (1 \pm n_q^\epsilon)\,(1 + N_M(\boldsymbol{k})), \tag{14.23}$$

where the products are over all particles and photons in the initial and final states respectively, and where the $\pm$ sign is for bosons and fermions respectively. The net probability of a transition in a medium is

$$\overline{w}_{i\to f} = w_{i\to f}\; d_i\, d_f\; D_f \tag{14.24}$$

where D_f is the density of final states factor (13.26).

In the factor $1+N_M(\boldsymbol{k})$ in (14.14) the unit term interpreted as describing *spontaneous* emission and the term $N_M(\boldsymbol{k})$ as describing *stimulated* or *induced* emission. The presence of bosons in the the final state also stimulates the transition due to the + sign in the particle factor in (14.14). However, the presence of fermions in the final state suppresses the transition. This may be attributed to the Pauli exclusion principle. If the final state for an electron or a positron is already occupied then the the transition cannot occur because it would require putting a second particle in a filled state. In the statistical description n_q^ϵ may be regarded as the probability that the fermion state is occupied, and the probability of a transition is proportional to the probability $1-n_q^\epsilon$ that the final state is unoccupied.

Lecture 15

Introductions to Feynman Diagrams

The idea in any diagrammatic approach is to set up a one-to-one correspondence between the terms in analytic expressions and elements in diagrams. In Feynman diagrams the elements correspond to terms in the QED interaction Hamiltonian density $: \hat{\overline{\Psi}}(x) \hat{\not{A}}(x) \hat{\Psi}(x) :$, and to terms in the S-matrix expansion corresponding to this interaction Hamiltonian. Additional types of diagrammatic element are required in quantum plasmadynamics to describe the interaction terms arising from the nonlinear responses of the medium.

In this Lecture and the next, rules are formulated for drawing diagrams and for writing down the *Feynman amplitude* corresponding to each diagram. These rules are such that the Feynman amplitude corresponds to the relevant term in the expansion of the S-matrix; this is either S_{fi} or M_{fi}, with the latter being relevant only for plane wave solutions. In this Lecture the electron/positron states are assumed to be arbitrary with quantum number ϵ, q and energy $\epsilon\varepsilon_q$.

1. Elements in Diagrams for QED

The elements in diagrams in QED are electron/positron lines (solid lines with arrows) photon lines (dashed lines) and vertices (dots) where these lines join. These are illustrated in Figure 15.1. Electron and positron lines are distinguished by the direction of the arrow: the convention is that on electron lines the arrow points to the left and on positron lines the arrow points to the right. The initial state corresponds to the right and the final state corresponds to the left of the diagram. Thus an incoming electron is represented by a solid line which enters from the right of the diagram and has an arrow pointing to the left, an incoming positron is represented by a solid line which enters from the right of the diagram and has an arrow pointing to the right, and an incoming photon is represented by a dashed line which enters from the right of the diagram. Final electrons, positrons and photons are represented by lines that leave the diagram on the left. *External lines*, which represent initial or final particles, have one free end and the other end is at a vertex. *Internal lines* have both ends at vertices and represent propagators between the space-time points of the two vertices. Electron propagators are represented by solid lines with arrows pointing to the left, positron propagators are represented by solid lines with arrows pointing to the right, and photon propagators are represented by dashed lines.

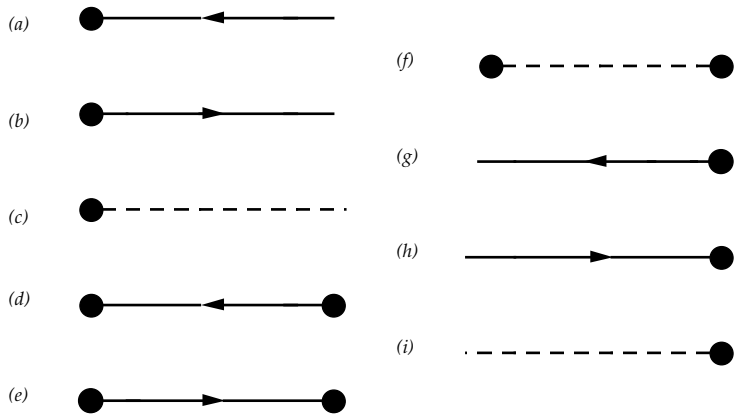

Figure 15.1 The elements in Feynman diagrams consist of: (a) incoming electron line, (b) incoming positron line, (c) incoming photon line, (d) internal electron line, (e) internal positron line, (f) internal photon line, (g) outgoing electron line, (h) outgoing positron line, and (i) outgoing photon line. At vertices, denoted by large dots for emphasis, two solid lines and a dashed line join such that the direction of the arrow along the solid line is continuous.

In QED the nth order S-matrix element is given by (14.17), viz.

$$\hat{S}^{(n)} = \frac{(ie)^n}{n!} \int d^4x_1 \int d^4x_2 \ldots \int d^4x_n \hat{\mathcal{T}}\{: \hat{\overline{\Psi}}(x_1)\hat{\not{A}}(x_1)\hat{\Psi}(x_1) : \\ : \hat{\overline{\Psi}}(x_2)\hat{\not{A}}(x_2)\hat{\Psi}(x_2) : \ \ldots \ : \hat{\overline{\Psi}}(x_n)\hat{\not{A}}(x_n)\hat{\Psi}(x)_n :\}. \tag{15.1}$$

This element in the S-matrix is represented by diagrams with n vertices. Each vertex may be labeled by the appropriate space-time coordinate. That is, the n vertices may be labeled by $x_1, x_2, \ldots x_n$.

Each vertex involves a photon operator and a γ-matrix in the combination $\hat{\not{A}}(x) = \gamma_\mu \hat{A}^\mu(x)$. It is convenient to combine the γ-matrix with one of the n numerical factors ie multiplying the element (15.1). Thus the vertices are to be labeled with 4-vector indices $\mu, \nu, \ldots$, with one index per vertex, and a term $ie\gamma_\mu$ in the Feynman amplitude is associated with the vertex labeled μ.

2. Initial and Final States

The description of the initial and final states in an arbitrary S-matrix element in QED may be understood by considering the first order terms. The elements may be written schematically, cf. (14.8) with (14.9). The first order term in the S-matrix is ie times the integral over d^4x of

$$: \hat{\overline{\Psi}}(x)\hat{\not{A}}(x)\hat{\Psi}(x) : = \sum_{\epsilon'\eta\epsilon=\pm} \sum_{q',M,q} \hat{f}^{\epsilon'\eta\epsilon}_{q'Mq}(\boldsymbol{x}) \exp[i(\epsilon'\varepsilon_{q'} - \eta\omega_M - \epsilon\varepsilon_q)t], \tag{15.2}$$

with

$$\hat{f}^{\epsilon'\eta\epsilon}_{q'Mq} = \overline{\Psi}^{\epsilon'}_{q'}(\boldsymbol{x})\not{A}^\eta_M(\boldsymbol{x})\, \Psi^\epsilon_q(\boldsymbol{x})\, \hat{a}^{\epsilon\dagger}_{q'}\hat{c}^\eta_M(\boldsymbol{k})\hat{a}^\epsilon_q, \tag{15.3}$$

with

$$A^{+\mu}_M(\boldsymbol{x}) = V\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, a_M(\boldsymbol{k})\, e^\mu_M(\boldsymbol{k})\, \exp[i\boldsymbol{k}\cdot\boldsymbol{x}], \qquad A^-_M(x) = [A^+_M(x)]^*, \tag{15.4}$$

and with

$$\hat{a}^\epsilon_q = \begin{cases} \hat{a}_q, & \text{for } \epsilon = 1, \\ \hat{b}_q, & \text{for } \epsilon = -1, \end{cases} \qquad \hat{c}^\eta_M(\boldsymbol{k}) = \begin{cases} \hat{c}_M(\boldsymbol{k}), & \text{for } \eta = 1, \\ \hat{c}^\dagger_M(\boldsymbol{k}), & \text{for } \eta = -1. \end{cases} \tag{15.5}$$

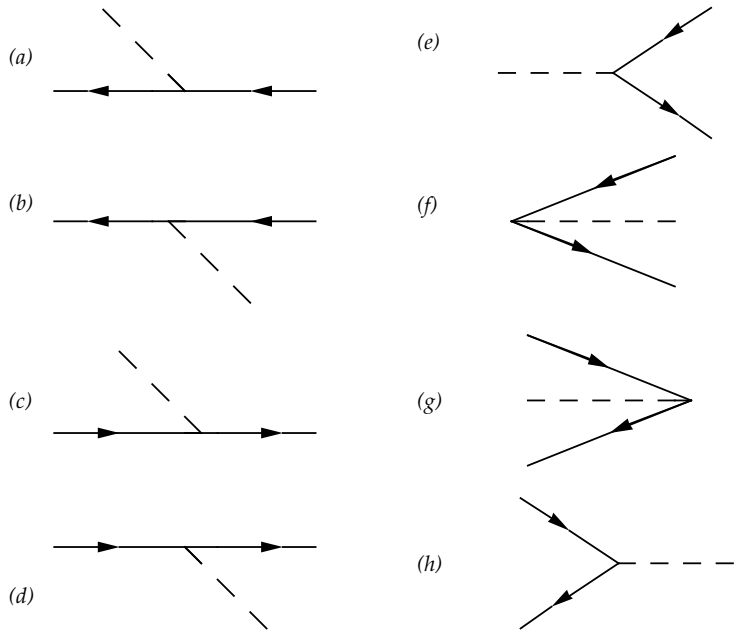

Figure 15.2 The Feynman diagrams for first order processes in QED are for: (a) Cerenkov emission by an electron, (b) Landau damping by an electron, (c) Cerenkov emission by a positron, (d) Landau damping by a positron, (e) pair annihilation into one photon, (f) simultaneous annihilation of a pair and a photon, (g) simultaneous creation of a pair and a photon, (h) decay of a photon into a pair. As discussed in §6.4, all these processes are possible in principle in a plasma.

Consider the term that describes Cerenkov emission, that is the term involving $\epsilon' = 1$, $\eta = -1$, $\epsilon = 1$. This is described by Figure 15.2a. In Cerenkov emission an electron in an initial state with quantum number q emits a photon in the mode M leaving the electron in the state with quantum numbers q'. The associated element in S_{fi} is $\overline{\Psi}^+_{q'}(\boldsymbol{x})\not{A}^-_M(\boldsymbol{x})\Psi^+_q(\boldsymbol{x})$. The wavefunctions $\overline{\Psi}^+_{q'}(\boldsymbol{x})$ and $\Psi^+_q(\boldsymbol{x})$ are 4-spinors and may be represented by a row matrix and a column matrix respectively. These are multiplied according to matrix multiplication along the direction opposite to the arrow in Figure 15.2a.

Next consider Cerenkov emission by a positron. This is described by Figure 15.2c. In this case q labels the final positron state and q' labels the initial state. The relevant element in $S_{\rm fi}$ is $\overline{\Psi}^-_{q'}(\boldsymbol{x})\not{A}^-_M(\boldsymbol{x})\Psi^-_q(\boldsymbol{x})$. The positron wavefunctions are multiplied according to matrix multiplication along the direction opposite to the arrow in Figure 15.2c. Inspection of the other terms in which there is a pair in the initial state or in the final

state (Figures 15.1e to 13.1h) shows that again the order of the wavefunctions is such that they correspond to matrix multiplication along the direction opposite to the arrow.

The inverse of Cerenkov emission is the absorption process (Landau damping) described by Figure 15.2b. The initial state of the electron is again labeled q and the final state labeled q'. However this is misleading because if $q \to q'$ corresponds to Cerenkov emission then $q \to q'$ cannot also correspond to the inverse absorption process because the final state must be different. One could choose to write $q \to q''$ in place of $q \to q'$ to take this fact into account. However it is much more convenient to consider the absorption process between the same two states as in emission. This involves a relabeling of the sums in (15.2) so that q labels the final state and q' labels the initial state. Then the absorption process corresponds to the transition $q' \to q$, with q and q' being the same states as involved in the emission process. Then by inspection the change from emission to absorption involves the replacement $\boldsymbol{k} \to -\boldsymbol{k}$.

An initial photon is represented by $a_M(\boldsymbol{k})\, e^{\mu}_M(\boldsymbol{k})$ and a final photon by the complex conjugate of this quantity, cf. (11.22). An exponential factor $\exp(\pm ik_M x)$ and the integral over $Vd^3\boldsymbol{k}/(2\pi)^3$ are also to be included in (11.22). The integral is not relevant for an initial photon, which is assumed to correspond to a unit range of $Vd^3\boldsymbol{k}/(2\pi)^3$. For a final photon the integral is included in the density of states factor (14.28). Thus the integral is to be ignored for present purposes.

3. Propagators: Internal Lines

In the S-matrix expansion propagators appear in connection with contractions. The integrand in (15.2) is to be written in normal order using Wick's theorem, and this leads to sets of terms with one, two, three, etc. contractions. A contraction involving operators at x_i and x_j is described by a line joining x_i and x_j, and this line represents the propagator. There are three types of contraction involving fields at x_i and x_j. One is a contraction over photon operators and this corresponds to the photon propagator. The other two types of contraction are over $\hat{\Psi}(x_i)$ and $\hat{\overline{\Psi}}(x_j)$ and over $\hat{\overline{\Psi}}(x_i)$ and $\hat{\Psi}(x_j)$. The operators in $\hat{\Psi}(x_i)$ either annihilate an electron or create a positron at x_i, and the operators in $\hat{\overline{\Psi}}(x_j)$ either create an electron or annihilate a positron at x_j. Thus the first of these contractions corresponds to an electron propagating from x_j to x_i or a positron propagating from x_i to x_j. These are both represented by solid lines with the arrow pointing from x_j to x_i. The other contraction corresponds to a positron propagating from x_j to x_i or an electron propagating from x_i to x_j. These are both represented by solid lines with the arrow pointing from x_i to x_j.

The Feynman propagator is defined so that one does not need to distinguish between electron and positron propagators. With positrons regarded as electrons propagating backwards in time the electron propagator from x_i to x_j is the same as the positron propagator from x_j to x_i; the propagator involves an electron for $t_j > t_i$ and a positron for $t_j < t_i$. The convention for arrows on the lines is chosen to reflect the fact that these two processes are described by the same propagator. The convention for electron and positron lines implies that the direction of the arrow along any electron/positron line is continuous irrespective of whether the propagator is associated with exchange of an electron or positron.

The relation (10.20) between the propagators and the vacuum expectation values, and

the equivalence of contractions and vacuum expectation values from Wick's theorem imply

$$\hat{\Psi}(x)\overline{\hat{\Psi}}(x') = iG(x,x'), \tag{15.6}$$

and

$$\hat{A}^{\mu}(x)\hat{A}^{\dagger\nu}(x') = -iD^{\mu\nu}(x-x'). \tag{15.7}$$

Note that in general, notably in the presence of a magnetostatic field, $G(x,x')$ depends separately on x and x', but that $D^{\mu\nu}(x-x')$ depends only on the difference $x-x'$ in all cases of interest.

If there are two or more electron propagators forming a closed fermion loop, then the arguments following (14.7) imply that an extra minus sign is to be included in S_{fi}. For the same reason, if two diagrams differ by having two external fermion lines interchanged, then their amplitudes differ by a sign.

4. External Fields

In the discussion so far, the field $A(x)$ is assumed to describe photons and to have been second quantized. In some applications the field $A(x)$ is identified as a static or slowly varying external field. Examples include the Coulomb field of a nucleus, and a uniform electric or magnetic field. An interaction term that involves such an external field is described by a vertex in which the photon line is replaced by a squiggly line connected to an "x". Such a vertex is illustrated in Figure 15.3.

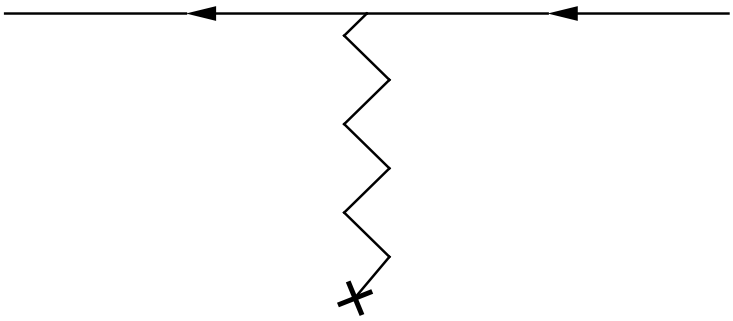

Figure 15.3 An external field is represented by a squiggly that joins an electron line at a vertex, and has its other end at an "x".

5. Multiple Photon Vertices

For completeness let me include the modifications that I have suggested to include nonlinear responses of a plasma in generalizing QED to quantum plasmadynamics there are additional interaction terms arising from quadratic and cubic nonlinear responses of the medium. The effect of the quadratic and cubic responses are included through the interaction Hamiltonians (14.21) and (14.22) respectively. The diagrammatic counterparts of these are m-photon vertices, which are shaded circles with m vertices connecting to photon lines, as illustrated in Figure 15.4. The quadratic response corresponds to a 3-photon vertex, and the cubic response to a 4-photon vertex.

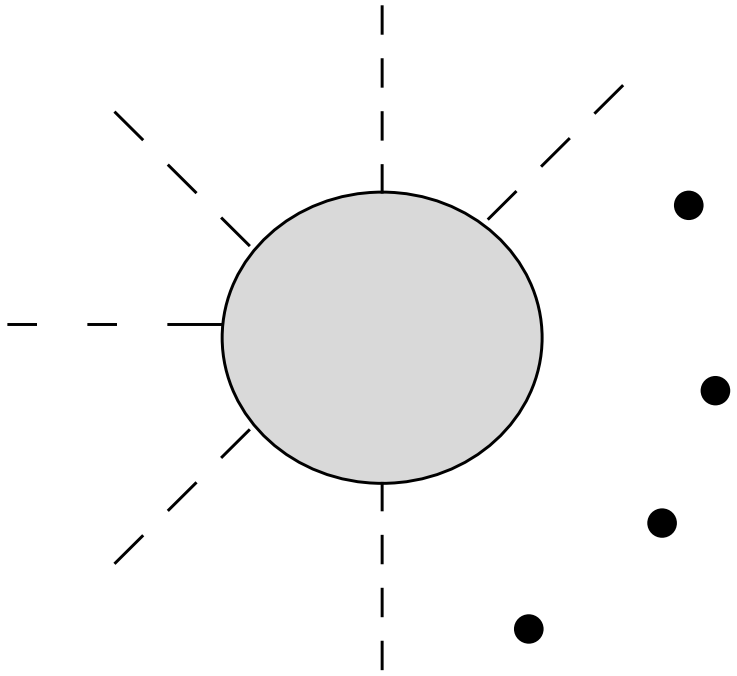

Figure 15.4 An m-photon vertex is a shaded circle with m external photon lines connecting to m vertices. Such a diagrammatic element represents a nonlinear response of the medium, with $m = 3, 4, \ldots$ corresponding to the quadratic, cubic, ... responses.

The quadratic response gives a first order term in the expansion of the S-matrix, and the cubic response gives a second order term. More generally a single m-photon vertex is to be regarded as an $(m-2)$th order element, i.e., it is of the same order as $(m-2)$ electron/photon vertices.

It is convenient to rewrite (14.21) or (14.22) in terms of a general $(m-2)$th order nonlinear interaction term. This is of the form

$$\begin{aligned}\int d^4x\, \hat{\mathcal{H}}_I^{nl(m-2)}(x) = \frac{1}{m} \int d^4x_0 \int d^4x_1 \cdots \int d^4x_m \int \frac{d^4k_0}{(2\pi)^4}\, \frac{d^4k_1}{(2\pi)^4} \cdots \frac{d^4k_m}{(2\pi)^4} \\ \alpha^{(m)\mu_0\mu_1\cdots\mu_m}(-k_0, k_1, \ldots, k_m)\hat{A}_{\mu_0}(x_0)\hat{A}_{\mu_1}(x_1)\ldots\hat{A}_{\mu_m}(x_m) \\ (2\pi)^4\, \delta^4(k_0 + k_1 + \cdots + k_m)\, \exp[i(k_0x_0 + k_1x_1 + \cdots + k_mx_m],\end{aligned} \tag{15.8}$$

where k has been rewritten as $-k_0$ to exhibit the symmetry property involving permutations of the affices $0, 1, \ldots, m$. Note that with this convention one has

$$\sum_{r=0}^{m} k_r = 0, \tag{15.9}$$

and also that the symmetry property (1.4.26) implies

$$\alpha^{(m)\mu_0\mu_1\cdots\mu_m}(-k_0, k_1, \ldots, k_m) = \alpha^{(m)\mu_1\mu_0\cdots\mu_m}(-k_1, k_0, \ldots, k_m). \tag{15.10}$$

The convention that the minus sign be associated with the first written argument results from the original definition of the nonlinear response tensor. One could redefine the nonlinear response tensors so that the minus sign does not appear at all, but this is not done here.

The contribution to S_{fi} from an m-photon vertex is $-i$ times the integral (15.8). The operators $\hat{A}(x)$ in (15.8) are to be expressed in terms of the photon wavefunctions as in the QED term, cf. (15.2) to (15.5), or in terms of an external $A(x)$ if this is appropriate.

6. Rules for Drawing Diagrams

The foregoing arguments lead to the following rules for drawing diagrams in quantum plasmadynamics.

(i) The initial state is to the right of the diagram and the final state is to the left. For a given process (specified initial and final states) all diagrams with the specified number and kind of particles and photons in the initial and final states are to be drawn.

(ii) Each vertex has a 4-tensor index $(\mu, \nu, \ldots)$ associated with it. In the coordinate space representation there is also a space-time point associated with each vertex.

(iii) An electron is represented by a solid line with an arrow pointing from right to left and a positron is represented by a solid line with an arrow pointing from left to right. The direction of the arrow along a solid line is continuous.

(iv) Photons are represented by dashed lines.

(v) A photon line begins or terminates at a vertex, which is represented by dot, either joining an electron/positron line (an *electron/photon* vertex) or an m-photon vertex.

(vi) An interaction with an external field $A(x)$ is described by a vertex with the photon line replaced by a squiggly line joined to an "**x**".

(vii) The nth order nonlinear response of the medium is represented by an $(n+1)$-photon vertex, which is a shaded circle with $(n+1)$ vertices. Each vertex has a 4-tensor index associated with it and (in the coordinate representation) a space-time point.

Some additions to these rules are given in the next Lecture, specifically in connection with diagrams in momentum space and in connection with closed fermion loops.

7. Rules for Constructing S_{fi}

Rules for writing down S_{fi} for a given diagram are as follows.

1) An nth order term in the S-matrix corresponds to a diagram with n electron/photon vertices. An m-photon vertex is regarded as an element of order $m-2$ in determining the order of the diagram. The contributions from all diagrams with the specified number and kind of initial and final particles are to be added in determining S_{fi}.

2) Each vertex corresponds to a factor $ie\,\gamma_\mu$, where μ is the 4-tensor index associated with the vertex. The integrals are to be performed over the space-time coordinates associated with each vertex.

3) Each incoming electron or positron line with quantum number q corresponds to a wavefunction $\Psi_q^+(\boldsymbol{x})\,e^{-i\varepsilon_q t}$ or $\overline{\Psi}_q^-(\boldsymbol{x})\,e^{-i\varepsilon_q t}$, respectively, and each outgoing electron or positron line with quantum number q' corresponds to a wavefunction $\overline{\Psi}_{q'}^+(\boldsymbol{x})\,e^{i\varepsilon_{q'} t}$ or $\Psi_{q'}^-(\boldsymbol{x})\,e^{i\varepsilon_{q'} t}$, respectively.

4) An incoming photon in the mode M has a factor $e_M^\mu(\boldsymbol{k})\,e^{-ik_M x}$ associated with it, where (x, μ) label the vertex to which the photon line connects. Each outgoing photon

line corresponds to a factor $e^{*\mu}_M(\boldsymbol{k})\, e^{ik_M x}$. In an interaction with an external field $A(x)$, then a factor $A^\mu(x)$ is included in place of these photon wavefunctions.

5) An internal electron/positron line pointing x_1 to x_2 corresponds to the propagator $iG(x_2, x_1)$. An internal photon line between vertices (x_1, μ) and (x_2, ν) corresponds to the propagator $-iD^{\mu\nu}(x_1 - x_2)$.

6) The Dirac spinors are written according to matrix multiplication along the direction opposite to the arrow. An extra minus sign is to be included for each closed electron/positron loop. Also, although the overall phase of is unimportant, two diagrams that differ only by the interchange of two external electron/positron lines must have opposite signs.

7) For an m-photon vertex, the m vertex factors 2) are to be omitted and replaced by a single factor

$$\frac{-i}{m}\int d^4x_0 \int d^4x_1 \cdots \int d^4x_m \int \frac{d^4k_0}{(2\pi)^4}\,\frac{d^4k_1}{(2\pi)^4}\cdots\frac{d^4k_m}{(2\pi)^4}$$
$$(2\pi)^4\,\delta^4(k_0 + k_1 + \cdots + k_m)\exp[i(k_0x_0 + k_1x_1 + \cdots + k_mx_m]$$
$$\alpha^{(m)\mu_0\mu_1\ldots\mu_m}(-k_0, k_1, \ldots, k_m).$$

Lecture 16

Diagrams in Momentum Space

It is usually more convenient to use Feynman diagrams in the the momentum space representation than in the coordinate space representation. As discussed in the Lecture 15, in the coordinate space representation the vertices are labeled with the appropriate space-time points, but in practice one is rarely interested in the relative positions of particles. In the momentum representation internal and external lines are labeled with the appropriate 4-momenta, and these are of direct relevance in most applications. The conversion to momentum space is straightforward in the absence of static fields; one simply expresses all wavefunctions and all propagators in terms of their Fourier transforms and integrates over all the space-time variable x. This procedure cannot be used in the presence of static field because the particle propagator $G(x, x')$ then depends independently on x and x' and its Fourier transform cannot be taken in the usual way.

1. Conversion to Momentum Space

The coordinate space form of any specific term in S_{fi} involves integrals over the space-time variables corresponding to all the vertices. In the absence of any static field these integrals may be carried out trivially as follows.

The initial and final particles are described by wavefunctions which vary as $\exp(\mp ipx)$. It is conventional to use the wavefunctions in the form (11.20) and (11.21), viz.

$$\hat{\Psi}(x) = V \sum_s \int \frac{d^3\boldsymbol{p}}{(2\pi)^3} \frac{1}{\sqrt{2\varepsilon V}} \Big[\hat{a}_s(\boldsymbol{p})\, u_s(\boldsymbol{p})\, e^{-ipx} + \hat{b}_s^\dagger(\boldsymbol{p})\, v_s(\boldsymbol{p})\, e^{ipx} \Big], \tag{16.1}$$

$$\hat{\overline{\Psi}}(x) = V \sum_s \int \frac{d^3\boldsymbol{p}}{(2\pi)^3} \frac{1}{\sqrt{2\varepsilon V}} \Big[\hat{a}_s^\dagger(\boldsymbol{p})\, \bar{u}_s(\boldsymbol{p})\, e^{ipx} + \hat{b}_s(\boldsymbol{p})\, \bar{v}_s(\boldsymbol{p})\, e^{-ipx} \Big]. \tag{16.2}$$

The factors $1/\sqrt{2\varepsilon V}$ are included in the definition (11.21) with (11.24) of the amplitude iM_{fi}. Similarly the photon wavefunction is written

$$\hat{A}^\mu_M(x) = V \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, a_M(\boldsymbol{k}) \Big[\hat{c}_M(\boldsymbol{k})\, e^\mu_M(\boldsymbol{k})\, e^{-ik_M x} + \hat{c}^\dagger_M(\boldsymbol{k})\, e^{*\mu}_M\, e^{ik_M x} \Big]. \tag{16.3}$$

The factors $a_M(\boldsymbol{k})$ are also included in the definition (11.21) with (11.24) of the amplitude iM_{fi}.

In the absence of a static field, the propagators may be expressed in terms of their Fourier transforms:

$$\begin{aligned} G(x, x') = G(x - x') &= \int \frac{d^4p}{(2\pi)^4}\, e^{-ip(x-x')}\, G(p), \\ D^{\mu\nu}(x - x') &= \int \frac{d^4k}{(2\pi)^4}\, e^{-ik(x-x')}\, D^{\mu\nu}(k). \end{aligned} \tag{16.4}$$

Then each space-time points x appears only in an exponential function, and the integral over x gives a δ-functions that may be interpreted as expressing conservation of 4-momentum.

Consider an electron/photon vertex at the space-time point x. There are three 4-momenta, p, p' and k say. Suppose firstly that all lines correspond to free particles. The exponential dependence has a minus sign (e.g., $\exp(-ipx)$) for initial electron, positron or photon, and a plus sign (e.g., $\exp(ipx)$) for a final electron, positron or photon. The δ-function then expresses the requirement $p = p' + k$, i.e., that 4-momentum be conserved at the vertex.

Now suppose that one or more of the lines corresponds to a propagator. Then (16.1) implies that the exponential dependence has the minus sign for the propagator from x' to x and the plus sign for the propagator from x to x'. These propagators are interpreted as carrying 4-momentum into and away from the vertex respectively, and with this interpretation the δ-function expresses the requirement that the total incoming 4-momentum balance the total outgoing 4-momentum at the vertex.

For an m-photon vertex, the contribution to S_{fi} is given by (15.8) which involves photon fields at each of the vertices. The field $A(x_i)$ at the ith vertex varies exponentially with the minus sign for an incoming photon and with the plus sign for an outgoing photon. On inserting these dependences for each photon field, the integral over each of $x_0, x_1, \ldots, x_m$ in (15.8) gives a δ-function, and the integrals over $k_0, k_1, \ldots, k_m$ in (15.8) are then trivially performed over these δ-functions. The δ-function in (15.8) remains, and this expresses the requirement $k_0 + k_1 + \cdots + k_m = 0$ with the each k_i equal either to the 4-momentum of an incoming photon or equal to minus the 4-momentum of an outgoing photon. Thus as for an electron/photon vertex, in momentum space an m-photon vertex has associated with it a δ-function that expresses conservation of 4-momentum at the vertex.

2. Loop Momenta

The foregoing arguments lead to the conclusion that in the momentum representation there are δ-functions that imply that 4-momentum is conserved at each vertex. There are also integrals over 4-momenta, with one integral over 4-momentum per propagator according to (16.4). If there are no closed loops in the diagram, then the integrals over each of these 4-momenta may be performed over one of the δ-functions expressing 4-momentum conservation at the vertices. Once all the integrals have been performed in this way, one δ-function remains. This expresses conservation of the net 4-momentum, and may be written in the for $\delta^4(p_f - p_i)$ that appears in the definition (11.21) with (11.24) of M_{fi}.

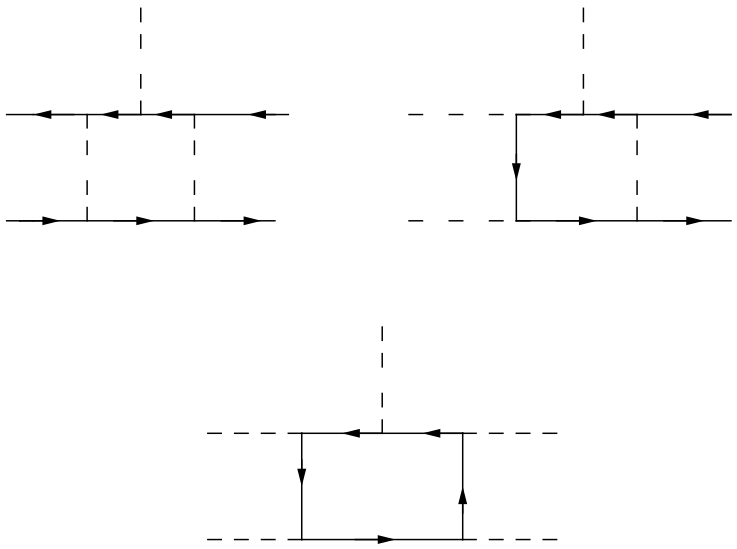

Figure 16.1 Three examples of closed loop diagrams are shown. Conservation of 4-momentum at each of the vertices determines the 4-momentum in each line around the loop only to within an additive constant called the loop momentum.

In general a diagram includes one or more closed loops. The *loop momentum* is the undetermined 4-momentum in a closed loop. The loop momentum may be identified with the 4-momentum in any line of the loop, and then the 4-momentum in any other line is determined by conservation of 4-momentum at the vertices of the loop. One is to integrate over each undetermined loop momentum. The argument for this is as follows.

First consider diagrams that involve no m-photon vertices. Let there be n vertices and g propagators. Each vertex joins three lines, so that if there were no propagators there would be $3n$ external lines. Each propagator corresponds to one internal line that joins onto two vertices. Thus there must be $3n - 2g$ external lines, each corresponding to an external particles whose 4-momentum is specified. If there are no closed loops then there are $n - 1$ internal lines, and thus the number propagators is determined by $g = n - 1$. If there is one closed loop then there is one more propagator, and if there are ℓ closed loops then there are ℓ more propagators than when there are no closed loops. Thus the number of closed loops is $g - n + 1$. There are therefore $g - n + 1$ undetermined loop momenta.

There are g integrals over the 4-momenta carried by the propagators. There are n relations between the 4-momenta from conservation of 4-momentum at each vertex. Conservation of 4-momentum ($p_f = p_i$) is already extracted in the definition (11.21) with (11.24) of M_{fi}. Thus there are $n - 1$ unused relations between the 4-momenta. The g integrals over the $n - 1$ δ-functions leave $g - n + 1$ remaining integrals over 4-momenta. These integrals may be identified with the integrals over the $g - n + 1$ undetermined loop momenta.

The inclusion of m-photon vertices does not alter this argument. This may be seen by noting that the geometric structure of a diagram with an m-photon vertex is unchanged if one replaces the m-photon vertex by a closed particle loop with m sides. (Here a "closed particle loop" involves a closed solid line in the present case, but the following arguments apply more generally.) There is no integral over any undetermined momentum associated with the m-photon vertex, and so in this equivalence one can say that the

integral over the loop momentum has effectively been performed. Thus inclusion of an m-photon vertex effectively adds m vertices and m propagators without changing the number of undetermined loop momenta. The number of undetermined loop momenta remains equal to the number of closed loops.

The nonlinear response tensors may be calculated by performing a statistical average over closed particle loop diagrams. The average over the "triangle" and "box" diagrams, cf. Figure 16.1, give the quadratic and cubic nonlinear responses respectively. When diagrams containing m-photon vertices are included these already include the contributions from closed particle loops implicitly. Thus diagrams containing closed particle loops are to be omitted as they have effectively been replaced by analogous diagrams with an m-photon vertex in place of the closed particle loop. Note that to every diagram containing a closed particle loop there is an analogous diagram that differs only in the direction of the arrow around the loop. These diagrams are to be treated together as a pair of similar diagrams. The contributions of each such a pair of diagrams is included in the analogous diagram with an m-photon. Thus the diagrams containing an m-photon vertex actually replaces a pair of diagrams containing a closed particle loop with m sides.

The replacement of diagrams containing closed particle loops by analogous diagrams containing m-photon vertices applies in the absence of a medium provided that the response of the vacuum is described in terms of vacuum contributions to the response tensors. In the absence of a static fields, these vacuum contributions are nonzero only for even values of m. (Furry's theorem implies that the m-photon vertices for the vacuum are zero for m odd due to cancelation of the contributions from the two diagrams with oppositely directed arrows.)

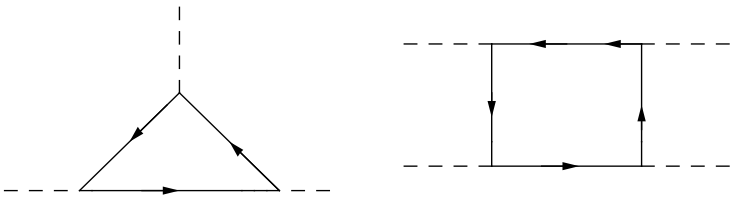

Figure 16.2 (a) The "triangle" diagram and (b) the "box" diagram are replaced by a 3-photon vertex and a 4-photon vertex respectively. There are in fact two "triangle" and two "box" diagrams that differ only in the direction of the arrow, and these pairs of diagrams are replaced by the 3-photon and 4-photon vertices respectively.

3. Static Fields

In the discussion so far it has been implicit that $A(x)$ describes a photon field. If $A(x)$ describes a static field, then it is to be expressed in terms of its Fourier transform. This provides the required exponential factor $\exp(-ikx)$ for the foregoing arguments to be extended to include a static field. The Fourier transform also contains an integral over $d^4k/(2\pi)^4$. The 4-momentum provided by the static field is undetermined, and the implication is that one integrates over this undetermined 4-momentum.

In practice a static field provides 3-momentum rather than 4-momentum. This is due to the fact that if $A(x)$ does not depend on t then its Fourier transform contains a factor $2\pi\,\delta(\omega)$ that ensures that the static field provides no energy to the system.

4. Sums over States of Polarization

The probability for specific processes involves the modulus squared of the Feynman amplitude. For each electron/positron line that joins an initial electron to a final electron the amplitude iM_{fi} includes a matrix product of the form $\bar{u}'\hat{O}u$, where $\hat{O}$ is an operator involving γ matrices and where u and u' denote $u_s(\boldsymbol{p})$ and $u_{s'}(\boldsymbol{p}')$ respectively. There are analogous products with u or $\bar{u}'$ replaced by v or $\bar{v}'$ for an electron/positron lines that join an initial positron to a final positron or for lines ending in initial or final pairs. In many situations one is not concerned with the states of polarization of the electrons or positrons, and it is then appropriate to average over the initial states of polarization and sum over the final states of polarization.

The sum over the two states of polarization is performed as follows. Consider a factor of the form $\bar{u}'\hat{O}u$ in iM_{fi}. Then $|M_{fi}|^2$ contains the product

$$(\bar{u}'\hat{O}u)\,(\bar{u}'\hat{O}u)^\dagger = \bar{u}'\hat{O}u\bar{u}\gamma^0\hat{O}^\dagger\gamma^0u' = \mathrm{Tr}(u'\bar{u}'\hat{O}u\bar{u}\gamma^0\hat{O}^\dagger\gamma^0). \tag{16.5}$$

In the last line in (16.5) the invariance of the trace of a product of matrices under cyclic permutations has been used. In this way the initial and final electron wavefunctions appear as the products $u\bar{u}$ and $u'\bar{u}'$. Similarly when initial or final positron states are involved $|M_{fi}|^2$ may be rewritten as a trace involving the wavefunctions only as the products $v\bar{v}$ or $v'\bar{v}'$. The sum may then be performed using (4.29) and (4.30), giving $\tilde{\not{p}} + m$ for electron states and $\tilde{\not{p}} - m$ for positron states, with $\tilde{p} = [\varepsilon, \boldsymbol{p}]$. (These sum are written again in (16.9) and (16.10) below.) The average over an initial state of polarization is half the sum over these states.

5. Additional Rules for Diagrams in Momentum Space

The foregoing discussion suggests the following rules for drawing diagrams in momentum space. These are in addition to the rules (i) to (vii) listed in Lecture 14.

(viii) For diagrams in momentum space all lines are labeled with the 4-momentum of the particles (rather than the vertices being labeled with the space-time points). At vertices 4-momentum is conserved.

(ix) The integral $(d^4q/(2\pi)^4)$ over any undetermined 4-momentum (q) in a closed loop or associated with an external field $A(q)$ is to be performed.

(x) Diagrams containing closed particle loops are to be omitted. (The contributions from such diagrams are included in corresponding diagrams with a closed particle loop with m sides replaced by an m-photon vertex, provided it is understood that the nonlinear responses of the vacuum are included along with the responses of a material medium.)

6. Rules for Constructing iM_{fi} and $w_{i\to f}$

The rules 1) to 7) given in Lecture 14 for writing down the form of S_{fi} for each diagram are modified for diagrams in momentum space. The following additional rules for writing down the form of iM_{fi}, defined by (11.21) with (11.24), and for the probability $w_{i\to f}$.

8) In constructing iM_{fi} each vertex is represented by a factor $ie\gamma_\mu$, each initial electron, positron and photon is represented by factors $u_s(\boldsymbol{p})$, $\bar{v}_s(\boldsymbol{p})$ and $e^\mu_M(\boldsymbol{k})$ respectively,

and each final electron, positron and photon is represented by factors $\bar{u}_s(\boldsymbol{p})$, $v_s(\boldsymbol{p})$ and $e^{*\mu}_M(\boldsymbol{k})$ respectively, with

$$a_M(\boldsymbol{k}) = \left(\frac{\mu_0 R_M(\boldsymbol{k})}{\omega_M(\boldsymbol{k}) V}\right)^{1/2}. \tag{16.6}$$

9) Internal electron and photon lines correspond to $-iG(p)$ and $iD_{\mu\nu}(k)$ respectively. An m-photon vertex corresponds to

$$\frac{1}{m}\,\alpha^{(m)\mu_0\mu_1\ldots\mu_m}(-k_0, k_1, \ldots, k_m)(2\pi)^4\,\delta^4(k_0 + k_1 + \cdots + k_m).$$

10) The transition probability per unit time is given by

$$w_{i\to f} = V\,(2\pi)^4\,\delta^4(p_f - p_i)|M_{fi}|^2 \left| \prod_i a_M(\boldsymbol{k})\,\frac{1}{\sqrt{2\varepsilon V}} \prod_f a_M(\boldsymbol{k})\,\frac{1}{\sqrt{2\varepsilon V}} \right|^2 D_f, \tag{16.7}$$

where the product is over all initial and final particles and photons, and where D_f is the density of final states factor

$$D_f = \prod_f \left(\frac{V d^3\boldsymbol{p}}{(2\pi)^3}\right) \prod_f \left(\frac{V d^3\boldsymbol{k}}{(2\pi)^3}\right). \tag{16.8}$$

11) For unpolarized electron or positrons one averages over the initial states of polarization and sums over the final states of polarization using the sums

$$\sum_{s=\pm} u_s(\boldsymbol{p})\bar{u}_s(\boldsymbol{p}) = \tilde{\not{p}} + m, \tag{16.9}$$

$$\sum_{s=\pm} v_s(\boldsymbol{p})\bar{v}_s(\boldsymbol{p}) = \tilde{\not{p}} - m, \tag{16.10}$$

with $\tilde{p} = [\varepsilon, \boldsymbol{p}]$. The average is half the sum.

Lecture 17

Cerenkov Emission

In vacuo all first order processes are forbidden. In a medium a with refractive index $\mu(\omega)$ greater than unity a particle with speed greater than the phase speed $v_\phi = c/\mu(\omega)$ does radiate. This is called *Cerenkov emission*, and it is a first order process. Cerenkov first observed such emission in the early 1930s as a blue glow from water near a radioactive source that produces *beta* rays, i.e., near a source of fast electrons.

1. The Probability for Cerenkov Emission

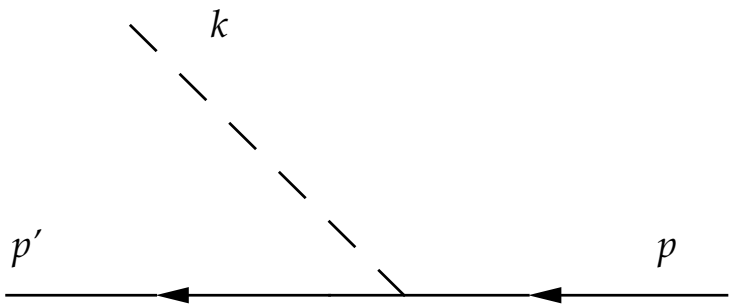

Figure 17.1 The Feynman diagram for Cerenkov emission.

Consider Cerenkov emission of a photon in a mode M with 4-momentum k by an electron in an initial state with 4-momentum p, so that the final 4-momemtum of the electron is $p' = p - k$. The Feynman diagram for this process is shown in Figure 17.1. The rules given in Lecture 15 allow one to write down the amplitude for this process. It is

$$iM_{fi} = ie\, e^*_{M\mu}(\boldsymbol{k})\, \bar{u}_{s'}(\boldsymbol{p'})\gamma^\mu u_s(\boldsymbol{p}) \tag{17.1}$$

where s and s' denote the initial and final spins of the electron.

The transition probability is given by (16.7), viz.

$$w_{i\to f} = V\,(2\pi)^4\,\delta^4(p_f - p_i)|M_{fi}|^2 \left| \prod_i a_M(\boldsymbol{k})\,\frac{1}{\sqrt{2\varepsilon V}} \prod_f a_M(\boldsymbol{k})\,\frac{1}{\sqrt{2\varepsilon V}} \right|^2 D_f. \tag{17.2}$$

The δ-function expressing conservation of 4-momentum is

$$(2\pi)^4\,\delta^4(p_f - p_i) \to (2\pi)^4\,\delta^4(p' + k - p).$$

There is no photon in the initial state and one in the final state, and there is an electron with energy ε in the initial states and an electron in the final state with energy ε' in the final state. Hence we have

$$\prod_i a_M(\boldsymbol{k})\,\frac{1}{\sqrt{2\varepsilon V}} \to \frac{1}{\sqrt{2\varepsilon V}} \quad \text{and} \quad \prod_f a_M(\boldsymbol{k})\,\frac{1}{\sqrt{2\varepsilon V}} \to a_M(\boldsymbol{k})\,\frac{1}{\sqrt{2\varepsilon' V}}.$$

The density of final states is

$$D_f \to \frac{V d^3\boldsymbol{p}'}{(2\pi)^3}\,\frac{V d^3\boldsymbol{k}}{(2\pi)^3}$$

Thus (17.2) reduces to

$$w_{i\to f} = V\,(2\pi)^4\,\delta^4(p'+k-p)\,|M_{fi}|^2\frac{|a_M(\boldsymbol{k})|^2}{2\varepsilon V 2\varepsilon' V}\frac{V d^3\boldsymbol{p}'}{(2\pi)^3}\,\frac{V d^3\boldsymbol{k}}{(2\pi)^3}. \tag{17.3}$$

It is convenient to write (17.3) in the form

$$w_{i\to f} = w_M(\boldsymbol{p},\boldsymbol{k})(2\pi)^3\,\delta^3(\boldsymbol{p}'+\boldsymbol{k}-\boldsymbol{p})\,\frac{d^3\boldsymbol{p}'}{(2\pi)^3}\,\frac{d^3\boldsymbol{k}}{(2\pi)^3}, \tag{17.4}$$

where

$$w_M(\boldsymbol{p},\boldsymbol{k}) = |M_{fi}|^2\frac{V|a_M(\boldsymbol{k})|^2}{2\varepsilon 2\varepsilon'}\,2\pi\,\delta(\varepsilon'+\omega_M(\boldsymbol{k})-\varepsilon) \tag{17.5}$$

is the *probability of Cerenkov emission.*

2. Average over polarization states

If one is not interested in the polarization of the electron, then on evaluating $|M_{fi}|^2$ one averages over the initial states of polarization and sums over the final states of polarization using (16.9). This gives

$$|M_{fi}|^2 = e^2\,e^*_{M\mu}(\boldsymbol{k})e_{M\nu}(\boldsymbol{k})\,\tfrac{1}{2}\mathrm{Tr}\,[(\not{p}'+m)\gamma^\mu(\not{p}+m)\gamma^\nu]\,, \tag{17.6}$$

with $p' = p-k$, and where (16.6) has been used.

The trace over the γ matrices may be evaluated using (9.8) to (9.11). First, using the definition (9.8) and the fact that the trace is zero for an odd number of γ matrices, one finds

$$\mathrm{Tr}\,[(\not{p}-\not{k}+m)\gamma^\mu(\not{p}+m)\gamma^\nu] = T^{\alpha\mu\beta\nu}(p-k)_\alpha p_\beta + m^2\,T^{\mu\nu}. \tag{17.7}$$

Using (9.11), and after some rearrangement, one finds

$$\tfrac{1}{4}\mathrm{Tr}\,[(\not{p}-\not{k}+m)\gamma^\mu(\not{p}+m)\gamma^\nu] = 2\left[(p-\tfrac{1}{2}k)^\mu(p-\tfrac{1}{2}k)^\nu + \frac{1}{4}\left(k^2 g^{\mu\nu}-k^\mu k^\nu\right)\right]. \tag{17.8}$$

The probability (17.5) reduces to

$$\begin{aligned} w_M(\boldsymbol{p},\boldsymbol{k}) = \frac{\mu_0 e^2 R_M(\boldsymbol{k})}{\varepsilon\varepsilon'\omega_M(\boldsymbol{k})}[|(\boldsymbol{p}-\tfrac{1}{2}\boldsymbol{k})\cdot\boldsymbol{e}_M(\boldsymbol{k})|^2 - \frac{1}{4}\omega_M^2(\boldsymbol{k}) \\ + \frac{1}{4}|\boldsymbol{k}\times\boldsymbol{e}_M(\boldsymbol{k})|^2]\,2\pi\,\delta(\varepsilon'+\omega_M(\boldsymbol{k})-\varepsilon), \end{aligned} \tag{17.9}$$

with $\varepsilon' = (m^2+|\boldsymbol{p}-\boldsymbol{k}|^2)^{1/2}$, and where we use the temporal gauge $e^\mu_M = [0,\boldsymbol{e}_M]$.

3. The Power Radiated in Transverse Waves

Consider Cerenkov emission of transverse waves in a isotropic medium. The power radiated by an electron can be evaluated by integrating

$$P_M = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\,\omega_M(\boldsymbol{k}) w_M(\boldsymbol{p},\boldsymbol{k}) \tag{17.10}$$

over $d^3\boldsymbol{k}$. The subscripts M are now omitted, and the wave properties, including the ratio of electric to total energy are given by (9.8).

It is convenient to write the integral in terms of polar coordinates in $\boldsymbol{k}$-space. Let θ be the angle between $\boldsymbol{k}$ and $\boldsymbol{v}$. One has

$$\int d^3\boldsymbol{k} = \int_0^{2\pi} d\phi \int_{-1}^{1} d\cos\theta \int d\omega\ \mu^2\omega^2\,\frac{d(\mu\omega)}{d\omega}, \tag{17.11}$$

where $|\boldsymbol{k}|$ has been rewritten as $\mu\omega$ with the ω of μ implicit. The integral over azimuthal angle $d\phi$ is trivial and gives 2π. The integral over $d\cos\theta$ may be performed over the δ-function in (17.9). One then needs to take account of the restrictions on the allowed range of $\cos\theta$ on the remaining ω-integral.

Using $\varepsilon' = (\varepsilon^2 - 2|\boldsymbol{p}|\,|\boldsymbol{k}|\cos\theta + |\boldsymbol{k}|^2)^{1/2}$, the integral over the δ-function reduces to

$$\int_{-1}^{1} d\cos\theta\,\delta(\varepsilon' + \omega - \varepsilon) = \frac{|\omega - \varepsilon|}{|\boldsymbol{p}|\,|\boldsymbol{k}|} \tag{17.12}$$

The δ-function in (17.9) determines the value of $\cos\theta$,

$$\cos\theta = \frac{1}{\mu v}\,\left[1 + (\mu^2 - 1)\frac{\omega}{2\varepsilon}\right], \tag{17.13}$$

and this value of $\cos\theta$ is to be substituted elsewhere in the integrand. Before this is done one sums over the two states of polarization of the photon to find

$$\sum\left(|(\boldsymbol{p} - \tfrac{1}{2}\boldsymbol{k})\cdot\boldsymbol{e}|^2 - \frac{1}{4}\omega^2 + \frac{1}{4}|\boldsymbol{k}\times\boldsymbol{e}|^2\right) = \left(|\boldsymbol{p}|^2\sin^2\theta - \frac{1}{2}\omega^2 + \frac{1}{2}\mu^2\omega^2\right), \tag{17.14}$$

and then substitutes for $\sin^2\theta = 1 - \cos^2\theta$, using (17.13).

One remaining step in the integration is to determine the limit on the ω-integration implied by $\cos\theta < 1$. This requires

$$\mu > \mu_- = \frac{\varepsilon v}{\omega} - \left[\left(\frac{\varepsilon v}{\omega}\right)^2 + 1 - \frac{2\varepsilon}{\omega}\right]^{1/2}. \tag{17.15}$$

The result obtained is

$$P = \frac{\mu_0 e^2 v}{4\pi}\int_{\mu>\mu_-} d\omega\omega\left[1 - \frac{1}{\mu^2 v^2}\,\left(1 + \frac{\mu^2\omega^2}{4\varepsilon^2 v^2}\right)\left(1 + (\mu^2 - 1)\frac{\omega}{2\varepsilon}\right)^2 + (\mu^2 - 2)\frac{\omega^2}{4\varepsilon^2 v^2}\right], \tag{17.16}$$

with $\cos\theta$ determined by (17.13) and with μ_- given by (17.15). In the non-quantum limit (17.16) reduces to a well known formula for Cerenkov emission:

$$P = \frac{\mu_0 e^2 v}{4\pi} \int_{\mu v > 1} d\omega\omega \left(1 - \frac{1}{\mu^2 v^2}\right), \tag{17.17}$$

where the ω dependence of the refractive index μ remains implicit. The fact that the integral weights high frequencies more strongly than low frequencies accounts for the blue color of Cerenkov emission.

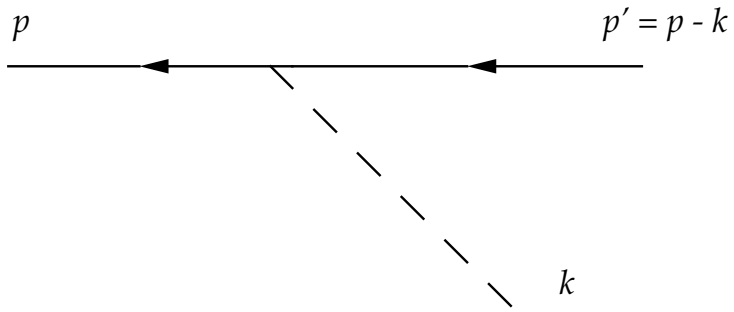

Figure 17.2 The Feynman diagram for Landau damping, which is the inverse of Cerenkov emission.

4. Induced Processes

The probability (17.5) applies to a transition $\boldsymbol{p} \to \boldsymbol{p} - \boldsymbol{k}$ with emission of a photon. It is straightforward to calculate the probability for the inverse transition $\boldsymbol{p} - \boldsymbol{k} \to \boldsymbol{p}$ with absorption of a photon. The Feynman diagram we consider is Figure 17.2. In place of (17.1) one has

$$iM_{fi} = ie\, e_{M\mu}(\boldsymbol{k})\, \bar{u}_{s''}(\boldsymbol{p}'')\gamma^\mu u_s(\boldsymbol{p}), \tag{17.18}$$

with $p'' = p + k$. To make this the inverse of the emission process described by (17.1) the replacements

$$p \to p' = p - k \quad \text{and} \quad p'' = p + k \to p$$

are made. One then finds that in place of (17.4) one has a transition probability

$$w_{i\to f} = w_M(\boldsymbol{p}, \boldsymbol{k})(2\pi)^3\, \delta^3(\boldsymbol{p}' + \boldsymbol{k} - \boldsymbol{p})\, \frac{V d^3\boldsymbol{p}'}{(2\pi)^3}, \tag{17.19}$$

with the probability $w_M(\boldsymbol{p}, \boldsymbol{k})$ again given by (17.5). The fact that the probabilities of emission and absorption between two states are equal is an example of the principle of *detailed balance*.

The probabilities considered so far, it has been implicit that there is only one electron and only one photon in the final state. Suppose that the occupation number of the initial state is $n(\boldsymbol{p})$ for the electrons and $N_M(\boldsymbol{k})$ for the photons. Then arguments given in Lecture 13, cf. (14.23) and (14.24) imply that the total probalitiy of emission (em) is

$$w_M^{\text{em}}(\boldsymbol{p}, \boldsymbol{k}) = 2w_M(\boldsymbol{p}, \boldsymbol{k})\, n(\boldsymbol{p})\, [1 - n(\boldsymbol{p} - \boldsymbol{k})]\, [1 + N_M(\boldsymbol{k})], \tag{17.20}$$

where the factor of 2 comes from the sum over the two polarization states of the electron. (Recall that we have averaged over these two states, and the sum is twice the average.) The total probability of absorption (abs) is

$$w_M^{\rm abs}(\boldsymbol{p},\boldsymbol{k}) = 2w_M(\boldsymbol{p},\boldsymbol{k})\,[1-n(\boldsymbol{p})]\,n(\boldsymbol{p}-\boldsymbol{k})\,N_M(\boldsymbol{k}). \tag{17.21}$$

As mentioned in Lecture 13, the unit term in the factor $[1+N_M(\boldsymbol{k})]$ in (17.20) is attributed to *spontaneous* emission and the term proportional to $N_M(\boldsymbol{k})$ is attributed to *stimulated* or *induced* emission.

In an emission event $N_M(\boldsymbol{k})$ increases by unity, and in an absorption event $N_M(\boldsymbol{k})$ decreases by unity. Hence, after integrating over all the initial electron states, the net rate of change of $N_M(\boldsymbol{k})$ is given by

$$\frac{dN_M(\boldsymbol{k})}{dt} = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\,[w_M^{\rm em}(\boldsymbol{p},\boldsymbol{k}) - w_M^{\rm abs}(\boldsymbol{p},\boldsymbol{k})]. \tag{17.22}$$

Then using (17.20) and (17.21) one obtains the kinetic equation for the photons:

$$\frac{dN_M(\boldsymbol{k})}{dt} = 2\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\,w_M(\boldsymbol{p},\boldsymbol{k})\,\{n(\boldsymbol{p})\,[1-n(\boldsymbol{p}-\boldsymbol{k})] + N_M(\boldsymbol{k})\,[n(\boldsymbol{p})-n(\boldsymbol{p}-\boldsymbol{k})]\}. \tag{17.23}$$

The terms independent of $N_M(\boldsymbol{k})$ on the right hand side of (17.23) describe spontaneous emission, and the terms proportional to $N_M(\boldsymbol{k})$ described the net effect of the induced terms of stimulated emission and true absorption. For $n(\boldsymbol{p}) > n(\boldsymbol{p}-\boldsymbol{k})$, which corresponds to an inverted energy population, then the net effect is one of amplification (negative absorption), which is also called *maser action*.

It is also possible to derive a kinetic equation for the electrons by similar arguments. Each emission event $(p \to p-k)$ decreases $n(\boldsymbol{p})$ by unity and each absorption event $(p-k \to p)$ increases $n(\boldsymbol{p})$ by unity. In this case one also needs to take into account transitions $p \to p+k$ and $p+k \to p$.

5. Crossed Processes

The processes related to Cerenkov emission by crossing have been identified in Lecture 13. Quite generally, the probability for two processes that differ by a particle or photon in the initial (or final) state being replaced by an anti-particle or photon in the final (or initial) state are related by $p \to -p$ or $k \to -k$ for the relevant particle or photon. Let us illustrate this by considering the decay of a pair into one photon. The Feynman diagram for this is illustrated in Figure 17.3.

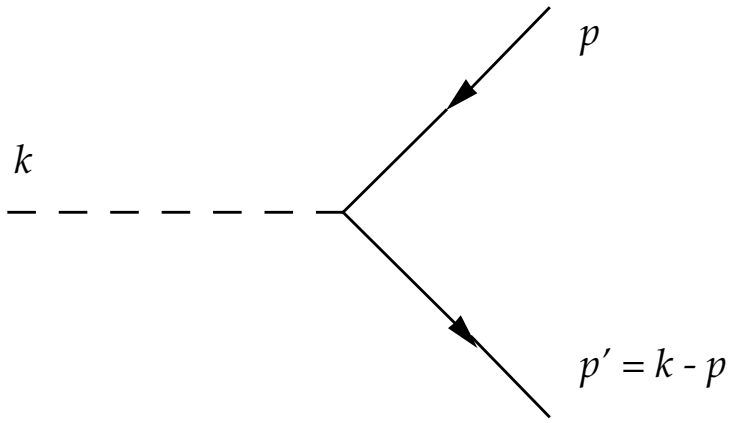

Figure 17.3 The Feynman diagram for decay of a pair into a single photon.

The matrix element, in place of (17.1) is

$$iM_{fi} = ie\, e^*_{M\mu}(\boldsymbol{k})\, \bar{v}_{s'}(\boldsymbol{p}')\gamma^\mu u_s(\boldsymbol{p}). \tag{17.24}$$

The argument of the δ-function in (17.2) becomes $(p_f - p_i) \rightarrow (p' + p - k)$. Then 3-momentum conservation implies $p' = [\varepsilon', -\boldsymbol{p} + \boldsymbol{k}]$, with $\varepsilon' = (m^2 + (\boldsymbol{p} - \boldsymbol{k})^2)^{1/2}$. The density of final states factor changes because there is now only a photon and no particle in the final state, but this is unimportant in the derivation of the probability.

The sum over the final states of polarization of an electron is replaced by an average over the initial states of polarization of a positron. By inspection of (16.9) and (16.10) this implies that in place of (17.6) we have

$$|M_{fi}|^2 = e^2\, e^*_{M\mu}(\boldsymbol{k}) e_{M\nu}(\boldsymbol{k})\, \frac{1}{4}\mathrm{Tr}\left[(\not{p}' - m)\gamma^\mu(\not{p} + m)\gamma^\nu\right], \tag{17.25}$$

with $p' = k - p$. The first term inside the trace is $(\not{p} - \not{k} + m)$ in the case of Cerenkov emission, and is $(-\not{p} + \not{k} - m)$ in the case of pair decay. The change in sign appears naturally, and is needed because the trace must be positive and it is positive for $p' = k - p$ only with the sign as in (17.25). If one wishes, then this change in sign could be attributed to a change in sign of the factor ε' in the denominator in (17.9). However, although this convention works for fermions, it does not work for bosons; there is no overall change in sign in replacing a boson in the initial (or final) state by an antiboson in the final (or initial) state.

It is then apparent by analogy with the derivation of (17.9) that the probability for pair decay (pd) into one photon is

$$\begin{aligned} w^{(\mathrm{pd})}_M(\boldsymbol{p}, \boldsymbol{p}') = \frac{\mu_0 e^2 R_M(\boldsymbol{k})}{2\varepsilon\varepsilon'\omega_M(\boldsymbol{k})}[-|(\boldsymbol{p} - \boldsymbol{p}')\cdot \boldsymbol{e}_M(\boldsymbol{k})|^2 + \omega^2_M(\boldsymbol{k}) \\ - |(\boldsymbol{p} + \boldsymbol{p}')\times \boldsymbol{e}_M(\boldsymbol{k})|^2]\, 2\pi\, \delta(\varepsilon' + \varepsilon - \omega_M(\boldsymbol{k})), \end{aligned} \tag{17.26}$$

with $\boldsymbol{k} = \boldsymbol{p} + \boldsymbol{p}'$.

The prescription for simply changing the sign of p' is correct, but it needs to be modified if a sum over polarization states has been performed.

6. Kinematic Restrictions

Finally, let us consider the kinematic conditions under Cerenkov emission and the crossed processes are allowed. The kinematic restrictions are implied by the δ-function in (17.4), together with the implicit relations $p^2 = m^2$ and $p^{'2} = m^2$ for the electron and the dispersion relation $\omega = \omega_M(\boldsymbol{k})$ for the waves. These require $p^{'2} = (p-k)^2 = p^2 - 2pk + k^2$, which implies $pk = \frac{1}{2}k^2$ or, equivalently,

$$\varepsilon(\omega - \boldsymbol{k}\cdot\boldsymbol{v}) = \tfrac{1}{2}\hbar(\omega^2 - |\boldsymbol{k}|^2), \tag{17.27}$$

where $\hbar$ has been included explicitly. The right hand side of (17.27) is a quantum correction. Writing $\boldsymbol{k}\cdot\boldsymbol{v} = |\boldsymbol{k}|\,|\boldsymbol{v}|\cos\theta$ in(17.27), the condition $\cos\theta \leq 1$ implies a threshold condition:

$$\varepsilon^2(\omega^2 - |\boldsymbol{k}|^2) - \varepsilon\omega(\omega^2 - |\boldsymbol{k}|^2) + |\boldsymbol{k}|^2 m^2 + \tfrac{1}{4}(\omega^2 - |\boldsymbol{k}|^2)^2 \geq 0. \tag{17.28}$$

Choosing the equality in (17.28) and solving for ε gives

$$\varepsilon = \tfrac{1}{2}\omega \pm \tfrac{1}{2}|\boldsymbol{k}|\left(\frac{\omega^2 - |\boldsymbol{k}|^2 - 4m^2}{\omega^2 - |\boldsymbol{k}|^2}\right)^{1/2}. \tag{17.29}$$

There are two classes of solution determined by the condition that the discriminant be real. One class is for $\omega^2 - |\boldsymbol{k}|^2 < 0$ with $\omega^2 - |\boldsymbol{k}|^2 - 4m^2$ then necessarily less than zero. Then only the + sign in (17.29) is acceptable, and one requires

$$\varepsilon > \tfrac{1}{2}\omega + \tfrac{1}{2}|\boldsymbol{k}|\left(\frac{4m^2 + |\boldsymbol{k}|^2 - \omega^2}{|\boldsymbol{k}|^2 - \omega^2}\right)^{1/2}. \tag{17.30}$$

which is equivalent to (17.15). This class applies to Cerenkov emission by an electron or a positron and to the corresponding absorption processes.

The second class of solution of (17.29) exist for $\omega^2 - |\boldsymbol{k}|^2 > 0$ with

$$\omega^2 > 4m^2 + |\boldsymbol{k}|^2. \tag{17.31}$$

This class corresponds to the crossed processes of decay of a photon into a pair or annihilation of a pair into a photon. Note that these process require $\omega > |\boldsymbol{k}|$, whereas Cerenkov emission requires $\omega < |\boldsymbol{k}|$.

Lecture 18

Gyromagnetic Processes

Classically an electron in a magnetic field emits due to its accelerated motion as it spirals around magnetic field lines. From a quantum viewpoint (Lecture 6) the motion perpendicular to the magnetic field lines is quantized as a simple harmonic oscillator, and the emission of radiation is attributed to transitions in which the electron jumps from a higher to a lower energy state. The generic name for this process is *gyromagnetic emission*. Gyromagnetic emission is analogous to Cerenkov emission in that it may be treated as a first order process in QED. Gyromagnetic emission is of relevance in several branches of physics, and quantum effects become of importance in very strong magnetic fields such as those of neutron stars. The crossed process of the decay of a single photon into a pair is of some interest in connection with neutron stars.

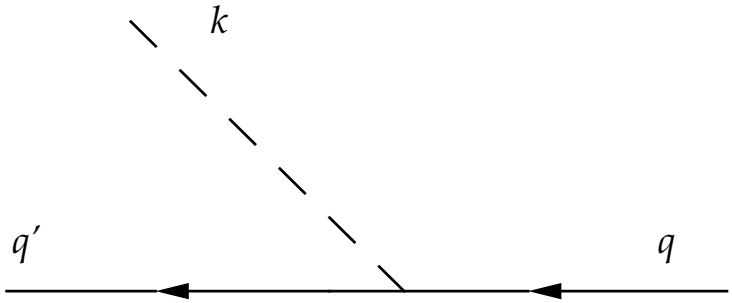

Figure 18.1 The Feynman diagram for gyromagnetic emission is the same as that for Cerenkov emission except for the labeling of the states.

1. The Probability of Gyromagnetic Emission

Wavefunctions for an electron in a magnetic field are written down in (6.20). With q denoting p_z, n, s) the solutions are are

$$\Psi^{\epsilon}_{q+}(t,\boldsymbol{x}) = \frac{\exp(-i\epsilon\varepsilon t + i\epsilon p_y y + i\epsilon p_z z)}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} (\epsilon\varepsilon_q + m)v_{n-1}(\xi) \\ 0 \\ \epsilon p_z v_{n-1}(\xi) \\ ip_n v_n(\xi) \end{pmatrix},$$
$$\Psi^{\epsilon}_{q-}(t,\boldsymbol{x}) = \frac{\exp(-i\epsilon\varepsilon t + i\epsilon p_y y + i\epsilon p_z z)}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} 0 \\ (\epsilon\varepsilon_q + m)v_n(\xi) \\ -ip_n v_{n-1}(\xi) \\ -\epsilon p_z v_n(\xi) \end{pmatrix}, \tag{18.1}$$

where $q+$ and $q-$ denote the quantum numbers with spin $s = +1$ and $s = -1$ respectively. (The spin operator is chosen implicitly for mathematical convenience; wavefunctions for physically meaningful spin operators are written down in Appendix B.) In (18.1) we have

$$p_n = (2neB)^{1/2}, \quad \varepsilon_q = (m^2 + p_z^2 + 2neB)^{1/2}, \quad \xi = (eB)^{1/2}(x + \epsilon p_y/eB),$$
$$v_\ell(\xi) = \frac{1}{(\sqrt{\pi}2^\ell l!)^{1/2}} H_\ell(\xi)\, e^{-\xi^2/2}, \qquad 2n = 2\ell + 1 + s. \tag{18.2}$$

The Feynman diagram for gyromagnetic emission is shown in Figure 18.1. In this case one cannot use the rules for diagrams in momentum space. However it is straightforward to use the S-matrix itself using the rules derived in Lecture 16. For the emission of a photon in the mode M, one finds

$$S_{fi} = ie \int d^4x\, a_M(\boldsymbol{k}) e^*_{M\mu}(\boldsymbol{k}) e^{ik_M x}\, \overline{\Psi}^+_{q'}(\boldsymbol{x}) e^{i\varepsilon_{q'}t} \gamma^\mu \Psi^+_q(\boldsymbol{x}) e^{-i\varepsilon_q t}. \tag{18.3}$$

It is convenient to introduce a vertex function

$$[\gamma^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})]^\mu := \int d^3\boldsymbol{x}\, \exp(-i\boldsymbol{k}\cdot\boldsymbol{x})\, \overline{\Psi}^{\epsilon'}_{q'}(\boldsymbol{x}) \gamma^\mu \Psi^\epsilon_q(\boldsymbol{x}). \tag{18.4}$$

Then (18.3) reduces to

$$S_{fi} = ie\, a_M(\boldsymbol{k}) e^*_{M\mu}(\boldsymbol{k}) [\gamma^{++}_{q'q}(\boldsymbol{k})]^\mu\, 2\pi\, \delta(\varepsilon_q - \varepsilon_{q'} - \omega_M(\boldsymbol{k})). \tag{18.5}$$

The vertex function (18.4) is dependent on the choice of gauge. One may isolate the gauge-dependent factors and define a vertex function $[\Gamma^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})]^\mu$ that is independent of the choice of gauge. It involves functions $J^n_\nu(x)$ with $x = k_\perp^2/2eB$. These functions are related to the generalized Laguerre polynomials $L^\nu_n(x)$ by

$$J^n_\nu(x) = \left(\frac{n!}{(n+\nu)!}\right)^{1/2} e^{-\frac{1}{2}x} x^{\frac{1}{2}\nu}\, L^\nu_n(x). \tag{18.6}$$

The detailed derivation of the vertex function and the properties of the functions (18.6) are summarized in Appendix E.

A subtle point is that the gauge-dependent quantum numbers (p_y in the present case) determine the position of the center of gyration of the electron, and this position changes in an emission event. Usually one is not interested in the position of the center of gyration and one averages over position to eliminate this dependence. After averaging over the position of the center of gyration, the probability for gyromagnetic emission must be independent of the choice of gauge, and it may be written in terms of the gauge-independent vertex function. The actual derivation of the probability is given in Appendix E.

The resulting expression for the *probability of gyromagnetic emission* is

$$[w^{\epsilon'\epsilon}_{q'q}]_M(\boldsymbol{k}) = \frac{\mu_0 e^2 R_M(\boldsymbol{k})}{\omega_M(\boldsymbol{k})} |e^*_{M\mu}(\boldsymbol{k}) [\Gamma^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})]^\mu|^2\, 2\pi\, \delta(\epsilon\varepsilon_q - \epsilon'\varepsilon_{q'} - \omega_M(\boldsymbol{k})). \tag{18.7}$$

Gyromagnetic emission by an electron is the specific case $\epsilon' = \epsilon = 1$, and the other choices of sign correspond to crossed processes.

2. Transition Rates

The probability (18.7) may be approximated in various ways to treat different gyromagnetic processes. The most important approximations are to the functions (18.5). There are approximations relevant to the classical limit (large n) and to the ultrarelativistic limit. Here we consider the opposite limit of small n when quantum effects need to be taken into account.

The limit of relevance for small $x \ll 1$ and small to moderate values of n, when the functions (18.5) may be approximated by zero or unity according to

$$J_\nu^n(0) = \delta_{\nu 0}. \tag{18.8}$$

Then the vertex function may be approximated by

$$\begin{aligned}\mathbf{\Gamma}^{++}_{q'q}(\boldsymbol{k}) =& \delta_{s's}\left[(-)^j\left(\frac{B}{2B_c}\right)^{1/2}\left(\frac{\ell!}{(\ell-j)!}\right)^{1/2}\frac{x^{\frac{1}{2}(j-1)}}{(j-1)!}\right](1,i,0)\\ &+\delta_{s'-s}\left[(-)^{(j-s)}\frac{B\cos\theta}{2B_c}\left(\frac{\ell!}{(\ell-j+s)!}\right)^{1/2}\frac{x^{\frac{1}{2}(j-s)}}{(j-s)!}(j,isj,-\tan\theta)\right],\end{aligned} \tag{18.9}$$

where θ is the angle between $\boldsymbol{k}$ and $\boldsymbol{B}$, and where

$$B_c = \frac{m^2}{e} = 4.4\times 10^9\,\mathrm{T} \tag{18.10}$$

is the so-called *critical* magnetic field.

The transitions may be classified according to the change in the orbital quantum number ℓ and the change in the spin quantum number s. Transitions in which s changes sign are said to involve a *spin flip*. For non-spin-flip transitions the change in the principal quantum number n is the same as the change in ℓ. The transition $\ell \to \ell - 1$ is the dominant transition; it leads to emission near the fundamental of the cyclotron frequency

$$\Omega = \frac{eB}{m}, \tag{18.11}$$

and corresponds to electric dipole transition. The transition $\ell \to \ell - 1$ leads to emission near the second harmonic of the cyclotron frequency and corresponds to electric quadrupole transition. The transitions $\ell \to \ell - j$ leads to emission near the jth harmonic of the cyclotron frequency and correspond to electric 2^j-multipole transition.

A spin-flip transition $s = 1 \to s = -1$ with $\ell' = \ell$ leads to emission near the fundamental of the cyclotron frequency and corresponds to magnetic dipole transition, and a transition $s = 1 \to s = -1$ with $\ell \to \ell - 1$ leads to emission near the second harmonic of the cyclotron frequency and corresponds to magnetic quadrupole transition. A spin-flip transition $s = -1 \to s = 1$ is possible only for $\ell \to \ell - j$ with $j \geq 2$.

The rate of transitions may be evaluated by summing over the density of states (including the sum over polarization states) of the emitted photon. For transverse waves in vacuo the resulting rate is

$$R_{q\to q'} = \tfrac{1}{2}\mu_0 e^2 \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\,\frac{1}{\omega}\left[|\boldsymbol{\Gamma}^{++}_{q'q}(\boldsymbol{k})|^2 - |\boldsymbol{\kappa}\cdot\boldsymbol{\Gamma}^{++}_{q'q}(\boldsymbol{k})|^2\right] 2\pi\,\delta(\varepsilon_q - \varepsilon_{q'} - \omega). \tag{18.12}$$

For the electric dipole transitions (18.12) reduces to

$$\begin{aligned} R_{\ell\to\ell-1} &= \tfrac{1}{2}\mu_0 e^2\, 2\pi \int \frac{d\omega\omega}{(2\pi)^3}\int_{-1}^{1} d\cos\theta\,\frac{B}{B_c}\,\frac{\ell!}{(\ell-1)!}\,(1+\cos^2\theta)\,2\pi\,\delta(\varepsilon_n - \varepsilon_{n-1} - \omega) \\ &= \ell\frac{\mu_0 e^2 B}{3\pi B_c}\,(\varepsilon_n - \varepsilon_{n-1}), \end{aligned} \tag{18.13}$$

with n related to ℓ by (18.2). The emission frequency is close to the cyclotron frequency $\varepsilon_n - \varepsilon_{n-1} \approx \Omega$, and then (18.13) reduces further to

$$R_{\ell\to\ell-1} = \ell\frac{\mu_0 e^4 B^2}{3\pi m^3}. \tag{18.14}$$

Electrons in a high energy state jump down to lower states, predominantly in a stepwise process in which n decreases by unity in each transition. In the absence of other effects, the electron eventually reaches its ground state. The lifetime for the decay to the ground state is of order the inverse of the transition rate for the slowest transition, which is the last transition from the first excited state to the ground state. Thus the lifetime is of order $3/B^2$ s, where B is in tesla. For the magnetic field near a pulsar on has $B \approx 0.1B_c = 4\ 10^8$ T. The lifetime for the decay is then extremely short $\approx 2\ 10^{-17}$ s. One expects all electrons to normally be in their ground state.

3. Gyromagnetic Masers

If the occupation number of a higher energy state exceeds that of a lower energy state then one expects stimulated emission to exceed absorption and allow maser action. Thus one would expect maser action is the occupation number increases with principal quantum number n over some range. Although this conclusion is basically correct, there is a subtle point that makes is less obvious than it appears. The existence of this type of maser depends intrinsically on a relativistic effect, specifically on the fact that the frequency difference $\omega = \varepsilon_n - \varepsilon_{n-1}$, for emission at the fundamental, depends on n. This effect has been called the *anharmonicity*.

Suppose, on the contrary, that relativistic effects are ignored so that one has $\varepsilon_n - \varepsilon_{n-1} = \Omega$ independent of n. Then transitions $n \to n-1$ for all n contribute to the (anticipated) maser action at frequency $\omega = \Omega$. The net rate of stimulated emission for the transitions $n \to n-1$ is proportional to the difference in occupation numbers $(n_n - n_{n-1})$. The total rate of transitions is found by summing over all possible transitions. The net rate is therefore proportional to

$$(n_1 - n_0) + (n_2 - n_2) + \cdots + (n_n - n_{n-1}) + \cdots = (n_\infty - n_0) = -n_0 < 0.$$

It follows that the net effect is always one of positive absorption and never one of negative absorption or maser action. When relativistic effects are included the frequency of each transition is different, and so for $n_n > n_{n-1}$ stimulated emission dominates, leading to maser action, at the frequency of this particular transition.

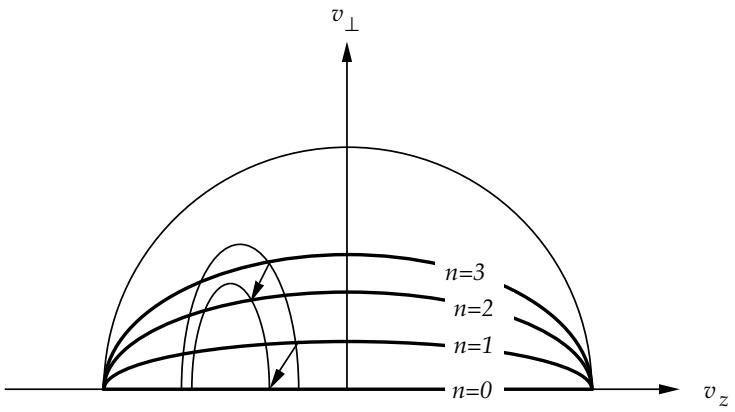

Figure 18.2 The allowed transitions due to gyromagnetic emission are illustrated schematically in a plane with velocity axes $v_\perp$ and v_z. The allowed states of the electron are labeled by there principal quantum numbers n. The initial states must lie on an ellipse and the final states must lie on another ellipse, both of which are determined by ω, k_z and the change in n. Two possible transitions are indicated by arrows.

The possibility of maser action at low values of n is illustrated in Figure 18.2. An artificial velocity space with component $v_\perp$ and v_z is introduced. On writing

$$p_z = \gamma m v_z, \quad p_n = \gamma m v_\perp, \quad \gamma = (1 - v_z^2 - v_\perp^2)^{-1/2}, \tag{18.15}$$

the allowed values of n and p_z correspond to ellipses, referred to a *n-ellipses*. Several n-ellipses are drawn in Figur 16.2 and labeled with the value of n.

The condition fro emission, expressed in the δ-function in the probability, is

$$(m^2 + p_z^2 + p_n^2)^{1/2} - (m^2 + {p'}_z^2 + p_{n'}^2)^{1/2} - \omega = 0, \tag{18.16}$$

with ${p'}_z = p_z - k_z$ and $n' = n-1$ here. On expressing p_z and p_n in terms of $v_\perp$ and v_z one obtains a curve on which the initial state must lie, and by expressing ${p'}_z$ and ${p'}_n$ in terms of $v_\perp$ and v_z one obtains a curve on which the final state must lie. These curves are also ellipses, and are referred to ar the *initial* and *final* ellipses respectively. Examples of such ellipses are shown schematically in Figure 18.2. The allowed initial state is where the initial ellipse intersect the n-ellipse, and the allowed final state is where the final ellipse intersect the $(n-1)$-ellipse. Two possible maser transitions are indicated in Figure 18.2.

Maser action occurs only if the higher energy of the two states determined by this graphical technique has a higher occupation number than the lower energy state. A pump mechanism is requires to cause such an inverted population.

4. One-Photon Pair Decay

The decay of one photon into a pair is possible in a *magnetized vacuum*, that is, in a vacuum in the presence of a magnetic field. This process may be treated using the probability (18.7) with $\epsilon = 1$ and $\epsilon' = -1$.

A magnetized vacuum is birefringent, that is, the transverse mode splits into two modes with different refractive indices and different polarizations. The wave properties may be approximated as follows. Let the wave modes be labeled + and −, and let the coordinate axes be oriented such that $\boldsymbol{B}$ is along the z-axis and $\boldsymbol{k}$ is in the x, z-plane, with

$$\boldsymbol{k} = |\boldsymbol{k}|\,\boldsymbol{\kappa}, \quad \boldsymbol{\kappa} = (\sin\theta, 0, \cos\theta). \tag{18.17}$$

The refractive indices for the modes of the *birefringent vacuum* are

$$\mu_\pm = 1 + \frac{e^2}{90\pi}\,\frac{B^2}{B_c^2}\,\sin^2\theta\,(11 \pm 3), \tag{18.18}$$

and the polarization vectors are

$$\boldsymbol{e}_+ = (0, 1, 0), \quad \boldsymbol{e}_- = (\cos\theta, 0, -\sin\theta). \tag{18.19}$$

It is convenient to refer to these as the *perpendicular* and *parallel* modes and to relabel them $+, - \to \perp, \parallel$, respectively.

The actual values of the refractive indices are not important in the present application. However, the decay rates into a pair for photons in the two modes are different due to their different polarizations.

The decay rate corresponds to the absorption coefficient $\gamma_M(\boldsymbol{k})$, which, for any mode M, is given by

$$\gamma_M(\boldsymbol{k}) = \sum_{\substack{n,n'\\ s,s'}} \frac{eB}{2\pi} \int \frac{dp_z}{2\pi}\, w^{\epsilon'\epsilon}_{Mq'q}(\boldsymbol{k}), \tag{18.20}$$

with $\epsilon = -\epsilon' = 1$ here. On inserting the probability (18.7) and the polarization vectors (18.19) in (18.20), the absorption coefficient reduces to

$$\gamma_{\perp,\parallel}(\boldsymbol{k}) = \sum_{\substack{n,n'\\ s,s'}} \frac{eB}{2\pi} \int \frac{dp_z}{2\pi}\, |[\Gamma]^{+-}_{qq'}(\boldsymbol{k})]^{2,3}|^2\, 2\pi\,\delta(\varepsilon_q + \varepsilon_{q'} - \omega). \tag{18.21}$$

Further reduction of (18.21) requires explicit expressions for the function $\boldsymbol{\Gamma}$, and suitable approximations to them. One simplifying approximation is that the pair is highly relativistic. In this limit one has

$$\gamma_{\perp,\parallel}(\boldsymbol{k}) \approx \frac{\mu_0 e^3 B}{4\pi\omega} \int_0^\infty du \int_0^\infty dw \frac{\omega^3 m \cosh u}{8e^2 B^2 \cosh^4 w} F_{\perp,\parallel}(u,w) \tag{18.22}$$

with

$$\begin{aligned} \begin{pmatrix} F_\perp(u,w) \\ F_\parallel(u,w) \end{pmatrix} = & \frac{32}{3\pi^2} \frac{m^4}{\omega^4} \left\{ \begin{pmatrix} \cosh^2 w \\ \sinh^2 w \end{pmatrix} \cosh^4 w \cosh^4 u \, K_{2/3}^2(z) \right. \\ & \left. + \tfrac{1}{2}[2\cosh^2 w \cosh^2 u - \sinh^2 u] \cosh^4 w \cosh^2 u \, K_{1/3}^2(z) \right\}, \end{aligned} \tag{18.23}$$

and with

$$z := \frac{2}{3\chi} \cosh^2 u \cosh^2 w, \qquad \chi := \frac{\omega e B}{2m^3} = \frac{\omega B}{m B_c}. \tag{18.24}$$

The detailed derivation of the result (18.22) with (18.23) is outlined in Appendix E.

Lecture 19

Mott Scattering and Bremsstrahlung

When an electron passes near an atom it is deflected by the Coulomb field of the atom. Classically this scattering is usually called Rutherford scattering; the quantum form is Mott scattering. During the encounter the electron is accelerated and so emits radiation. This is called bremsstrahlung, which is literally "braking radiation".

1. Scattering of an Electron by a Coulomb Field

Let us assume that the field of the nucleus may be described as an external field $A^\mu_{\rm ext}(x)$. Further let us assume that it is the Coulomb field (in the Coulomb gauge)

$$\boldsymbol{A}_{\rm ext}(x) = 0, \quad \phi_{\rm ext}(x) = -\frac{Ze}{4\pi\varepsilon_0 r}, \tag{19.1}$$

where Ze is the nuclear charge and where the nucleus is assumed to be at the origin (r=0). An exact treatment would require solving Dirac's equation, for example, as in Lecture 7. (In practice one needs to solve by separating in parbolic rather than spherical polar coordinates.) However, for many purposes it is adequate to use perturbation theory, which is what is done here.

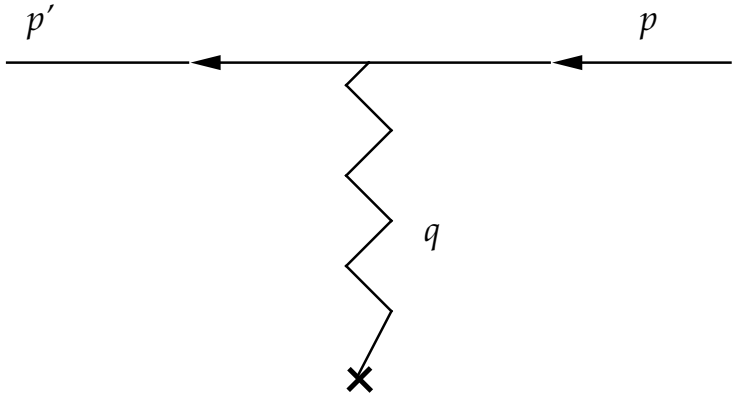

Figure 19.1 The Feynman diagram for scattering of an electron by an external field.

The Feynman diagram for scattering of an electron by an external field is Figure 19.1. The S-matrix element follows from the rules given in Lecture 16:

$$S_{fi} = \sum_{qq'} ie \int d^4x\, \overline{\Psi}^+_{q'}(\boldsymbol{x})\, e^{i\varepsilon_{q'}t} \not{A}_{\rm ext}(x)\, \Psi^+_q(\boldsymbol{x})\, e^{-i\varepsilon_q t}. \tag{19.2}$$

On writing $d^4x = dtd^3\boldsymbol{x}$ the t-integral is trivial in the case of a static field. It gives $2\pi\delta(\varepsilon'_q - \varepsilon_q)$, implying that the energy of the electron is unchanged. On inserting the field (19.1), (19.2) reduces to

$$S_{fi} = \sum_{qq'} -\frac{iZe^2}{4\pi\varepsilon_0} 2\pi\delta(\varepsilon_{q'} - \varepsilon_q) \int d^3\boldsymbol{x}\, \frac{1}{r}\, \overline{\Psi}^+_{q'}(\boldsymbol{x})\, \gamma^0\, \Psi^+_q(\boldsymbol{x}). \tag{19.3}$$

Now we assume plane wave solutions, as in (16.1) and (16.2). The integral in (19.3) is then of the form

$$\int d^3\boldsymbol{x}\,\frac{1}{r}\,\exp(-i\boldsymbol{q}\cdot\boldsymbol{x}) = \frac{4\pi}{|\boldsymbol{q}|^2}, \tag{19.4}$$

with $\boldsymbol{q} = \boldsymbol{p}' - \boldsymbol{p}$. Thus (19.3) gives

$$S_{fi} = -\frac{iZe^2}{4\pi\varepsilon_0}\,\frac{2\pi\delta(\varepsilon'-\varepsilon)}{V\sqrt{2\varepsilon' 2\varepsilon}}\,\frac{4\pi}{|\boldsymbol{p}'-\boldsymbol{p}|^2}\,\bar{u}_{s'}(\boldsymbol{p}')\,\gamma^0\,u_s(\boldsymbol{p}). \tag{19.5}$$

The transition probability per unit $w_{i\to f}$ time is given by $|S_{fi}|^2/T$, which with $|2\pi\,\delta(\varepsilon'-\varepsilon)|^2 = T\,2\pi\,\delta(\varepsilon'-\varepsilon)$ reduces to

$$w_{i\to f} = \frac{Z^2e^4}{(4\pi\varepsilon_0)^2}\,\frac{2\pi\,\delta(\varepsilon'-\varepsilon)}{2\varepsilon' 2\varepsilon V^2}\left(\frac{4\pi}{|\boldsymbol{p}'-\boldsymbol{p}|^2}\right)^2 |\bar{u}_{s'}(\boldsymbol{p}')\,\gamma^0\,u_s(\boldsymbol{p})|^2. \tag{19.6}$$

2. The Mott Cross Section

The kinematics require $\varepsilon' = \varepsilon$, implying that the energy, momentum $|\boldsymbol{p}|$ and the speed v of the electron are unchanged. Let θ be the *scattering angle*, that is the angle between $\boldsymbol{p}'$ and $\boldsymbol{p}$. Then one has

$$|\boldsymbol{p}'-\boldsymbol{p}|^2 = 2|\boldsymbol{p}|^2(1-\cos\theta) = 4|\boldsymbol{p}|^2\sin^2\tfrac{1}{2}\theta. \tag{19.7}$$

The average over the initial states and the sum over the final of polarization (denoted by an overline) give

$$\begin{aligned}\overline{|\bar{u}_{s'}(\boldsymbol{p}')\,\gamma^0\,u_s(\boldsymbol{p})|^2} &= \tfrac{1}{2}\mathrm{Tr}[(\not{p}'+m)\gamma^0(\not{p}+m)\gamma^0]\\ &= 2(p'^0p^0 + \boldsymbol{p}'\cdot\boldsymbol{p} + m^2)\\ &= 4\varepsilon^2(1-v^2\sin^2\tfrac{1}{2}\theta),\end{aligned} \tag{19.8}$$

where (9.11) has been used to evaluate the trace. Then (19.6) reduces to

$$w_{i\to f} = \frac{Z^2e^4}{(4\pi\varepsilon_0)^2}\,\frac{2\pi\,\delta(\varepsilon'-\varepsilon)}{V^2}\left(\frac{4\pi}{4|\boldsymbol{p}|^2\sin^2\frac{1}{2}\theta}\right)^2(1-v^2\sin^2\tfrac{1}{2}\theta). \tag{19.9}$$

The differential cross section for this process is found by multiplying by the density of final states $d^3\boldsymbol{p}'/(2\pi)^3$, carrying out the integral over $|\boldsymbol{p}'|$, and dividing by the flux of incoming particles v/V. This gives

$$\frac{d\sigma}{d^2\boldsymbol{\Omega}} = \frac{V^2}{v}\int_0^\infty \frac{d|\boldsymbol{p}'|\,|\boldsymbol{p}'|^2}{(2\pi)^3}\,\frac{Z^2e^4}{(4\pi\varepsilon_0)^2}\,\frac{2\pi\,\delta(\varepsilon'-\varepsilon)}{V^2}\left(\frac{4\pi}{4|\boldsymbol{p}|^2\sin^2\frac{1}{2}\theta}\right)^2(1-v^2\sin^2\tfrac{1}{2}\theta). \tag{19.10}$$

The integral reduces to

$$\int \frac{d|\boldsymbol{p}'|\,|\boldsymbol{p}'|^2}{(2\pi)^3}\,2\pi\,\delta(\varepsilon'-\varepsilon)=\frac{|\boldsymbol{p}|^2}{(2\pi)^2 v}, \tag{19.11}$$

where $|\boldsymbol{p}|=\varepsilon v$ and $d\varepsilon/d|\boldsymbol{p}|=v$ is used. Thus one obtains the *Mott cross section*

$$\frac{d\sigma}{d^2\boldsymbol{\Omega}}=\frac{Z^2e^4(1-v^2\sin^2\frac{1}{2}\theta)}{4(4\pi\varepsilon_0)^2|\boldsymbol{p}|^2v^2\sin^4\frac{1}{2}\theta} \tag{19.12}$$

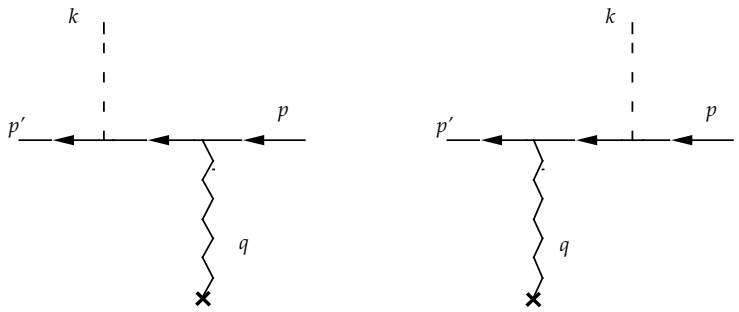

Figure 19.2 The Feynman diagrams for bremsstrahlung due to scattering of an electron by an external field.

3. The Bethe-Heitler Formula

The Feynman diagrams for bremsstrahlung are illustrated in Figure 19.2. Proceeding as in the treatment of Mott scattering the S-matrix element is

$$\begin{aligned} S_{fi}=&\frac{iZe^3}{4\pi\varepsilon_0|\boldsymbol{p}'+\boldsymbol{k}-\boldsymbol{p}|^2}\,\frac{2\pi\,\delta(\varepsilon'+\omega-\varepsilon)}{\sqrt{2\varepsilon'2\varepsilon}V}\,e^*_\mu(\boldsymbol{k})\bar{u}_{s'}(\boldsymbol{p}')\\ &\left[\frac{\gamma^\mu(\not{p}'+\not{k}+m)\gamma^0}{(p'+k)^2-m^2}+\frac{\gamma^0(\not{p}-\not{k}+m)\gamma^\mu}{(p-k)^2-m^2}\right]u_s(\boldsymbol{p}). \end{aligned} \tag{19.13}$$

The cross section for bremsstrahlung may be defined analogously to that for Mott scattering, except that the density of final states now includes a factor $d^3\boldsymbol{k}/(2\pi)^3$. The density of final states is written

$$D_f=\frac{Vd^3\boldsymbol{p}'}{(2\pi)^3}\,\frac{Vd^3\boldsymbol{k}}{(2\pi)^3}=V^2\left(\frac{\omega^3|\boldsymbol{p}'|^2d\varepsilon'}{(2\pi)^6v'^2}\right)\frac{d\omega}{\omega}d^2\boldsymbol{\Omega}_{\rm ph}d^2\boldsymbol{\Omega}_{\rm el}, \tag{19.14}$$

where the two solid angles are for the photon and the final electron, respectively. The integral over $d\varepsilon'$ is performed over the δ-function. The remainder of the calculation is quite lengthy, and will not be given in detail here. However, there are several points in the general derivation that warrant mention.

On evaluating $w_{i\to f}$ as in (10.9), one sums over the final state of polarization of the photon using

$$\sum_{\text{poln}} e^*_\mu e_\nu = -g_{\mu\nu}. \tag{19.15}$$

The justification of (19.15) presupposes that the tensor, $T^{\mu\nu}(k)$ say, with which $e^*_\mu e_\nu$ is contracted satisfies the gauge-invariance relations $k_\mu T^{\mu\nu}(k) = 0 = k_\nu T^{\mu\nu}(k)$. Then (10.9) in the temporal gauge implies

$$\sum_{\text{poln}} e^*_\mu e_\nu T^{\mu\nu}(k) = T^{11}(k) + T^{22}(k) + T^{33}(k) - \kappa_i \kappa_j T^{ij}(k),$$

where a mixed tensor notation is used in the last term. The gauge-invariance condition implies $\kappa_i \kappa_j T^{ij}(k) = T^{00}(k)$, which completes the justification of (19.15).

The average over the initial states and sum over final states of polarization of the electron requires that one evaluate the trace (after setting $p^2 = m^2,\ k^2 = 0$)

$$\begin{aligned} w_{i\to f} \propto \tfrac{1}{2}\mathrm{Tr}\Bigg\{ (\not{p}' + m)\left[\gamma^\mu \frac{(\not{p}' + \not{k} + m)}{2p'k}\gamma^0 - \gamma^0 \frac{(\not{p} - \not{k} + m)}{2pk}\gamma^\mu\right] \\ \times (\not{p} + m)\left[\gamma^0 \frac{(\not{p}' + \not{k} + m)}{2p'k}\gamma^\nu - \gamma^\nu \frac{(\not{p} - \not{k} + m)}{2pk}\gamma^0\right]\Bigg\}. \end{aligned} \tag{19.16}$$

The resulting calculation is quite lengthy, and it is worthwhile looking for tricks to simplify the calculation. One trick is to note that (19.13) may be simplified by using the relations

$$(\not{p} - m)u = 0, \quad \bar{u}'(\not{p}' - m) = 0. \tag{19.17}$$

Then the trace in (19.16) simplifies to

$$\begin{aligned} w_{i\to f} \propto \tfrac{1}{2}\mathrm{Tr}\Bigg\{ (\not{p}' + m)\left[\frac{(2p'^\mu + \gamma^\mu \not{k})}{2p'k}\gamma^0 - \gamma^0 \frac{(2p^\mu - \not{k}\gamma^\mu)}{2pk}\right] \\ \times (\not{p} + m)\left[\gamma^0 \frac{(2p'^\nu + \not{k}\gamma^\nu)}{2p'k} - \frac{(2p^\nu - \gamma^\nu \not{k})}{2pk}\gamma^0\right]\Bigg\}. \end{aligned} \tag{19.18}$$

The resulting evaluation is still quite cumbersome in the general case.

The final expression is the *Bethe Heitler formula*

$$\begin{aligned} d\sigma = \frac{d\omega}{\omega} d^2\boldsymbol{\Omega}_{\text{ph}} d^2\boldsymbol{\Omega}_{\text{el}} \frac{Z^2 e^6}{(4\pi\varepsilon_0)^3} \frac{|\boldsymbol{p}'|}{|\boldsymbol{p}|\,|\boldsymbol{q}|^4} \\ \Bigg\{ \frac{|\boldsymbol{p}'|^2 \sin^2\theta'}{(\varepsilon' - |\boldsymbol{p}'|\cos\theta')^2}(4\varepsilon^2 - |\boldsymbol{q}|^2) + \frac{|\boldsymbol{p}|^2 \sin^2\theta}{(\varepsilon - |\boldsymbol{p}|\cos\theta)^2}(4\varepsilon'^2 - |\boldsymbol{q}|^2) \\ + \frac{2\omega^2(|\boldsymbol{p}|^2\sin^2\theta + |\boldsymbol{p}'|^2\sin^2\theta')}{(\varepsilon - |\boldsymbol{p}|\cos\theta)^2(\varepsilon' - |\boldsymbol{p}'|\cos\theta')^2} \\ - \frac{2|\boldsymbol{p}|\,|\boldsymbol{p}'|\sin\theta\sin\theta'\cos\phi}{(\varepsilon - |\boldsymbol{p}|\cos\theta)^2(\varepsilon' - |\boldsymbol{p}'|\cos\theta')^2}(4\varepsilon\varepsilon' - |\boldsymbol{q}|^2 + 2\omega^2)\Bigg\}, \end{aligned} \tag{19.20}$$

where θ and θ' are the angles between $\boldsymbol{p}$, $\boldsymbol{k}$ and $\boldsymbol{p}'$, $\boldsymbol{k}$, respectively, where the angle ϕ is defined by writing

$$\boldsymbol{k} \times \boldsymbol{p} \cdot \boldsymbol{k} \times \boldsymbol{p}' = |\boldsymbol{k}|^2 |\boldsymbol{p}| \, |\boldsymbol{p}'| \, \sin\theta \sin\theta' \cos\phi, \tag{19.21}$$

and with

$$\boldsymbol{q} := \boldsymbol{p}' + \boldsymbol{k} - \boldsymbol{p}. \tag{19.22}$$

Remarkably, it is possible to carry out the integrals over solid angles to find the cross section per unit frequency range. The resulting expression is lengthy. In the nonrelativistic limit a calculation in the Born approximation gives

$$\frac{d\sigma}{d\omega} = \frac{16 Z^2 e^6}{3(4\pi\varepsilon_0)^2 m^2 v^2 \omega} \ln\frac{v + v'}{v - v'}, \tag{19.23}$$

which applies for $Ze^2/4\pi\varepsilon_0 v \ll 1$, $Ze^2/4\pi\varepsilon_0 v' \ll 1$. The analogous result in the ultrarelativistic limit is

$$\frac{d\sigma}{d\omega} = \frac{4 Z^2 e^6}{(4\pi\varepsilon_0)^2 m^2 \omega} \frac{\varepsilon'}{\varepsilon} \left(\frac{\varepsilon}{\varepsilon'} + \frac{\varepsilon'}{\varepsilon} - \frac{2}{3} \right) \left(\ln\frac{2\varepsilon\varepsilon'}{m\omega} - \frac{1}{2} \right) \tag{19.24}$$

4. Bremsstrahlung Emission of Soft Photons

One case that may be evaluated relatively simply is for "soft" photons. The specific assumption made is that the terms involving k in the parentheses in (19.18) may be neglected in comparison with the terms involving p or p'. With the approximation the cross section for bremsstrahlung becomes proportional to that for Mott scattering. One finds

$$\frac{(d\sigma/d^2\Omega)_{\text{softbrems}}}{(d\sigma/d^2\Omega)_{\text{Mott}}} = \frac{e^2}{4\pi\varepsilon_0} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \frac{1}{2\omega} \left(\frac{p'}{p'k} - \frac{p}{pk} \right)^2. \tag{19.25}$$

The integral over $d^3\boldsymbol{k}/(2\pi)^3$ in (19.25) may be evaluated by writing it as

$$\frac{d^3\boldsymbol{k}}{(2\pi)^3} = d\omega \omega^2 d^2\boldsymbol{\Omega}/(2\pi)^3$$

where $d^2\boldsymbol{\Omega}$ denotes the integral over solid angles about the direction $\boldsymbol{\kappa}$ of $\boldsymbol{k}$. One has

$$pk = \varepsilon\omega(1 - \boldsymbol{\kappa} \cdot \boldsymbol{v}), \quad p'k = \varepsilon\omega(1 - \boldsymbol{\kappa} \cdot \boldsymbol{v}'). \tag{19.26}$$

Two integrals need to be evaluated. One is

$$\int d^2\boldsymbol{\Omega} \, \frac{1}{(1 - \boldsymbol{\kappa} \cdot \boldsymbol{v})^2} = 4\pi \left(\frac{\varepsilon}{m} \right)^2, \tag{19.27}$$

with a second integral of the same form with primed quantities. The other integral may be evaluated using a trick that is widely used in QED. The integral is

$$\int d^2\boldsymbol{\Omega}\,\frac{1-\boldsymbol{v}\cdot\boldsymbol{v}'}{(1-\boldsymbol{\kappa}\cdot\boldsymbol{v})(1-\boldsymbol{\kappa}\cdot\boldsymbol{v}')} = \int_0^1 dx \int d^2\boldsymbol{\Omega}\,\frac{1-\boldsymbol{v}\cdot\boldsymbol{v}'}{[1-x\boldsymbol{\kappa}\cdot\boldsymbol{v}-(1-x)\boldsymbol{\kappa}\cdot\boldsymbol{v}']^2} \tag{19.28}$$

$$= \frac{2\pi(1-\boldsymbol{v}\cdot\boldsymbol{v}')}{X}\,\ln\left|\frac{X+1-\boldsymbol{v}\cdot\boldsymbol{v}'}{X-1+\boldsymbol{v}\cdot\boldsymbol{v}'}\right|,$$

where the introduction of the x-integral is referred to as a *Feynman parametrization*, and with

$$X = [(\boldsymbol{v}-\boldsymbol{v}')^2 - (\boldsymbol{v}\times\boldsymbol{v}')^2]^{1/2}. \tag{19.29}$$

The final result is

$$\frac{(d\sigma/d^2\Omega)_{\text{softbrems}}}{(d\sigma/d^2\Omega)_{\text{Mott}}} = \frac{e^2}{4\pi\varepsilon_0}\,\frac{2v^2\sin^2\frac{1}{2}\theta}{3\pi}\,\ln\frac{\omega_{\max}}{\omega_{\min}}, \tag{19.30}$$

where the integral is cut off at $\omega_{\max}$ and $\omega_{\min}$. The evaluation of these cutoff frequencies involves a detailed investigation; the resulting expression for the logarithmic term is usually expressed in terms of the *Gaunt factor*.

The divergence of the (19.30) at large ω ($\omega_{\max}\to\infty$) is nonphysical, as energy conservation requires $\omega_{\max} < \varepsilon$. The divergence at small ω is unavoidable; it is one of a class of "infra-red" divergences in QED. This divergence is of no significance in practice; for example, emission is impossible below the plasma frequency (because there are no transverse waves below the plasma frequency) and the plasma frequency must be non-zero whenever there are free electrons to emit bremsstrahlung.

Pair Creation by a Photon in a Coulomb Field

Pair formation due to a photon passing near a nucleus is related to bremsstrahlung by a crossing symmetry. The Feynman diagrams in Figure 19.2 for bremsstrahlung are modified to those in Figure 19.3. The following results apply in the Born approximation, cf. Berestetskii *et al.* (1971, pp. 327-330).

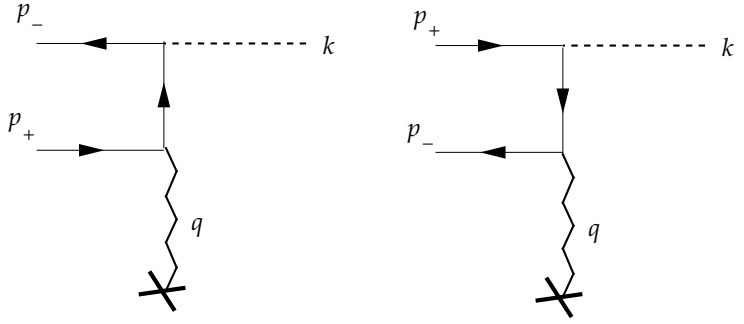

Figure 19.3 The Feynman diagrams for pair creation due to a photon in an external field.

The pair production processes may be described by schematically by $Z + \gamma \to Z + e^- + e^+$, where bremsstrahlung is described by $Z + e^- \to Z + e^- + \gamma$. Hence the probability for bremsstrahlung needs to be modified by the replacements $p \to -p_+$, to replace the initial electron by the final positron, and $k \to -k$ to move the photon from the initial to the final state. It is also convenient to relabel the final electrons: $p' \to p_-$. The angles in (19.20) are then replaced according to $\theta \to \pi - \theta_+$, $\theta' \to \theta_-$, $\phi \to \phi - \pi$. The cross section in a form analogous to (19.20) is

$$\begin{aligned}
\mathrm{d}\sigma = -\mathrm{d}\varepsilon_+ \mathrm{d}\cos\theta_+ \mathrm{d}\cos\theta_- \mathrm{d}\phi \, \frac{Z^2 e^6}{(4\pi\varepsilon_0)^3 2\pi} \, \frac{|\boldsymbol{p}_+||\boldsymbol{p}_-|}{\omega^3 |\boldsymbol{q}|^4} \Bigg\{ & \frac{|\boldsymbol{p}_+|^2 \sin^2\theta_+}{(\varepsilon_+ - |\boldsymbol{p}_+|\cos\theta_+)^2} \, (4\varepsilon_-^2 - |\boldsymbol{q}|^2) \\
& + \frac{|\boldsymbol{p}_-|^2 \sin^2\theta_-}{(\varepsilon_- - |\boldsymbol{p}_-|\cos\theta_-)^2} \, (4{\varepsilon_+}^2 - |\boldsymbol{q}|^2) - \frac{2\omega^2(|\boldsymbol{p}_-|^2 \sin^2\theta_- + |\boldsymbol{p}_+|^2 \sin^2\theta_+)}{(\varepsilon_- - |\boldsymbol{p}_-|\cos\theta_-)(\varepsilon_+ - |\boldsymbol{p}_+|\cos\theta_+)} \\
& - \frac{2|\boldsymbol{p}_-|\,|\boldsymbol{p}_+|\sin\theta_-\sin\theta_+\cos\phi}{(\varepsilon_- - |\boldsymbol{p}_-|\cos\theta_-)(\varepsilon_+ - |\boldsymbol{p}_+|\cos\theta_+)} \, (2\varepsilon_-^2 + 2\varepsilon_+^2 - |\boldsymbol{q}|^2) \Bigg\},
\end{aligned} \tag{19.31}$$

with $|\boldsymbol{q}|^2 = (\mathbf{p}_+ + \mathbf{p}_- - \mathbf{k})^2$, $\varepsilon_- + \varepsilon_+ = \omega$.

After integrating over the angles, the counterpart of (19.22) is

$$\begin{aligned}
\mathrm{d}\sigma_{\mathrm{pair}} = \frac{Z^2 e^6}{(4\pi\varepsilon_0)^3 m^2} \, \frac{|\boldsymbol{p}_+||\boldsymbol{p}_-|}{\omega^3} \, \mathrm{d}\varepsilon_+ \Bigg\{ & -\frac{4}{3} - 2\varepsilon_-\varepsilon_+ \, \frac{|\boldsymbol{p}_-|^2 + |\boldsymbol{p}_+|^2}{|\boldsymbol{p}_-|^2 |\boldsymbol{p}_+|^2} \\
& + m^2 \left(\ell_- \, \frac{\varepsilon_+}{|\boldsymbol{p}_-|^3} + \ell_+ \, \frac{\varepsilon_-}{|\boldsymbol{p}_+|^3} - \frac{\ell_-\ell_+}{|\boldsymbol{p}_-||\boldsymbol{p}_+|} \right) \\
& + L \Bigg[-\frac{8\varepsilon_-\varepsilon_+}{3|\boldsymbol{p}_-||\boldsymbol{p}_+|} + \frac{\omega^2}{|\boldsymbol{p}_-|^3|\boldsymbol{p}_+|^3} \, (\varepsilon_-^2\varepsilon_+^2 + |\boldsymbol{p}_-|^2|\boldsymbol{p}_+|^2 - m^2\varepsilon_-\varepsilon_+) \\
& - \frac{m^2\omega}{2|\boldsymbol{p}_-||\boldsymbol{p}_+|} \, (\ell_- \, \frac{\varepsilon_-\varepsilon_+ - |\boldsymbol{p}_-|^2}{|\boldsymbol{p}_-|^3} + \ell_+ \, \frac{\varepsilon_-\varepsilon_+ - |\boldsymbol{p}_+|^2}{|\boldsymbol{p}_+|^3} \Bigg] \Bigg\},
\end{aligned} \tag{19.32}$$

$$L = \ln \frac{\varepsilon_-\varepsilon_+ + |\boldsymbol{p}_-||\boldsymbol{p}_+| + m^2}{\varepsilon_-\varepsilon_+ - |\boldsymbol{p}_-||\boldsymbol{p}_+| + m^2}, \qquad \ell_\pm = \ln \frac{\varepsilon_\pm + |\boldsymbol{p}_\pm|}{\varepsilon_\pm - |\boldsymbol{p}_\pm|}.$$

In the ultrarelativistic case, $\varepsilon_\pm \gg m$, the cross section (19.32) may be approximated by

$$\mathrm{d}\sigma_{\mathrm{pair}} = 4Z^2\alpha_f r_0^2 \, \frac{d\varepsilon_+|}{\omega^3} \, (\varepsilon_+^2 + \varepsilon_-^2 + \tfrac{2}{3}\varepsilon_+\varepsilon_-) \left(\ln \frac{2\varepsilon_+\varepsilon_-}{m\omega} - \frac{1}{2} \right). \tag{19.33}$$

Integration over ε_+ from m to $\omega \gg m$ gives the total cross section for pair production by a photon of given energy

$$\sigma_{\mathrm{pair}} = \frac{28Z^2\alpha_f r_0^2}{9} \left(\ln \frac{2\omega}{m} - \frac{109}{42} \right). \tag{19.34}$$

For a nonrelativistic pair (19.33) and (19.34) are replaced by

$$\mathrm{d}\sigma_{\mathrm{pair}} = \frac{2Z^2\alpha_f r_0^2}{3m^3}\,(\omega - 2m)\,[(\varepsilon_+ - m)(\varepsilon_- - m)]^{1/2}\,d\varepsilon_+, \tag{19.35}$$

$$\sigma_{\mathrm{pair}} = \frac{\pi Z^2\alpha_f r_0^2}{12}\,\left(\frac{\omega - 2m}{m}\right)^3. \tag{19.36}$$

Lecture 1

Nonrelativistic Quantum Mechanics

In this Lecture nonrelativistic quantum mechanics is reviewed briefly following an approach originally due to Dirac.

1. Lagrangian and Hamiltonian Mechanics

The most convenient form of classical mechanics as a basis for generalization to quantum mechanics is Hamiltonian mechanics. Starting from a Lagrangian system, Hamilton's equations are introduced by making a Legendre transformation, that is, by changing the independent variables.

In Lagrangian mechanics, a holonomic system with s degrees of freedom is described by a set of s generalized coordinates $q_1, \ldots, q_s$ and the associated generalized velocities $\dot{q}_1, \ldots, \dot{q}_s$, where the dots denote the time derivatives. The Lagrangian $L(q, \dot{q}, t)$ is a function of these variables, where q and $\dot{q}$ denote, respectively, $q_1, \ldots, q_s$ and $\dot{q}_1, \ldots, \dot{q}_s$ collectively. The equations of motion in Lagrangian form are

$$\left[\frac{d}{dt}\frac{\partial}{\partial \dot{q}_i} - \frac{\partial}{\partial q_i}\right] L(q, \dot{q}, t) = 0, \qquad \text{for} \quad i = 1, \ldots, s. \tag{1.1}$$

The generalized momenta $p_1, \ldots, p_s$ are defined by

$$p_i := \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}, t), \qquad i = 1, \ldots, s. \tag{1.2}$$

The change of independent variables is from the q and $\dot{q}$ to the q and p. If one writes

$$H = \sum_{i=1}^{s} p_i \dot{q}_i - L, \tag{1.3}$$

then in the differential

$$dH = \sum_{i=1}^{s} \Big\{ p_i \, d\dot{q}_i + \dot{q}_i \, dp_i \Big\} - \sum_{i=1}^{s} \left\{ \frac{\partial L}{\partial q_i} \, dq_i + \frac{\partial L}{\partial \dot{q}_i} \, d\dot{q}_i \right\}$$

the terms involving $d\dot{q}_i$ cancel, so that the q and p are the new independent variables, as required. The equations of motion (1.1) and the relation implied by (1.3) between the $\dot{q}$ and p become

$$\dot{q}_i = \frac{\partial}{\partial p_i} H(q, p, t), \qquad \dot{p}_i = -\frac{\partial}{\partial q_i} H(q, p, t), \qquad \text{for} \quad i = 1, \ldots, s. \tag{1.4}$$

These are Hamilton's equations.

2. Canonical Transformations

Within the framework of classical mechanics, a major advantage of the Hamiltonian formalism over the Lagrangian formalism is the ability to make *canonical transformations*. Such a tranformation is from variables q, p and Hamiltonian $H(q,p)$ to another set of variables Q, P and Hamiltonian $K(Q,P)$, where Q and P are functions of q and p. The new variables and the new Hamiltonian satisfy Hamilton's equations.

A special case of a canonical transformation is a *point transformation*, which corresponds to the new coordinates Q being functions only of the old coordinates q, and not of the old momenta p. The new Hamiltonian is then equal to the old Hamiltonian. A simple example of a point transformation is the transformation from cartesian to spherical polar or cylindrical polar coordinates. Note that, although the new coordinates do not depend on the old momenta, the new momenta do depend on the old coordinates as well as on the old momenta in general; for example, the radial momentum in spherical polar coordinates is $p_r = (xp_x + yp_y + zp_z)/r$.

3. Poisson Brackets

The detailed form of Hamilton's equations depends on the choice of coordinates (and on the choice of gauge in the presence of an electromagnetic field). It is desirable to develop much of the theory in a way that is independent of the specific choices made. One such development is based on the use of Poisson brackets. Let f and g be any functions of q, p, t. Their Poisson bracket is defined by

$$[f,g]_{\mathrm{P}} := \sum_{i=1}^{s} \left[\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right]. \tag{1.5}$$

The value of $[f,g]_{\mathrm{P}}$ is unchanged under a canonical transformation.

Setting f and g equal to one of the q or p leads to the *fundamental Poisson brackets*

$$[q_i,q_j]_{\mathrm{P}} = 0, \quad [p_i,p_j]_{\mathrm{P}} = 0, \quad [q_i,p_j]_{\mathrm{P}} = \delta_{ij}, \qquad i,j = 1,\ldots,s. \tag{1.6}$$

These relations may be reinterpreted as a definition of a canonical set (q, p). Also the time evolution of any function is determined by its explicit time dependence and its Poisson bracket with the Hamiltonian:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{i=1}^{s} \left[\dot{q}_i \frac{\partial f}{\partial q_i} + \dot{p}_i \frac{\partial f}{\partial p_i} \right] = \frac{\partial f}{\partial t} + [f,H]_{\mathrm{P}}, \tag{1.7}$$

where (1.4) and (1.5) are used.

4. The State Function and Operators

There are several steps in the generalization from classical mechanics to quantum mechanics. Firstly a conceptual change needs to be made. This involves describing the particular system of interest in terms of a state function.

To allow for interference and diffraction effects, one needs to introduce a representation of a state of a system that can be decomposed into two or more parts that can interfere with each other. Dirac introduced such a state function in an abstract way in terms of his bra, $\langle\ |$, and ket, $|\ \rangle$, notation. What is known about the system is included in the state function, and it is convenient to include this information as arguments, so that $|\,1\,\rangle$ and $|\,2\,\rangle$ describe states labelled 1 and 2, respectively. The ket $|\ \rangle$ is regarded as a vector in a complex vector space, specifically a Hilbert space $\mathcal{H}$, and the corresponding bra $\langle\ | = \{|\ \rangle\}^\dagger$ as the corresponding vector in the adjoint Hilbert space $\mathcal{H}^\dagger$ (in which all quantities are complex conjugated). The ket for a mixed state (such as in a two slit diffraction experiment) is the sum of the kets for the states of which it is composed.

The inner product is a bracket formed from a bra and a ket. The inner product between states 2 and 1 is written

$$\langle 2|1\rangle = \langle 1|2\rangle^*$$

where * denotes complex conjugation.

One extracts information on the system by performing a set of operations (measurements or observations) on the system. Each set of operations is described by an operator in the Hilbert space. Operators are denoted by hats, with $\hat{Q}$ the operator corresponding to measuring a quantity Q. An operation can change the state of a system. The number

$$Q_{21} = \langle 2|\,\hat{Q}|1\rangle \tag{1.8}$$

is referred to as the *matrix element* between states 1 and 2 for the operator $\hat{Q}$. If the operation of $\hat{Q}$ on a system leads to a definite value Q, then the system is in an eigenstate of $\hat{Q}$ with eigenvalue Q. Observable values are real numbers, and hence the eigenvalues of an observable quantity must be real. This requires that the operators corresponding to observable quantities be self adjoint: $\hat{Q}^\dagger = \hat{Q}$.

Two other formal requirements are that the eigenkets of any self adjoint operator span the Hilbert space, and that eigenkets corresponding to different eigenvalues be orthogonal. For continuous eigenvalues this orthogonality condition led Dirac to introduce the δ-function. Thus if q and q' are two different eingenvalues of and operator $\hat{q}$ with a continuous spectrum of eigenvalues, then one requires

$$\langle\, q'|\,\hat{q}\,|\,q\,\rangle = q\delta(q - q'). \tag{1.9}$$

These ideas are summarized in the following axioms:

> *The state of a system is described by a vector (a ket $|\ \rangle$) in a Hilbert space $\mathcal{H}$, and an observable Q is described by a self adjoint operator $\hat{Q}$ in $\mathcal{H}$. The eigenkets of any self adjoint operator span the Hilbert space, and eigenkets corresponding to different eigenvalues are orthogonal.*

The link between this formalism and the classical description is expressed in terms of the relation between the quantum mechanical commutator between two operators and the corresponding Poisson bracket. The commutator is

$$[\hat{f}, \hat{g}] = \hat{f}\,\hat{g} - \hat{g}\,\hat{f}. \tag{1.10}$$

The relation is expressed in the following axiom (retaining $\hbar$ explicitly for the present):

> *The commutator* $[\hat{f}, \hat{g}]$ *of any two operators is equal to* $i\hbar$ *times the corresponding Poisson bracket* $[f, g]_{\mathrm{P}}$.

The basic Poisson brackets (1.6) then imply the *fundamental commutation relations*

$$[\hat{q}_i, \hat{q}_j] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0, \quad [\hat{q}_i, \hat{p}_j] = i\hbar\,\delta_{ij}, \qquad i, j = 1, \ldots, s. \tag{1.11}$$

5. Representations

In detailed calculations it is usually convenient to use a specific representation of the Hilbert space. The most widely used representation is the *coordinate* or *Schrödinger* representation. This representation is implicit in the familiar Schrödinger equation of nonrelativistic quantum mechanics, and in the standard forms for the Klein Gordon and Dirac equations. Other representations include the momentum representation, which is related to the coordinate representation by a Fourier transform and is sometimes useful when plane wave solutions are involved, and matrix representations, which are useful when the eigenvalues are discrete.

The coordinate representation is introduced by using the eigenkets $|\,q\,\rangle$ $(= |q_1, \ldots, q_s\rangle)$ of the coordinate operators $\hat{q}_i$ as basis vectors in the Hilbert space. One has

$$\hat{q}_i\,|q_1, \ldots, q_s\rangle = q_i\,|q_1, \ldots, q_s\rangle, \qquad i = 1, \ldots, s. \tag{1.12}$$

The matrix elements of the coordinate operators are diagonal:

$$\langle q'_1, \ldots, q'_s|\,\hat{q}_i\,|q_1, \ldots, q_s\rangle = q_i\,\delta(q_1 - q'_1)\,\ldots\,\delta(q_s - q'_s), \tag{1.13}$$

and this implies that the operator $\hat{q}_i$ is represented by multiplication by q_i. An arbitrary ket $|\,\rangle$ is represented by the so called wavefunction

$$\Psi(q) = \langle\,q\,|\,\rangle, \tag{1.14}$$

where the argument q represents $q_1, \ldots, q_s$ collectively. Finally, to satisfy the commutation relations (1.11), the operator $\hat{p}_i$ must be represented by $-i\hbar\partial/\partial q_i$.

6. Pictures for the Time Evolution

A final major step in generalizing classical mechanics to quantum mechanics involves including the time evolution. The classical evolution of any variable f may be described in the form (1.7). To avoid possible confusion, let us assume that f does not depend explicitly on time ($\partial f/\partial t = 0$). Then the classical evolution of f is due only to the dependence of f on the changing coordinates and momenta of the particles in the system. We are concerned with the quantum mechanical counterpart of this evolution.

In quantum mechanics f is replaced by the matrix elements of the corresponding operator $\hat{f}$, and one is free to choose to include the time evolution in the operator or in the kets and bras, or in a mixture of the two. All three choices are convenient for different purposes. Any specific choice is referred to as a *picture* for the time evolution. The *Schrödinger picture* has all the time evolution in the kets and bras, and the *Heisenberg picture* has all the evolution in the operators. The *interaction picture* has part of the evolution in the bras and kets and another part in the operators.

Note that the some authors refer to *representations* of the time evolution. Here *representation* is used only to refer to representations of the kets and operators in the Hilbert space, and *picture* is used exclusively to refer to the time evolution.

In the Schrödinger picture the operators do not depend on time and the evolution of the kets implied by (1.7), with the Poisson bracket replaced by the commutator as postulated, is

$$i\hbar \frac{d}{dt} \left| t \right\rangle_{\mathrm{S}} = \hat{H} \left| t \right\rangle_{\mathrm{S}}, \tag{1.15}$$

where the subscript S denotes the Schrödinger picture. In the Heisenberg picture the kets do not depend on time and the operators evolve according to

$$i\hbar \frac{d}{dt} \hat{f}_{\mathrm{H}} = [\hat{f}_{\mathrm{H}}, \hat{H}], \tag{1.16}$$

where the subscript H denotes the Heisenberg picture.

7. The Schrödinger Equation

The Schrödinger equation applies in the Schrödinger picture and the coordinate representation. Writing $H(q, p, t)$ for the Hamiltonian, including only one coordinate explicitly, the coordinate representation of (1.15) becomes

$$\left[i\hbar \frac{\partial}{\partial t} - \hat{H}(q, -i\hbar\partial/\partial q, t) \right] \Psi(q, t) = 0, \tag{1.17}$$

which is a general form of the *Schrödinger equation*.

The specific systems of interest here are a free particle and a particle in an electromagnetic field. The Hamiltonian for a nonrelativistic free particle is $H = \boldsymbol{p}^2/2m$, in which case (1.17) becomes

$$\left[i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \nabla^2 \right] \Psi(\boldsymbol{x}, t) = 0, \tag{1.18}$$

with

$$\nabla^2 = \text{div grad} = (\partial/\partial x)^2 + (\partial/\partial y)^2 + (\partial/\partial z)^2.$$

The Hamiltonian for a nonrelativistic particle with charge q in an electromagnetic field is $H = (\boldsymbol{p} - q\boldsymbol{A})^2/2m + q\phi$. Then (1.18) is replaced by

$$\left[i\hbar\,\frac{\partial}{\partial t} + \frac{\hbar^2}{2m}\,\nabla^2 - \frac{i\hbar q}{m}\,\boldsymbol{A}\cdot\text{grad} - \frac{i\hbar q}{2m}\,\text{div}\boldsymbol{A} + \frac{q^2}{2m}\,\boldsymbol{A}^2 + q\phi\right]\Psi(\boldsymbol{x},t) = 0. \tag{1.19}$$

The prescription for generalizing relativistic Hamiltonians to include the electromagnetic field is analogous to the nonrelativistic case: $\boldsymbol{p}$ is replaced by $\boldsymbol{p} - q\boldsymbol{A}$ and a term $q\phi$ is added to H.

8. Inclusion of Spin

The spin of the electron needs to be included artificially in the generalization of the Schrödinger equation to the Schrödinger-Pauli equation. One introduces a complex two dimensional spin space, thereby formally doubling the dimensionality of the Hilbert space, and introduces a spin vector $\frac{1}{2}\hbar\boldsymbol{\sigma}$ in this space. The Hamiltonian H for an electron is replaced by

$$H = H' - \mu_e\,\boldsymbol{\sigma}\cdot\boldsymbol{B}, \tag{1.20}$$

where

$$\mu_e = \frac{e\hbar}{2m} \tag{1.21}$$

is the magnetic moment of the electron.

The components of $\boldsymbol{\sigma}$ are the Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.22}$$

They satisfy the relations

$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0, \qquad \sigma_x\sigma_y = i\sigma_z, \qquad (\sigma_x)^2 = 1, \tag{1.23}$$

together with other identities obtained by cyclic permutations of x,y,z. In tensor notation (1.23) becomes

$$\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k. \tag{1.24}$$

Lecture 20

Compton Scattering and Related Processes

The lowest order processes that are allowed in vacuo are of second order, that is they correspond to Feynman diagrams with two vertices. These processes include Compton scattering, Møller scattering and bremsstrahlung, as well as processes related to these by crossing symmetries. In Compton and Møller scatterings and related processes, there are two particles (regarding a photon as a particle for present purposes) in both the initial and final states. We start by discussing the kinematics of such processes in general.

1. Invariant Kinematics

Consider a process involving two particles in the initial state and two particles in the final state. Let these have 4-momenta p_1, p_2, p_3, p_4. Conservation of 4-momentum requires

$$p_1 + p_2 = p_3 + p_4, \tag{20.1}$$

so that there are only three independent 4-momenta. Let the masses of the four particles be m_1, m_2, m_3, m_4. Then we have

$$p_1^2 = m_1^2, \quad p_2^2 = m_2^2, \quad p_3^2 = m_3^2, \quad p_4^2 = m_4^2. \tag{20.2}$$

It is convenient to define three invariants:

$$s := (p_1 + p_2)^2, \quad t := (p_1 - p_3)^2, \quad u := (p_1 - p_4)^2. \tag{20.3}$$

One refers to the process in which the particles of masses m_1, m_2 are in the initial (or final) state as the *s-channel*, the process in which the particles of masses m_1, m_3 are in the initial (or final) state as the *t-channel*, the process in which the particles of masses m_1, m_4 are in the initial (or final) state as the *u-channel*. Using (20.1) one finds that these three invariants are related by

$$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2. \tag{20.4}$$

There are limits on the physically allowed ranges of these variables. For example one has $p_i p_j \geq m_i m_j$, as may be seen by considering the rest frame of one (the jth say) of the particles, where one has $p_i p_j = \varepsilon_i m_j \geq m_i m_j$; then because $p_i p_j$ is an invariant, this inequality must apply in all frames. In the s-channel the inequalities $p_i p_j \geq m_i m_j$ imply

$$\begin{aligned} s &\geq (m_1 + m_2)^2, (m_3 + m_4)^2, \\ t &\leq (m_1 - m_3)^2, (m_2 - m_4)^2, \\ u &\leq (m_1 - m_4)^2, (m_2 - m_3)^2. \end{aligned} \tag{20.5}$$

Similar limits may be derived for the t- and u-channels.

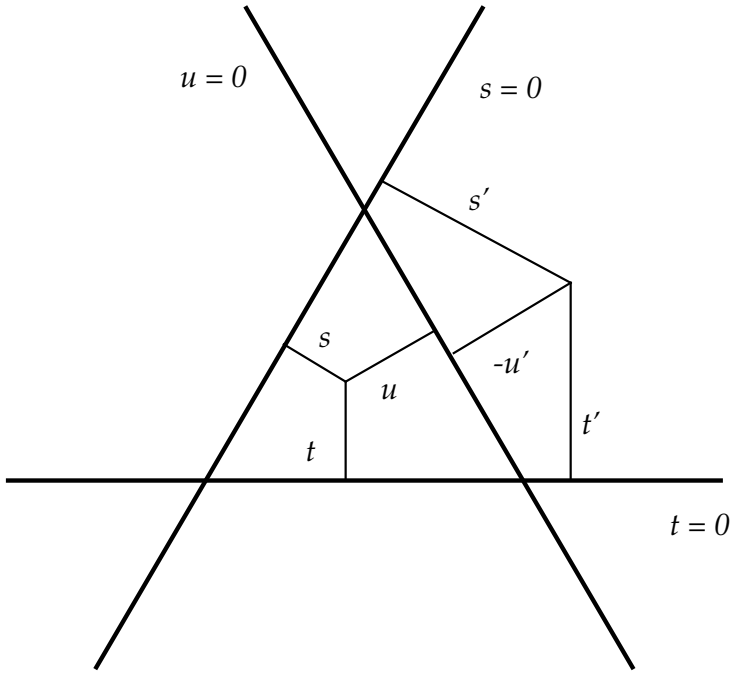

Figure 20.1 In the Mandelstam plane a given value of s, t and u is represented by a point which is a vertical distance s, t and u from the three sides of an equilateral triangle of height $h = m_1^2 + m_2^2 + m_3^2 + m_4^2$. For a point outside the triangle, as illustrated by the point labeled s', t' and u', one or more of the distances is negative.

The limits (20.5) on s may be derived in terms of a so-called *Gram* determinant

$$\begin{vmatrix} p_1^2 & p_1p_2 \\ p_2p_1 & p_2^2 \end{vmatrix} \geq 0, \quad \begin{vmatrix} p_3^2 & p_3p_4 \\ p_4p_3 & p_4^2 \end{vmatrix} \geq 0. \tag{20.6}$$

More general inequalities may be derived by considering other Gram determinants formed from the 4-vectors. Specifically, one has

$$\begin{vmatrix} p_1^2 & p_1p_2 & p_1p_3 \\ p_2p_1 & p_2^2 & p_2p_3 \\ p_3p_1 & p_3p_2 & p_3^2 \end{vmatrix} \geq 0, \tag{20.6}$$

and similarly for any other three 4-momenta. Such inequalities lead to a further restriction of the allowed range of the variables s, t, u. This range may be expressed algebraically by

$$stu \geq a_{12}s + a_{13}t + a_{14}u, \tag{20.7}$$

$$a_{12} = \frac{(m_1^2m_2^2 - m_3^2m_4^2)(m_1^2 + m_2^2 - m_3^2 - m_4^2)}{(m_1^2 + m_2^2 + m_3^2 + m_4^2)}, \tag{20.8}$$

and similarly for the other two coefficients. A graphical presentation of the allowed range in the *Mandelstam plane* is shown in Figure 20.1.

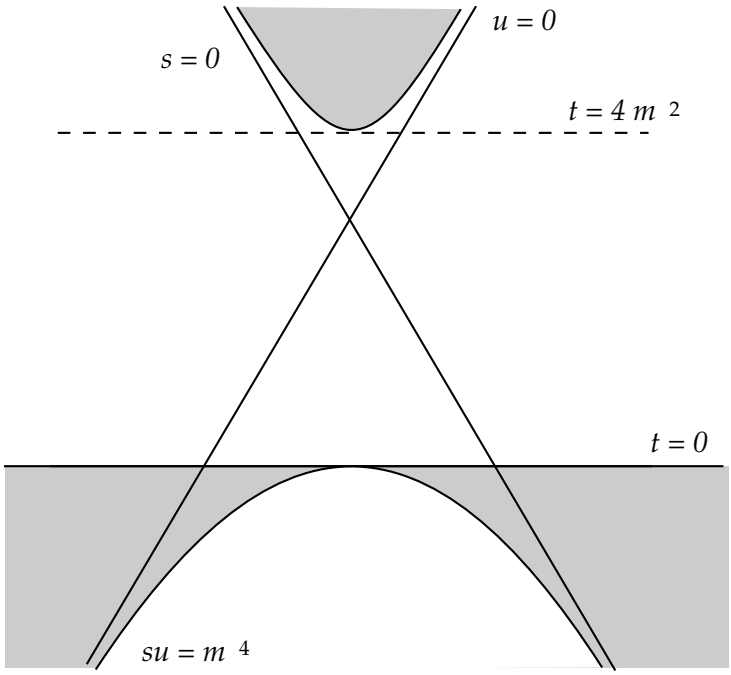

Figure 20.2 The Mandelstam plane for Compton scattering is illustrated. The physical regions are determined by the curve (20.7), which reduces to $su \geq m^4$ and (20.5), which reduces to $s \geq m^2$, $t \leq 0$. This is the shaded region on the lower left of the figure. The other physical regions are the shaded regions on the lower right, which is for the u-channel, and on the top of the figure, which is for the t-channel.

2. Compton Scattering

The Feynman diagrams for Compton scattering in vacuo are shown in Figure 20.3. The corresponding term in the scattering amplitude is

$$iM_{fi} = ie^2\, e^*_\nu(\boldsymbol{k}')e_\mu(\boldsymbol{k})\,\bar{u}(\boldsymbol{p}')\,\{\gamma^\mu G(p-k')\gamma^\nu + \gamma^\nu G(p+k)\gamma^\mu\}\,u(\boldsymbol{p}), \tag{20.9}$$

where subscripts M to denote the wave mode are irrelevant when considering waves in vacuo. The transition probability may be written in the form, cf. (17.2),

$$w_{i\to f} = \frac{1}{V}\,w(\boldsymbol{p},\boldsymbol{k}',\boldsymbol{k})\,(2\pi)^3\,\delta^3(\boldsymbol{p}'+\boldsymbol{k}'-\boldsymbol{p}-\boldsymbol{k})\,\frac{d^3\boldsymbol{p}'}{(2\pi)^3}\,\frac{d^3\boldsymbol{k}'}{(2\pi)^3}. \tag{20.10}$$

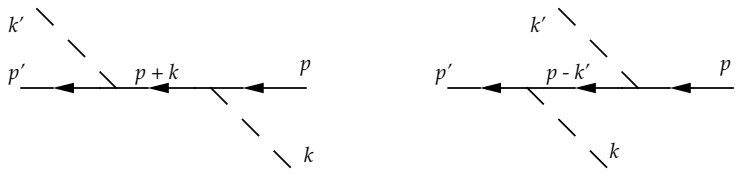

Figure 20.3 The two Feynman diagrams for Compton scattering in vacuo are illustrated.

On averaging over the initial states of polarization and summing over the final states of polarization, one has

$$w(\boldsymbol{p}, \boldsymbol{k}', \boldsymbol{k}) = \frac{\mu_0^2 e^4 X}{2\omega\omega'\varepsilon\varepsilon'}\, 2\pi\, \delta(\varepsilon' + \omega' - \varepsilon - \omega), \tag{20.11}$$

with, after a tedious calculation which is outlined at the end of this lecture,

$$X = \left(\frac{m^2}{s-m^2} + \frac{m^2}{u-m^2}\right)^2 + \frac{m^2}{s-m^2} + \frac{m^2}{u-m^2} - \frac{1}{4}\left(\frac{s-m^2}{u-m^2} + \frac{u-m^2}{s-m^2}\right). \tag{20.12}$$

3. The Cross Section

The cross section is defined by dividing the transition rate by the flux of particles (the number arriving per unit area per unit time), and performing the sums and integrals over the final state. In the center-of-mass frame the flux of incoming particles may be identified as their relative velocity $|\boldsymbol{v}_1 - \boldsymbol{v}_2|$ divided by the volume V. A more general expression for the flux may be obtained by noting that in the center-of-mass frame one has

$$|\boldsymbol{v}_1 - \boldsymbol{v}_2| = \frac{[(p_1 p_2)^2 - m_1^2 m_2^2]^{1/2}}{\varepsilon_1 \varepsilon_2}, \tag{20.13}$$

and that one may re-express $p_1 p_2$ in terms of $s = m_1^2 + m_2^2 + 2p_1 p_2$. Thus in the case of Compton scattering one has $p_1 \to p$, $p_2 \to k$, and hence the flux j is given by

$$j = \frac{s - m^2}{2\omega\varepsilon V}. \tag{20.14}$$

The differential cross section is then given by

$$d\sigma = \frac{2\omega\varepsilon}{s-m^2}\, \frac{\mu_0^2 e^4 X}{2\omega\omega'\varepsilon\varepsilon'}\, (2\pi)^4\, \delta^4(p' + k' - p - k)\, \frac{d^3\boldsymbol{p}'}{(2\pi)^3}\, \frac{d^3\boldsymbol{k}'}{(2\pi)^3}. \tag{20.15}$$

This is an invariant quantity; the factors $\omega\varepsilon$ cancel, and the differentials $d^3\boldsymbol{p}'/(2\pi)^3\varepsilon'$ and $d^3\boldsymbol{k}'/(2\pi)^3\omega'$ are invariants, as may be seen by considering the manifestly invariant integrals

$$\int \frac{d^4p'}{(2\pi)^4}\, 2\pi\, \delta(p'^2 - m^2) = \int \frac{d^3\boldsymbol{p}'}{(2\pi)^3}\, \frac{1}{2\varepsilon'}, \qquad \int \frac{d^4k'}{(2\pi)^4}\, 2\pi\, \delta(k'^2) = \int \frac{d^3\boldsymbol{k}'}{(2\pi)^3}\, \frac{1}{2\omega'}.$$

Let us choose the center-of-mass frame to evaluate the integral. In this frame one has $\boldsymbol{p} = -\boldsymbol{k}$ and $\boldsymbol{p}' = -\boldsymbol{k}'$, and hence $\varepsilon' = (m^2 + \omega'^2)^{1/2}$. The $\boldsymbol{k}'$-integral is written in spherical polar coordinates. The integral over $|\boldsymbol{k}'| = \omega'$ performed using the δ-function:

$$\int d\omega'\, \delta(\varepsilon' + \omega' - \varepsilon - \omega) = \frac{\varepsilon'}{\varepsilon + \omega}.$$

The integral over solid angle involves a trivial integral over azimuthal angle, which gives 2π, and an integral over $d\cos\chi$, where χ is the angle between $\boldsymbol{k}$ and $\boldsymbol{k}'$: $\boldsymbol{k}\cdot\boldsymbol{k}' = \omega\omega'\cos\chi$. This integral may be replaced by one over the invariant $t = (p-p')^2 = (k-k')^2$:

$$d\cos\chi = \frac{dt}{2\omega\omega'}. \tag{20.16}$$

The integral is then in a manifestly invariant form

$$d\sigma = \frac{\mu_0^2 e^4 X}{2\pi}\,\frac{m^2 dt}{(s-m^2)^2}. \tag{20.17}$$

It is convenient to introduce the *classical radius of the electron*; in normal units this is given by

$$r_0 = \frac{e^2}{4\pi\varepsilon_0 mc^2} = \frac{\mu_0 e^2}{4\pi m}. \tag{20.18}$$

3. The Klein Nishina Formula

The invariant form (20.17) may be used to derive the differential cross section in any particular frame, and it may be used to derive the total cross section.

Consider the frame in which the initial electron is at rest. Let the angle between $\boldsymbol{k}$ and $\boldsymbol{k}'$ in this frame be θ. By considering the square of $p' = p+k-k'$, with $p'^2 = p^2 = m^2$, $k'^2 = k^2 = 0$, in this frame one finds

$$m(\omega-\omega') - \omega\omega'(1-\cos\theta) = 0. \tag{20.19}$$

Also one has

$$s - m^2 = 2m\omega, \quad u - m^2 = -2m\omega', \quad t = -2\omega\omega'(1-\cos\theta). \tag{20.20}$$

The integral over t may be converted back into an integral over solid angle ($d^2\boldsymbol{\Omega}$ in this new frame) by writing

$$dt \to 2\omega\omega'\,\frac{d^2\boldsymbol{\Omega}}{2\pi}.$$

Then (20.17) becomes

$$d\sigma = \tfrac{1}{2}r_0^2\left(\frac{\omega'}{\omega}\right)^2\left(\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta\right)d^2\boldsymbol{\Omega}, \tag{20.21}$$

which is the differential form of the *Klein Nishina* cross section.

The non-quantum limit corresponds to $\omega' = \omega$ in (20.21). The resulting expression $d\sigma = \frac{1}{2}r_0^2\,(1+\cos^2\theta)\,d^2\boldsymbol{\Omega}$ is the differential cross section for *Thomson scattering*, which is the classical counterpart of Compton scattering. Technically, *Compton scattering* is the scattering of an electron by a photon, and *inverse Compton scattering* is the scattering

of a photon by an electron, which is what corresponds to Thomson scattering in the non-quantum limit. However, *Compton scattering* is also used in the generic sense to refer to the effects of electron-photon scattering.

The integral over solid angle in the non-quantum limit gives the *Thomson cross section*

$$\sigma_T = \frac{8\pi r_0^2}{3}. \tag{20.22}$$

The generalization of this to the fully relativistic quantum case may be obtained by performing the integral over t in (20.17). The general kinematic restrictions discussed above require

$$s + t + u = 2m^2, \quad \text{and} \quad su \geq m^4. \tag{20.23}$$

The integral then reduces to

$$\begin{aligned} \frac{8\pi r_0^2}{x^2} \int_{x/(1+x)}^{x} dy \left[\left(\frac{1}{x} - \frac{1}{y} \right)^2 + \frac{1}{x} - \frac{1}{y} + \frac{1}{4} \left(\frac{y}{x} - \frac{x}{y} \right) \right] \\ = \frac{2\pi r_0^2}{x} \left[\left(1 - \frac{4}{x} + \frac{8}{x^2} \right) \ln(1+x) + \frac{1}{2} + \frac{8}{x} - \frac{1}{2(1+x)^2} \right], \end{aligned} \tag{20.24}$$

with $x = (s - m^2)/m^2$. In the frame in which the initial electron is at rest (20.20) implies $x = 2\omega/m$.

4. Pair Annihilation into Two Photons

The annihilation of an electron/positron pair into two photons is one of the processes related to Compton scattering by crossing symmetry. The cross section for this and other crossed processes differ from (20.17) only in that the flux j of particles is different, and hence (20.14) needs to be revised appropriately.

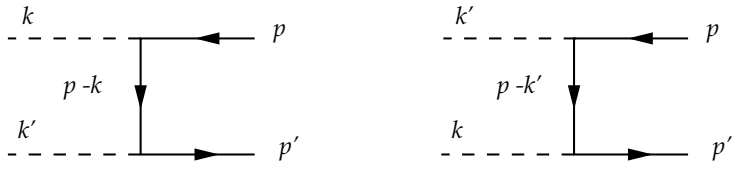

Figure 20.4 The two Feynman diagrams for pair annihilation are illustrated.

The Feynman diagrams for the annihilation of an electron/positron pair into two photons are illustrated in Figure 20.4. This annihilation process corresponds to the t-channel, whereas Compton scattering corresponds to the s-channel. In the s-channel the value of s is specified by the initial conditions, and either t or u is an independent variable. In the t-channel it is the value of t that is specified by the initial conditions and either s or u is an independent variable.

In the t-channel one has $p_1 \to p$, $p_2 \to p'$, and hence the flux j is given by

$$j = \frac{1}{\varepsilon\varepsilon' V} \left[(pp')^2 - m^4 \right]^{1/2} = \frac{1}{4\varepsilon\varepsilon' V} \, t(t - 4m^2). \tag{20.25}$$

Choosing the independent variable to be s the restrictions (20.4) and (20.7) imply that the limits of the allowed range of s for given t are determined by $su = m^4$ and $s+t+u = 2m^2$. The range is

$$-\tfrac{1}{2}t - \tfrac{1}{2}[t(t-4m^2)]^{1/2} \leq s - m^2 \leq -\tfrac{1}{2}t + \tfrac{1}{2}[t(t-4m^2)]^{1/2}. \tag{20.26}$$

The total cross section for the annihilation of a pair into two photons is then given by integrating

$$d\sigma = 8\pi r_0^2 \, \frac{m^2 X \, ds}{t(t-4m^2)} \tag{20.27}$$

over the range (20.26). The integral is elementary, although somewhat lengthy. The result is

$$\sigma = \frac{2\pi r_0^2}{\tau^2(\tau-4)} \left\{ (\tau^2+4\tau-8) \, \ln \frac{\tau^{1/2}+(\tau-4)^{1/2}}{\tau^{1/2}-(\tau-4)^{1/2}} - (\tau+4)[\tau(\tau-4)]^{1/2} \right\}, \tag{20.28}$$

with $\tau := t/m^2$.

In a frame in which the electron is at rest, and the positron has speed v and Lorentz factor γ, one has $\tau = 2(1+\gamma)$. For a nonrelativistic positron (20.28) gives

$$\sigma \approx \frac{\pi r_0^2}{(\tau-4)^{1/2}}, \tag{20.29}$$

with $(\tau-4)^{1/2} \approx v \ll 1$. The cross section diverges for $v \to 0$, that is, for slow positrons.

Fast positrons are slowed down by their interactions with electrons, and in practice the timescale for slowing down is faster than the timescale for annihilation in flight. Th slow positrons are then captured by electrons to form *positronium*, and annihilation of pairs is dominated by the annihilation of positronium. The rate at which positronium annihilates may be estimated using (20.29). In the limit $v \to 0$, one estimates the rate of decay by multiplying σv by the density of states factor $4/\pi a^3$ for positronium, where $a = 8\pi\varepsilon_0\hbar^2/me^2$ is the Bohr radius for the ground state. This gives a decay rate (in ordinary units)

$$\frac{1}{\tau_0} = \frac{mc^2\alpha^5}{2\hbar} \tag{20.30}$$

where $\alpha = e^2/4\pi\varepsilon_0\hbar c \approx 1/137$ is the fine structure constant. The half life of positronium is thus $\tau_0 = 1.23\ 10^{-10}$ s.

Appendix. Evaluation of the Matrix Element

The matix element that needs to be evaluated is the average over states of polarization of the modulus squared of, cf. (20.9),

$$M_{fi} = e^2 \, \bar{u}(\mathbf{p}') \left[\not{\epsilon} \, \frac{\not{p} - \not{k}'}{(p-k')^2 - m^2} \, \not{\epsilon}'^* + \not{\epsilon}'^* \, \frac{\not{p} + \not{k}}{(p+k)^2 - m^2} \, \not{\epsilon} \right] u(\mathbf{p}), \tag{20A.1}$$

where the arguments of $\not{e} = \gamma^\mu e_\mu(\mathbf{k})$ and $\not{e}'^* = \gamma^\nu e^*_\nu(\mathbf{k}')$ are omitted. Before squaring we can simplify in the following ways (Feynman 1962, pp. 125–129). First, evaluating the denominators for photons in vacuo ($k^2 = 0 = k'^2$) gives

$$(p - k')^2 - m^2 = -2pk', \qquad (p + k)^2 - m^2 = 2pk. \tag{20A.2}$$

Next, Dirac's equation implies $\not{p}\, u(\mathbf{p}) = mu(\mathbf{p})$, and the relation $\not{a}\not{b} = -\not{b}\not{a} + 2ab$ is used to write

$$\not{p}\not{e}'^* = -\not{e}'^*\not{p} + 2e'^* p, \qquad \not{p}\not{e} = -\not{e}\not{p} + 2ep. \tag{20A.3}$$

Thus one finds

$$M_{fi} = e^2\, \bar{u}(\mathbf{p}') \left[\frac{\not{e}\not{k}'\not{e}'^* - 2\not{e}e'^* p}{2pk'} + \frac{\not{e}'^*\not{k}\not{e} + 2\not{e}'^* ep}{2pk} \right] u(\mathbf{p}), \tag{20A.4}$$

The remaining simplification is to choose a convenient frame and gauge. In the frame in which the initial electron is at rest and in the temporal gauge, one has

$$pk' = m\omega', \qquad pk = m\omega, \qquad ep = 0 = e'^* p. \tag{20A.5}$$

Then one has

$$M_{fi} = \frac{e^2}{2m}\, \bar{u}(\mathbf{p}')\Phi u(\mathbf{p}), \qquad \Phi = \left[\frac{\not{e}\not{k}'\not{e}'^*}{\omega'} + \frac{\not{e}'^*\not{k}\not{e}}{\omega} \right]. \tag{20A.6}$$

The probability for unpolarized electrons is proportional to

$$\tfrac{1}{2}\overline{M^*_{fi} M_{fi}} = \tfrac{1}{2} \left(\frac{e^2}{2m} \right)^2 \mathrm{Tr}\left[\bar{\Phi}(\not{p}' + m)\Phi(\not{p} + m) \right], \tag{20A.7}$$

where the overline indicates the sum over the two spin states, and with $p' = p + k - k'$. The trace can be separated into four different terms by writing (20A.7) in the form

$$\tfrac{1}{2}\, \overline{M^*_{fi} M_{fi}} = \frac{e^4}{8m^2} \left(\frac{A}{\omega'^2} + \frac{B + C}{\omega'\omega} + \frac{D}{\omega^2} \right), \tag{20A.8}$$

$$\begin{aligned} A &= \mathrm{Tr}\left[\not{e}'\not{k}'\not{e}^*(\not{p}' + m)\not{e}\not{k}'\not{e}'^*(\not{p} + m) \right], \\ B &= \mathrm{Tr}\left[\not{e}'\not{k}'\not{e}^*(\not{p}' + m)\not{e}'^*\not{k}\not{e}(\not{p} + m) \right], \\ C &= \mathrm{Tr}\left[\not{e}^*\not{k}\not{e}'(\not{p}' + m)\not{e}\not{k}'\not{e}'^*(\not{p} + m) \right], \\ D &= \mathrm{Tr}\left[\not{e}^*\not{k}\not{e}'(\not{p}' + m)\not{e}'^*\not{k}\not{e}(\not{p} + m) \right]. \end{aligned} \tag{20A.9}$$

The problem is thereby reduced to evaluating the four traces in (20A.9).

The evaluation of the traces involves using

$$\mathrm{Tr}[\not{a}\not{b}] = 4ab, \qquad \mathrm{Tr}[\not{a}\not{b}\not{c}\not{d}] = 4(ab\; cd - ac\; bd + ad\; bc), \tag{20A.10}$$

and the fact that the trace of an odd number of γ-matrices is zero. Feynman (1962) used $\not{e}^*\not{e} = -1$, but this is correct only for linearly polarized radiation. In the following this

assumption is not made. The gauge condition is chosen so that $ep = 0$ and $e'p = 0$ are satisfied, and then one has $\not{p}\not{e} = -\not{e}\not{p}$, $\not{p}\not{e}' = -\not{e}'\not{p}$. One also has $ek = 0 = e'k'$, and hence $p' = p + k - k'$ then implies $ep' = -ek'$, $e'p' = e'k$.

One may evaluate A by moving the two factors $\not{k}'$ together, and using $\not{k}'\not{k}' = k'^2 = 0$. Hence one finds

$$A = \mathrm{Tr}\,[\not{e}'\not{k}'\not{e}^*(\not{p}' + m)\not{e}\not{k}'\not{e}'^*(\not{p} + m)] = -2pk'\,[a_1 + a_2 + a_3], \tag{20A.11}$$

$$a_1 = \mathrm{Tr}[\not{e}'\not{k}'\not{e}^*\not{p}\not{e}\not{e}'^*], \qquad a_2 = \mathrm{Tr}[\not{e}'\not{k}'\not{e}^*\not{k}\not{e}\not{e}'^*], \qquad a_2 = -\mathrm{Tr}[\not{e}'\not{k}'\not{e}^*\not{k}'\not{e}\not{e}'^*]. \tag{20A.12}$$

Explicit evaluation gives

$$\begin{aligned} a_1 &= -pk'\,\mathrm{Tr}[\not{e}'\not{e}^*\not{e}\not{e}'^*]), \\ a_2 &= -e'^*k\,\mathrm{Tr}[\not{e}\not{e}'\not{k}'\not{e}^*] + e'k\,\mathrm{Tr}[\not{e}\not{e}'^*\not{k}'\not{e}^*] - kk'\,\mathrm{Tr}[\not{e}\not{e}'^*\not{e}'\not{e}^*], \\ a_3 &= 2e^*k'\,\mathrm{Tr}[\not{e}'\not{k}'\not{e}\not{e}'^*]. \end{aligned} \tag{20A.13}$$

The final term in a_2 may be combined with a_1 using $pk' + kk' = p'k' = pk$, where $k'^2 = 0$ is used. Finally, one evaluating the remaining traces, one finds

$$\begin{aligned} A = {} & 8\,pk\ pk'\ (|e^*e'|^2 - |ee'|^2 + 1) \\ & - 16\,pk'\ \big[\mathrm{Re}\,(ee'^*\ e^*k'\ e'k - ee'\ e^*k'\ e'^*k) - |ek'|^2\big]. \end{aligned} \tag{20A.14}$$

The evaluation of D is simliar, and gives

$$\begin{aligned} D = {} & 8\,pk\ pk'\ (|e^*e'|^2 - |ee'|^2 + 1) \\ & + 16\,pk'\ \big[\mathrm{Re}\,(ee'^*\ e^*k'\ e'k - ee'\ e^*k'\ e'^*k) + |e'k|^2\big]. \end{aligned} \tag{20A.15}$$

Evaluating B is a similar fashion, one writes

$$\begin{aligned} B &= B_1 + B_2 + B_3, \\ B_1 &= \mathrm{Tr}[\not{e}'\not{k}'\not{e}^*(\not{p} + m)\not{e}'^*\not{k}\not{e}(\not{p} + m)], \\ B_2 &= \mathrm{Tr}[\not{e}'\not{k}'\not{e}^*\not{k}\not{e}'^*\not{k}\not{e}\not{p}], \\ B_3 &= -\mathrm{Tr}[\not{e}'\not{k}'\not{e}^*\not{k}'\not{e}'^*\not{k}\not{e}\not{p}]. \end{aligned} \tag{20A.16}$$

Moving $\not{p}$ to the right in B_1, bringing the factors $\not{k}$ together in B_2 and the factors $\not{k}'$ together in B_3 gives

$$\begin{aligned} B_1 &= 2\,pk\ pk'\ \mathrm{Tr}[\not{e}'\not{e}^*\not{e}'^*\not{e}], \\ B_2 &= 2\,e'^*k\,\big\{ -pk\ \mathrm{Tr}[\not{e}'\not{k}'\not{e}^*\not{e}] - pk'\ \mathrm{Tr}[\not{e}'\not{e}^*\not{k}\not{e}]\big\}, \\ B_3 &= -2\,e^*k'\,\big\{ -pk\ \mathrm{Tr}[\not{e}'\not{k}'\not{e}'^*\not{e}] - pk'\ \mathrm{Tr}[\not{e}'\not{e}'^*\not{k}\not{e}]\big\}. \end{aligned} \tag{20A.17}$$

On evaluating the remaining traces, one finds

$$\begin{aligned} B = {} & 8\,pk\ pk'\ (|e^*e'|^2 - 1 + |ee'|^2) - 8\,pk'\,|e'k|^2 + 8\,pk\,|ek'|^2 \\ & - 8\,pk\ e'^*k\,(ee'\ e^*k' - e^*e'\ ek') + 8\,pk'\ e^*k'\,(ee'\ e'^*k - ee'^*\ e'k). \end{aligned} \tag{20A.18}$$

The evaluation of C is similar and leads to

$$C = 8\,pk\ pk'\ (|e^*e'|^2 - 1 + |ee'|^2) - 8\,pk'\,|e'k|^2 + 8\,pk\,|ek'|^2 \\ + 8\,pk\ e'k\,(ee'^*\ e^*k' - e^*e'^*\ ek') - 8\,pk'\ ek'\,(e^*e'\ e'^*k - e^*e'^*\ e'k)\big]. \tag{20A.19}$$

On collecting terms, one finds, cf. (20A.8),

$$\frac{A}{(pk')^2} + \frac{B+C}{pk\ pk'} + \frac{D}{(pk)^2} = 8\left(\frac{pk}{pk'} + \frac{pk'}{pk}\right)(|e^*e'|^2 - |ee'|^2 + 1) + 16(|e^*e'|^2 - 1 + |ee'|^2). \tag{20A.20}$$

In the case of linearly polarized photons one has $|ee'|^2 = |e^*e'|^2$ and (20A.20) simplifies to the result obtained by Feynman (1962). Jauch and Rohrlich (1980, p. 231) wrote down a generalization for an arbitrary gauge, but their result is also restricted to the case of linear polarization.

For an unpolarized electron one sums over the initial states of polarization and averages over the final states of polarization. In (20A.20) this involves multiplying the polarization-independent factors by 2 and making the replacements $|e^*e'|^2 \to \frac{1}{2}(1+\cos^2\theta)$, $|ee'|^2 \to \frac{1}{2}(1+\cos^2\theta)$, where θ is the scattering angle. Thus one has

$$\left(\frac{A}{\omega'^2} + \frac{B+C}{\omega\omega'} + \frac{D}{\omega^2}\right)_{\text{unpol}} = 16\left(\frac{\omega}{\omega'} + \frac{\omega'}{\omega} - \sin^2\theta\right) = 64X, \tag{20A.21}$$

where we now revert to the frame in which the initial electron is at rest. This reproduces the expression that appears in (20.21). The final expression, $64X$, in (20A.21) follows by comparison with (20.12) after substituting the invariant forms (20.20).

Lecture 2

4-Tensor Notation

The 4-tensor notation used here has greek indices running over $0, 1, 2, 3$ or t, x, y, z. The signature is -2, i.e., the metric tensor is diagonal with components $1, -1, -1, -1$. Where appropriate latin indices are used to denote the space components $1, 2, 3$ or x, y, z. To introduce this notation in a formal way, let us define what is meant by a 4-tensor equation.

1. 4-Tensor Equations

A 4-tensor equation involves elements which are either kernel symbols, or products of kernel symbols, with each symbol having zero, one or more indices. The indices are written in space (one space per index) after the kernel symbol, and any index may be either raised (i.e., as a superscript), denoting a *contravariant component*, or lowered (i.e., as a subscript), denoting a *covariant component*. The indices may have affices, e.g., primes or numerical or other subscripts, and two indices are the same only if they have the same affices. In each element of a 4-tensor equation, an index occurs only either once, when it is called a *free index*, or twice, when it is called a *dummy index.* Each pair of dummy indices must consist of one raised and one lowered index. The *summation convention* is that the sum (from 0 to 3) over each pair of dummy indices is implied.

Each kernel symbol may be regarded as describing a tensor. The *rank* of a tensor is defined as the number of its free indices; 4-tensors of rank zero are called invariants and 4-tensors of rank one are called 4-vectors. Similarly the rank of a tensor equation is equal to the number of the free indices in each of its elements. The number and kind of free indices must be the same in all elements of a tensor equation.

There are three elementary manipulations which may be performed on any tensor equation: raising or lowering an index, relabeling indices and contracting over two indices. A lowered index ν may be converted into a raised index μ by using the contravariant form $g^{\mu\nu}$ of the metric tensor, and a raised index ν may be converted into a lowered index μ by using the covariant form $g_{\mu\nu}$ of the metric tensor. For a 4-vector a, these operations are

$$a^\mu = g^{\mu\nu} a_\nu, \qquad a_\mu = g_{\mu\nu} a^\nu. \tag{2.1}$$

One may relabel any free index, provided the relabeling is made in every element of the tensor equation. One may relabel any pair of dummy indices, and also interchange the raised and lowered index. A *contraction* may be performed on any tensor equation of rank two or higher. It involves converting two free indices into a pair of dummy indices, thereby reducing the rank of the equation by two.

It is convenient to write a 4-vector a in terms of its time component a^0 and its space components in the form of a 3-vector $\boldsymbol{a}$:

$$a^\mu = [a^0, \boldsymbol{a}], \qquad a_\mu = [a^0, -\boldsymbol{a}] \tag{2.2}$$

Note that the three Cartesian components of the 3-vector $\boldsymbol{a}$ are identified with the contravariant space components a^1, a^2, a^3 of the 4-vector, and that the covariant space components a_1, a_2, a_3 of the 4-vector are equal to minus the Cartesian components of $\boldsymbol{a}$.

2. Important 4-Vectors

The basic 4-vector is a space-time point $x^\mu = [t, \boldsymbol{x}]$, also called an *event*. (Note that natural units with $c = 1$ are used here, so that the basic unit of distance is a light second.) The following 4-vectors appear frequently below:

$$\text{event:} \qquad x^\mu = [t, \boldsymbol{x}], \tag{2.3}$$

$$\text{4-velocity:} \qquad u^\mu = [\gamma, \gamma\boldsymbol{v}], \tag{2.4}$$

$$\text{4-momentum:} \qquad p^\mu = [\varepsilon, \boldsymbol{p}], \tag{2.5}$$

$$\text{wave 4-vector:} \qquad k^\mu = [\omega, \boldsymbol{k}], \tag{2.6}$$

$$\text{4-current density:} \qquad J^\mu = [\rho, \boldsymbol{J}], \tag{2.7}$$

$$\text{4-potential:} \qquad A^\mu = [\phi, \boldsymbol{A}], \tag{2.8}$$

where $\boldsymbol{v}$ is the 3-velocity, $\gamma = (1 - v^2)^{-1/2}$ is the *Lorentz factor*, $\varepsilon = \gamma m$ is the energy, $\boldsymbol{p}$ is the 3-momentum, ω is the frequency, $\boldsymbol{k}$ is the wave vector, ρ is the charge density, $\boldsymbol{J}$ is the 3-current density, ϕ is the electric potential and $\boldsymbol{A}$ is the vector potential. Another important 4-vector quantity is the operator

$$\text{4-gradient:} \qquad \partial_\mu := \partial/\partial x^\mu = [\partial/\partial t, \partial/\partial \boldsymbol{x}], \tag{2.9}$$

which appears naturally in terms of covariant components. (The symbol ":=" denotes a definition.)

The invariant formed from two 4-vectors a and b is denoted ab:

$$ab := a^\mu b_\mu = a^0 b^0 - \boldsymbol{a} \cdot \boldsymbol{b}. \tag{2.10}$$

Similarly, the invariant formed from a single 4-vector a is $a^2 = (a^0)^2 - \boldsymbol{a}^2$.

3. Lorentz Transformations

The 4-tensor character of a physical quantity is defined in terms of its transformation properties under Lorentz transformations. Let K and K' be two inertial frames. An event in K is described by x^μ and the same event is described in K' by $x^{\mu'}$. These components are related by an equation of the form

$$x^{\mu'} = a^{\mu'} + L^{\mu'}{}_{\mu}\, x^\mu. \tag{2.11}$$

The constant 4-vector $a^{\mu'}$ relates the origins in space and time in K and K'; it corresponds to a generalization of a translational motionfrom three to four dimensions. The other term in (2.1) corresponds to a generalization of a rotation from three to four dimensions. In most cases below, a Lorentz transformation refers only to the rotational part. The translational part is of interest here only for some formal purposes, and is otherwise neglected.

The coefficients $L^{\mu'}{}_{\nu}$ may be written in matrix form. Note however, that they are not tensors: they have two indices but these two indices refere to different frames of reference. Irrespective of whether the indices are raised or lowered, the matrix convention is that

the first-written index labels rows and the second-written index labels columns. In order to preserve the invariant a^2 the determinant of this matrix must be equal to unity to within a sign. For *proper* Lorentz transformations this sign is positive; improper Lorentz transformations involve either reflection of a coordinate axis (parity transformation) or of the time axis (time reversal transformation). The inverse of the transformation matrix $L^{\mu'}{}_{\mu}$ is is written $L^{\mu}{}_{\mu'}$. Then for any 4-vector a^μ the transformation properties of the contravariant ant covariant components are

$$a^{\mu'} = L^{\mu'}{}_{\mu}\, a^\mu, \qquad a^\mu = L^{\mu}{}_{\mu'}\, a^{\mu'}, \tag{2.12}$$

$$a_{\mu'} = L^{\mu}{}_{\mu'}\, a_\mu \qquad a_\mu = L^{\mu'}{}_{\mu}\, a_{\mu'}. \tag{2.13}$$

The transformation matrices satisfy

$$L^{\mu}{}_{\mu'}\, L^{\mu'}{}_{\nu} = \delta^{\mu}{}_{\nu}, \qquad L^{\mu'}{}_{\mu}\, L^{\mu}{}_{\nu'} = \delta^{\mu'}{}_{\nu'}, \tag{2.14}$$

where

$$\delta^{\mu}{}_{\nu} = \begin{cases} 1 & \text{for } \mu = \nu, \\ 0 & \text{for } \mu \neq \nu. \end{cases} \tag{2.15}$$

is the unit 4-tensor. (Note that it is conventional to use $\delta^{\mu}{}_{\nu}$ rather than the mixed components $g^{\mu}{}_{\nu}$ of the metric tensor.)

An equation in 4-tensor form is said to be in a *manifestly covariant* form. This means that the form is obviously unchanged under a Lorentz transformation, so that the equation manifestly satisfies the requirement of the special theory of relativity. Under a transformation from frame K to frame K', a tensor equation transforms simply by adding primes to all the free indices.

In the case in which the axes in K and K' are parallel, and K' is moving along the z-axis of K at velocity $\boldsymbol{\beta}$, the explicit forms for the transformation matrices are

$$L^{\mu'}{}_{\mu} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix}, \qquad L^{\mu}{}_{\mu'} = \begin{pmatrix} \gamma & 0 & 0 & \beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \beta\gamma & 0 & 0 & \gamma \end{pmatrix}, \tag{2.16}$$

with $\gamma = (1 - \beta^2)^{-1/2}$.

4. Covariant Form of Maxwell's Equations

Maxwell's equations may be written in the covariant form

$$\partial^\mu\, F^{\nu\rho}(x) + \partial^\rho\, F^{\mu\nu}(x) + \partial^\nu\, F^{\rho\mu}(x) = 0, \tag{2.17}$$

$$\partial_\mu\, F^{\mu\nu}(x) = \mu_0\, J^\nu(x), \tag{2.18}$$

where $F^{\mu\nu}(x)$ is the Maxwell tensor, and where the argument x denotes $(t, \boldsymbol{x})$. The Maxwell tensor is antisymmetric

$$F^{\mu\nu}(x) = -F^{\nu\mu}(x), \tag{2.19}$$

and is related to the Cartesian components of the electric field $\boldsymbol{E}$ and the magnetic induction $\boldsymbol{B}$ by

$$F^{\mu\nu}(x) := [\boldsymbol{E}, \boldsymbol{B}] = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{2.20}$$

(The Cartesian components of the 3-vectors are denoted by x, y, z rather than $1, 2, 3$ to emphasize that relations such as (2.20) are between two different notations; such relations are prescriptions for relating the quantities in the two different notations.) The first of the pair of Maxwell's equations may be written more concisely in terms of the dual of the Maxwell tensor. The dual of any second rank tensor $T^{\mu\nu}$ is defined by

$$T^{D\,\mu\nu} = \tfrac{1}{2}\,\epsilon^{\mu\nu\alpha\beta}\,T_{\alpha\beta}, \tag{2.21}$$

where $\epsilon^{\alpha\beta\gamma\delta}$ is the permutation symbol:

$$\epsilon^{\alpha\beta\gamma\delta} = \begin{cases} +1 & \text{for } \alpha\beta\gamma\delta \text{ an even permutation of } 0123, \\ -1 & \text{for } \alpha\beta\gamma\delta \text{ an odd permutation of } 0123, \\ 0 & \text{otherwise.} \end{cases} \tag{2.22}$$

Note that the permutation symbol with lowered indices ϵ_{0123} is numerically equal to minus ϵ^{0123} (proof: $\epsilon^{0123} = g^{00}g^{11}g^{22}g^{33}\epsilon_{0123}$, and $g^{00} = 1$, $g^{11} = g^{22} = g^{33} = -1$). Many authors choose the sign convention $\epsilon_{0123} = 1$, which is opposite to that chosen here.

The dual of the Maxwell tensor is

$$F^{D\,\mu\nu}(x) = [\boldsymbol{B}, -\boldsymbol{E}] = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z & -E_y \\ B_y & -E_z & 0 & E_x \\ B_z & E_y & -E_x & 0 \end{pmatrix}. \tag{2.23}$$

Then (2.17) may be replaced by

$$\partial_\mu\, F^{D\,\mu\nu}(x) = 0. \tag{2.24}$$

There are two independent invariants that may be constructed from the Maxwell tensor. These are

$$F^{\mu\nu}\,F_{\mu\nu} = -2\,(\boldsymbol{E}^2 - \boldsymbol{B}^2) \tag{2.25}$$

and

$$F^{\mu\nu}\,F^{D}{}_{\mu\nu} = -4\,\boldsymbol{E}\cdot\boldsymbol{B}. \tag{2.26}$$

5. Continuity Equations for Charge and Energy

On operating on (2.18) with ∂_ν, the antisymmetry property (2.19) implies the equation of charge continuity

$$\partial_\mu J^\mu(x) = 0. \tag{2.27}$$

There is another continuity equation that may be derived directly from Maxwell's equations. This is

$$\partial_\mu \Theta^{\mu\nu}(x) = J_\alpha(x) F^{\alpha\nu}(x), \tag{2.28}$$

where

$$\Theta^{\mu\nu}(x) = \varepsilon_0 (F^\mu{}_\alpha F^{\alpha\nu} + \tfrac{1}{4} g^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta}) \tag{2.29}$$

is the symmetric energy-momentum tensor. (Note that we have $\mu_0 \varepsilon_0 = 1$ in natural units so that either ε_0 or $1/\mu_0$ may be used interchangeably.) On interpreting (2.28) as a continuity equation for energy, one identifies the energy density, W, momentum density, $\boldsymbol{P}$, energy flux, $\boldsymbol{F}$ and the stress 3-tensor, $\boldsymbol{T}$, for the electromagnetic field in vacuo:

$$\begin{array}{ll} W = \frac{1}{2}\varepsilon_0(\boldsymbol{E}^2 + \boldsymbol{B}^2), & \boldsymbol{P} = \varepsilon_0 \boldsymbol{E} \times \boldsymbol{B}, \\ \boldsymbol{F} = \varepsilon_0 \boldsymbol{E} \times \boldsymbol{B}, & \boldsymbol{T} = W\mathbf{1} - \varepsilon_0(\boldsymbol{E}\boldsymbol{E} + \boldsymbol{B}\boldsymbol{B}), \end{array} \tag{2.30}$$

where $\mathbf{1}$ denotes the unit 3-tensor.

6. Gauge Transformations

Equation (2.17) may be satisfied exactly by expressing $F^{\mu\nu}$ in terms of the 4-potential A^μ:

$$F^{\mu\nu}(x) = \partial^\mu A^\nu(x) - \partial^\nu A^\mu(x). \tag{2.31}$$

The choice $A(x)$ of 4-potential is not unique, and any other choice $A'(x)$ related to $A(x)$ by a gauge transformation is equally acceptable. A gauge transformation is of the form

$$A'^\mu(x) = A^\mu(x) + \partial^\mu \psi(x), \tag{2.32}$$

where $\psi(x)$ is an arbitrary differentiable function.

The freedom to make gauge transformations allows one to impose a gauge condition. All relevant gauge conditions are of the form

$$\hat{G}_\alpha A^\alpha(x) = 0, \tag{2.33}$$

where $\hat{G}_\alpha$ is a differential operator in general. Specific gauge conditions correspond to

$$\text{Lorentz gauge :} \qquad \hat{G}_\alpha = \partial_\alpha, \tag{2.34}$$

$$\text{Coulomb gauge :} \qquad \hat{G}_\alpha = [0, \partial/\partial \boldsymbol{x}], \tag{2.35}$$

$$\text{temporal gauge :} \qquad \hat{G}_\alpha = [1, \mathbf{0}]. \tag{2.36}$$

It is desirable to make the general theory independent of the choice of gauge as far as is possible. An equation that maintains its form under a gauge transformation is said to be in a *manifestly gauge-independent* form. It is desirable to develop a form of QED that is manifestly gauge-independent, as well as being manifestly covariant.

Lecture 3

The Klein Gordon and Dirac Equations

When considering relativisitic generalizations of nonrelativistic quantum mechanics, it is convenient to choose the coordinate representation and the Schrödinger picture. Then the wavefunction depends on t and $\boldsymbol{x}$ and these may be regarded as the 4-vector x. The immediate problem encountered in seeking a relativistic generalization of the Schrödinger equation (1.17) is that the Hamiltonian for a relativistic particle includes a square root, and there is no obvious way of taking the square root of an operator. In the Klein Gordon equation this is overcome by considering an equation involving a second order time derivative, and in the Dirac equation it is overcome by introducing matrices that effectively allow one to take the square root.

1. The Klein Gordon Equation

The Hamiltonian for a free relativistic particle is

$$H = [m^2 + \boldsymbol{p}^2]^{1/2}. \tag{3.1}$$

Schrödinger's equation (1.17) may be written $(i\partial/\partial t - \hat{H}), \Psi(x) = 0$, and if one premultiplies this by $(i\partial/\partial t + \hat{H})$, the square of the Hamiltonian appears. On inserting the forms (3.1) and adoptinng the coordinate representation ($\boldsymbol{p}^2 \to -\nabla^2$), one obtains an equation that may be written in the covariant form (in natural units)

$$(\partial^\mu \partial_\mu + m^2)\, \Psi(x) = 0. \tag{3.2}$$

This is the *Klein Gordon equation.* An electromagnetic field, described by its 4-potential $A(x)$, is included by making the *minimal coupling* replacement

$$\partial^\mu \to D^\mu := \partial^\mu + iqA^\mu(x) \tag{3.3}$$

in (3.2).

The fact that the Klein Gordon equation is second order in the time derivative makes it qualitatively different from the Schrödinger equation, which is first order in the time derivative. To follow the time evolution, the Klein Gordon equation requires that both the wavefunction and its time derivative be specified initially, and it is not obvious how one reconciles such an evolution with that implied by a first order equation for which only the wavefunction itself needs to be specified initially. Another difficulty with the Klein Gordon equation is that it has both positive and negative energy solutions. The currently accepted interpretation is that the negative energy solutions describe antiparticles, and that the two initial conditions that need to be imposed are equivalent to specifying the initial values of the wavefunctions for the particles and the antiparticles.

The Klein Gordon equation describes particles of spin zero. If $\Psi(x)$ is replaced by a 4-vector $\Psi^\mu(x)$ with $\partial_\mu \Psi^\mu(x) = 0$, then the resulting equation describes particles of spin one. Thus the Klein Gordon equation is unsatisfactory as a relativistic generalization of the Schrödinger equation for electrons, which have spin $\frac{1}{2}$.

2. The Dirac Equation

Dirac wrote down an equation of first order in the time derivative. An alternative form of the equation he wrote down is

$$(i\gamma^\mu \partial_\mu - m)\,\Psi(x) = 0. \tag{3.4}$$

This may be written more concisely as

$$(i\not{\partial} - m)\,\Psi(x) = (\not{p} - m)\,\Psi(x) = 0, \tag{3.5}$$

where the slash notation is defined by

$$\not{A} := \gamma^\mu A_\mu \tag{3.6}$$

for any 4-vector A. The quantities γ^μ are postulated to be such that on operating on (3.4) with $(i\not{\partial} + m)$ the resulting equation reduce to the Klein Gordon equation (3.2). This requires

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}. \tag{3.7}$$

On seeking a matrix representation, one finds that the minimum dimensionality allowing (3.7) to be satisfied is four. There is considerable freedom in choosing explicit representation of these 4×4 γ-matrices.

One specific choice is the *standard representation* which corresponds to

$$\begin{aligned}
\gamma^0 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \\
\gamma^2 &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \qquad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.
\end{aligned} \tag{3.8}$$

A convenient way of writing these and related 4×4 matrices is in terms of block matrices. Let $\mathbf{0}$ and $\mathbf{1}$ be the null and unit 2×2 matrices. Then on writing

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma} \end{pmatrix}, \quad \rho_x = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}, \quad \rho_y = \begin{pmatrix} \mathbf{0} & -i\mathbf{1} \\ i\mathbf{1} & \mathbf{0} \end{pmatrix}, \quad \rho_z = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \tag{3.9}$$

an alternative form of (3.8) is

$$\gamma^\mu = [\rho_z, i\rho_y\boldsymbol{\Sigma}]. \tag{3.10}$$

3. The Dirac Hamiltonian

The Dirac Hamiltonian may be identified by rewriting (3.4) in the form (1.17). This is achieved by multiplying (3.4) by γ^0 and using $(\gamma^0)^2 = 1$, as implied by (3.7). One finds

$$(\hat{p}^0 - \hat{H})\,\Psi(x) = 0, \tag{3.11}$$

with $\hat{p}^0 = i\partial/\partial t$. (The form (3.11) is the one actually written down originally by Dirac.) The Hamiltonian is identified as

$$\hat{H} = \boldsymbol{\alpha}\cdot\hat{\mathbf{p}} + \beta m, \tag{3.12}$$

with

$$\boldsymbol{\alpha} := \gamma^0\boldsymbol{\gamma}, \qquad \beta := \gamma^0, \tag{3.13}$$

and with $\hat{\mathbf{p}} = -i\partial/\partial\boldsymbol{x}$ in the coordinate representation. In the standard representation one has

$$\boldsymbol{\alpha} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}. \tag{3.14}$$

The Hamiltonian is an operator corresponding to an observable (the energy) and hence must be self adjoint, or hermitian in the present context. This implies a restriction on acceptable choices of the γ-matrices. Any choice must lead to hermitian forms for $\boldsymbol{\alpha}$ and β; that is, one must have

$$\boldsymbol{\alpha}^\dagger = \boldsymbol{\alpha}, \qquad \beta^\dagger = \beta, \tag{3.15}$$

and this requires

$$(\gamma^0)^\dagger = \gamma^0, \qquad \boldsymbol{\gamma}^\dagger = -\boldsymbol{\gamma}. \tag{3.16}$$

The conditions (3.16) may be expressed more concisely in the form

$$(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0. \tag{3.17}$$

4. The Dirac Conjugate

The adjoint of the Dirac equation in the form (3.4) is

$$\Psi^\dagger(x)\,(\hat{p}^0 - \boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}} - \beta m) = 0, \tag{3.18}$$

where the operators operate to the left. This may be written in the form

$$\overline{\Psi}(x)\,(\not{p} - m) = 0 \tag{3.19}$$

provided we define the "Dirac adjoint" by

$$\overline{\Psi}(x) := \Psi^\dagger(x)\gamma^0. \tag{3.20}$$

5. Particles and Antiparticles

The probability density $\Psi^*(x)\Psi(x)$ for the Klein Gordon field or $\overline{\Psi}(x)\gamma^0\Psi(x)$ for the Dirac field need not be positive. Historically this was one of the difficulties in the interpretation of the Klein Gordon equation. Dirac suggested that the negative energy solutions could be interpreted physically by postulating that in the vacuum all negative energy states are filled. This is sometimes called the *sea* of negative energy states. Dirac suggested that a sufficiently energetic photon ($\varepsilon = \hbar\omega > 2mc^2$) could knock an electron from the negative energy sea into a positive energy state. This produces an electron and a *hole*. The hole acts like a positively charged electron. Dirac's hole theory thus predicts the existence of the positron, which is an anti-electron. It is now accepted that every particle has a corresponding antiparticle, with the only exceptions being strictly neutral particles, which are their own antiparticles.

Dirac's hole theory relies on Fermi statistics: the negative energy sea can be filled because there can be only one electron per state according to the Pauli exclusion principle. Taken literally the hole theory presents several serious difficulties. One is that the vacuum would have an infinite charge density and an infinite energy density. Another is that is does not apply to the negative energy states of particles with even spin which obey Bose Einstein statistics. These difficulties are overcome in quantum field theory, essentially by formulating the theory to remove the difficulties by redefining the vacuum and by requiring a symmetry between particle and antiparticle operators.

Nowadays the negative probability density is simply accepted. It is interpreted by giving antiparticles a negative number density, so that a negative probability density implies that a positron is more likely to be found than an electron. The Klein Gordon and Dirac theories are thus interpreted as single charge rather than single particle theories. The creation of particle/antiparticle pairs is allowed and thereby allows the possibility of negative probability densities.

Lecture 4

Plane Wave Solutions

Free particles are represented by plane waves. Plane wave solutions of the Klein Gordon and Dirac equations are written down in this lecture.

1. Free Spin 0 and Spin 1 Particles

A plane wavefunction may be written in the form

$$\Psi(x) = \varphi(P)\, e^{-iPx}, \tag{4.1}$$

where P^μ denotes a set of quantum numbers that form a 4-vector.

A plane wave solution for a spin 0 satisfies the Klein Gordon equation. On substituting (4.1) into (3.2) one obtains

$$(P^2 - m^2)\, \varphi(P) = 0. \tag{4.2}$$

In general $\varphi(P)$ is not zero, and for a solution of (4.1) to exist, one requires $P^2 = m^2$. On solving $P^2 = m^2$ for the energy eigenvalues, the solutions are

$$P^0 = \pm\, [m^2 + \boldsymbol{P}^2]^{1/2}. \tag{4.3}$$

It is convenient to write the sign as $\epsilon = \pm$ and $P^\mu = \epsilon p^\mu$.

The normalization condition is to one particle in the normalization volume. This requires that the spatial integral of the energy density be equal to ε. The energy density is the $\mu = 0$, $\nu = 0$ component of the energy momentum tensor $T^{\mu\nu}(x)$, which is derived in the next lecture. The normalization requires

$$\int d^3\boldsymbol{x} \left[(\partial_0 \Psi^*)(\partial_0 \Psi) + (\mathrm{grad}\Psi^*) \cdot (\mathrm{grad}\Psi) + m^2 \Psi^* \Psi \right] = \varepsilon. \tag{4.4}$$

On inserting (4.1), the left hand side gives $V[(p^0)^2 + \boldsymbol{P}^2 + m^2]|\phi(P)|^2$, and hence the normalized solution corresponds to

$$\varphi(P) = \frac{1}{\sqrt{2\varepsilon V}}\ , \tag{4.5}$$

where V is the normalization volume, and with $\varepsilon = [m^2 + \boldsymbol{p}^2]^{1/2}$.

2. Free Spin $\frac{1}{2}$ Particles

In the case of the Dirac equation, $\varphi(P)$ in (4.1) is a Dirac spinor, that is, a column matrix with four elements. It must satisfy

$$(\not{P} - m)\varphi(P) = 0. \tag{4.6}$$

One may regard (4.6) as a set of four coupled equations for the four spinor components of $\varphi(P)$. The condition for a solution to exist is that the determinant of the coefficients vanish:

$$\det(\not{P} - m) = (P^2 - m^2)^2 = 0. \tag{4.7}$$

The solutions are given by (4.3), and the fact that $(P^2 - m^2)$ appears squared in (4.7) indicates that the solutions are doubly degenerate.

Solutions may be constructed from the matrix of cofactors of $\not{P} - m$. The matrix of cofactors of a matrix with elements A_{ij} is the matrix a_{ij} such that the determinant of $\det A$ of A_{ij} is given by $A_{ij}a_{jl} = \delta_{il} \det A$. Note that a_{ij} is the cofactor of A_{ji}. The matrix of cofactors is proportional to $\not{P} + m$. Four independent solutions may be obtained by choosing any two columns of $\not{P} + m$ (or any two linearly independent combinations of the four columns), and setting $P = \epsilon p$. The normalization is to one particle in the normalization volume. With the form of $T^{\mu\nu}(x)$ derived in the next lecture, one requires

$$\int d^3\boldsymbol{x} \Big[-\tfrac{i}{2}(\partial^0\overline{\Psi})\gamma^0\Psi + \tfrac{i}{2}\overline{\Psi}\gamma^0(\partial^0\Psi)\Big] = \epsilon\varepsilon. \tag{4.8}$$

Note that the sign ϵ needs to be included on the right hand side of (4.8) but not of (4.4) because $T^{00}(x)$ is and odd function of p^0 for spin $\frac{1}{2}$ and an even function for spin 0.

In the standard representation one has

$$\not{P} + m = \begin{pmatrix} P^0 + m & 0 & -P_z & -P_- \\ 0 & P^0 + m & -P_+ & P_z \\ P_z & P_- & -P^0 + m & 0 \\ P_+ & -P_z & 0 & -P^0 + m \end{pmatrix}, \tag{4.9}$$

$$P_\pm := P_x \pm iP_y, \tag{4.10}$$

where P_x, P_y, P_z are the components of the 3-vector $\boldsymbol{P}$. On choosing the first two columns of (4.9) one obtains the four solutions

$$\varphi^\epsilon_+(\epsilon\boldsymbol{p}) = \frac{1}{\sqrt{2\epsilon\varepsilon(\epsilon\varepsilon + m)V}} \begin{pmatrix} \epsilon\varepsilon + m \\ 0 \\ \epsilon p_z \\ \epsilon p_+ \end{pmatrix}, \quad \varphi^\epsilon_-(\epsilon\boldsymbol{p}) = \frac{1}{\sqrt{2\epsilon\varepsilon(\epsilon\varepsilon + m)V}} \begin{pmatrix} 0 \\ \epsilon\varepsilon + m \\ \epsilon p_- \\ -\epsilon p_z \end{pmatrix}, \tag{4.11}$$

with $\epsilon = \pm 1$.

The labels + and − correspond to "spin" eigenvalues $\pm\frac{1}{2}$ and may be identified as the $s = \pm 1$ solutions. However, the "spin" operator of which (4.11) are eigenfunctions is as yet unidentified. In the case of a particle at rest, the "spin" as identified here does correspond to the spin of a nonrelativistic particle in the usual sense. However, there is a formal problem that needs to be addressed to define spin in a meaningful sense in a Lorentz covariant theory. Our understanding of spin in nonrelativistic quantum mechanics is based on a 3-tensor notation, and we need to identify the spin as a 4-vector or 4-tensor to give it meaning in a covariant theory. This turns out to be a far from trivial exercise, as described in Appendix C.

Thus the "spin" implicit in (4.11) corresponds to the eigenvalues of some operator which we have not identified. Even when the relevant operator is identified explicitly, its interpretation is of no obvious significance. Despite the spin-dependence of the solutions (4.11) being of questionable physical significance, the solutions are used widely, basically because they are algebraically relatively simple. The solutions (4.11) are useful in two cases: for nonrelativistic particles, when the spin reduces to its nonrelativistic counterpart, and in situations where the spin is of no intrinsic interest and one ultimately sums or averages over the spin states.

A general form for the spin eigenfunctions is

$$\varphi_s^\epsilon(\epsilon\boldsymbol{p}) = \frac{1}{\sqrt{2\epsilon\varepsilon(\epsilon\varepsilon+m)V}} \left[C_s^\epsilon \begin{pmatrix} \epsilon\varepsilon+m \\ 0 \\ \epsilon p_z \\ \epsilon p_+ \end{pmatrix} + D_s^\epsilon \begin{pmatrix} 0 \\ \epsilon\varepsilon+m \\ \epsilon p_- \\ -\epsilon p_z \end{pmatrix} \right], \tag{4.12}$$

$$|C_s^\epsilon|^2 + |D_s^\epsilon|^2 = 1. \tag{4.13}$$

The values of C_s^ϵ and D_s^ϵ are determined (apart from phase factors) by the specific spin operator. The wavefunctions for two well defined spin operators are written down in Appendix B.

Irrespective of the particular spin operator, the eigensolutions satisfy a set of orthogonality relations and a set of completeness relations. These express the requirements that eigensolutions corresponding to different eigenvalues are orthogonal, and that the eigensolutions span the Hilbert space respectively. Two equivalent forms of the orthogonality relations are

$$\varphi_s^{\epsilon\dagger}(\epsilon\boldsymbol{p})\varphi_{s'}^{\epsilon'}(\epsilon\boldsymbol{p}) = \frac{1}{V}\,\delta^{\epsilon\epsilon'}\,\delta_{ss'}, \tag{4.14}$$

$$\overline{\varphi}_s^{\epsilon}(\epsilon\boldsymbol{p})\varphi_{s'}^{\epsilon'}(\epsilon'\boldsymbol{p}) = \frac{m}{\epsilon\varepsilon V}\,\delta^{\epsilon\epsilon'}\,\delta_{ss'}. \tag{4.15}$$

The completeness relations are related to the identity

$$\sum_{s=\pm} \varphi_s^\epsilon(\epsilon\boldsymbol{p})\overline{\varphi}_s^\epsilon(\epsilon\boldsymbol{p}) = \frac{1}{2\varepsilon V}\,(\tilde{\not{p}} + \epsilon m), \tag{4.16}$$

with $\tilde{p} = [\varepsilon, \boldsymbol{p}]$.

3. The Spinors $u_s(\boldsymbol{p})$ and $v_s(\boldsymbol{p})$

A widely used notation corresponds to

$$\varphi_s^+(\boldsymbol{p}) = \frac{u_s(\boldsymbol{p})}{\sqrt{2\varepsilon V}}\,, \qquad \varphi_s^-(-\boldsymbol{p}) = \frac{v_s(\boldsymbol{p})}{\sqrt{2\varepsilon V}}\,. \tag{4.17}$$

In terms of these spinors, the orthogonality relation (4.14) or (4.15) imply

$$\bar{u}_s(\boldsymbol{p})u_{s'}(\boldsymbol{p}) = -\bar{v}_s(\boldsymbol{p})v_{s'}(\boldsymbol{p}) = 2m\delta_{ss'}, \tag{4.18}$$

$$u_s^\dagger(\boldsymbol{p})u_{s'}(\boldsymbol{p}) = v_s^\dagger(\boldsymbol{p})v_{s'}(\boldsymbol{p}) = 2\varepsilon\delta_{ss'}, \tag{4.19}$$

$$\bar{u}_s(\boldsymbol{p})v_{s'}(\boldsymbol{p}) = \bar{v}_s(\boldsymbol{p})u_{s'}(\boldsymbol{p}) = 0. \tag{4.20}$$

The completeness relation (4.16) implies

$$\sum_{s=\pm}[u_s(\boldsymbol{p})\bar{u}_s(\boldsymbol{p}) - v_s(\boldsymbol{p})\bar{v}_s(\boldsymbol{p})] = 2m, \tag{4.21}$$

$$\sum_{s=\pm} u_s(\boldsymbol{p})\bar{u}_s(\boldsymbol{p}) = \tilde{\not{p}} + m, \tag{4.22}$$

$$\sum_{s=\pm} v_s(\boldsymbol{p})\bar{v}_s(\boldsymbol{p}) = \tilde{\not{p}} - m, \tag{4.23}$$

with $\tilde{p} = [\varepsilon, \boldsymbol{p}]$.

4. The Zitterbewegung

Consider the velocity of an electron in Dirac's theory. The velocity operator is

$$\hat{\boldsymbol{v}} = \partial\hat{H}/\partial\hat{\boldsymbol{p}} = \boldsymbol{\alpha}, \tag{4.24}$$

where $\hat{H}$ is given by (3.12). The mean value of the velocity for an arbitrary wavefunction $\Psi(x)$ is

$$\langle\boldsymbol{v}\rangle = \int d^3\boldsymbol{x}\,\Psi^\dagger(x)\boldsymbol{\alpha}\Psi(x). \tag{4.25}$$

The wavefunction may be expanded in plane wave solutions:

$$\Psi(x) = \sum_{\epsilon s}\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\,e^{-i\epsilon px}\,a_s^\epsilon(\epsilon\boldsymbol{p})\varphi_s^\epsilon(\epsilon\boldsymbol{p}), \tag{4.26}$$

where $a_s^{\dagger\epsilon}(\epsilon\boldsymbol{p})$ are arbitrary coefficients. When (4.26) is inserted in (4.25) the integrand contains terms that have no exponential variation plus term that vary as $\exp[ipx]$ and as $\exp[-ipx]$. For a nonrelativistic particle these correspond to very rapid variations as $\exp[\pm imt]$, plus slower variations.

These rapid variations in the velocity are referred to as the *Zitterbewegung*. In practice the Zitterbewegung is not important for nonrelativistic particles because it involves fluctuations of small amplitude over distances smaller than a Compton wavelength ($\hbar/mc$). A nonrelativistic electron is not well localized on this scalelength.

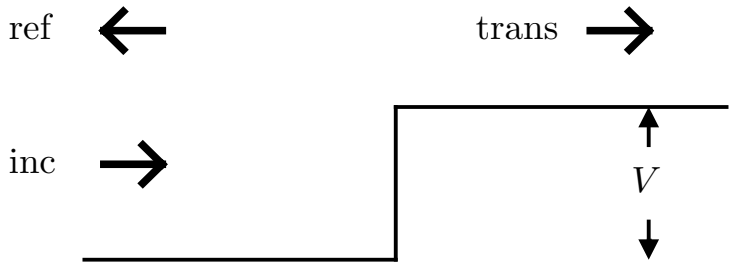

Figure 4.1 An electron is incident from the left onto a potential jump of magnitude V. The wavefunction has incident (inc), reflected (ref) and transmitted (trans) components.

6. The Klein Paradox

From an historical viewpoint, a difficulty with the Dirac equation involves the situation illustrated in Figure 4.1. From a knowledge of nonrelativistic quantum mechanics one expects there to be a probability of reflection and a probability of transmission from the potential jump, with these being given by the squares of the wavefunctions for the reflected and transmitted components. Furthermore these probabilities should sum to unity as there is no possibility other than reflection or transmission.

For simplicity let us assumed that the incident electron has spin up and that there is no spin flip. (It is left an as exercise to prove that there is no spin flip, that is, to prove that the reflected and transmitted wavefunctions are zero for spin down.) Using (4.11) in (4.1), the incident electron has the wavefunction

$$\Psi_{\rm inc}(x) = \frac{e^{-i\varepsilon t + ip_z z}}{\sqrt{2\varepsilon(\varepsilon+m)}} \begin{pmatrix} \varepsilon+m \\ 0 \\ p_z \\ 0 \end{pmatrix}, \tag{4.27}$$

where the electron is assumed to be propagating along the z-axis. The wavefunction for the reflected component, which has an oppositely directed momentum to the incident component, is

$$\Psi_{\rm ref}(x) = r\,\frac{e^{-i\varepsilon t - ip_z z}}{\sqrt{2\varepsilon(\varepsilon+m)}} \begin{pmatrix} \varepsilon+m \\ 0 \\ -p_z \\ 0 \end{pmatrix}, \tag{4.28}$$

where r is the reflection amplitude. The transmitted component has energy $\varepsilon' = \varepsilon - V$ and momentum

$$q = [\varepsilon'^2 - m^2]^{1/2}. \tag{4.29}$$

Its wavefunction is

$$\Psi_{\rm trans}(x) = \frac{\alpha\, e^{-i(\varepsilon'+V)t+iqz}}{\sqrt{2\varepsilon'(\varepsilon'+m)}} \begin{pmatrix} \varepsilon'+m \\ 0 \\ q \\ 0 \end{pmatrix}, \tag{4.30}$$

where α is the transmission amplitude.

Continuity of the wavefunctions at the potential jump at $z = 0$ leads to the condition

$$\frac{1}{\sqrt{2\varepsilon(\varepsilon+m)}} \begin{pmatrix} (\varepsilon+m)(1+r) \\ 0 \\ p_z(1-r) \\ 0 \end{pmatrix} = \frac{\alpha}{\sqrt{2\varepsilon'(\varepsilon'+m)}} \begin{pmatrix} \varepsilon'+m \\ 0 \\ q \\ 0 \end{pmatrix}, \tag{4.31}$$

On eliminating α, one solves for r and hence for the reflection coefficient

$$R = |r|^2 = \left|\frac{1-A}{1+A}\right|^2, \quad A := \frac{q}{p_z}\,\frac{\varepsilon+m}{\varepsilon-V+m}. \tag{4.32}$$

For $V < \varepsilon - m$ there is no problem with the solution. For $\varepsilon - m < V < \varepsilon + m$, q is imaginary implying $R = 1$, and this implies no transmitted component, as in the nonrelativistic case. For $V > \varepsilon + m$, q is real and R is greater than unity. This result is unacceptable in a single particle theory. This is the *Klein paradox*.

The modern view is that the Klein paradox is resolved by regarding the Dirac theory as a single charge theory rather than a single particle theory. One or more pairs can be created at the potential jump, and this allows a reflection coefficient that is greater than unity.

Lecture 5

Field Lagrangians

A Lagrangian formulation of field theory allows one to derive certain properties, such as the energy momentum tensor and the 4-current, in a systematic manner. The specific form of the Lagrangian is to be chosen such that the Euler-Lagrangian equations reproduce the field equations, that is, to the Klein Gordon equation or the Dirac equation in the present context.

N.B. The material in this lecture will not be covered in detail in lectures and will not be examinable. It is included here for reference in relation to the normalization conditions for the wave functions and the derivation of the expressions for the 4-current.

1. Euler-Lagrange Equations for a Field

The usual definition of a Lagrangian for a field, $\Psi(x)$ say, involves regarding Ψ and Ψ^* as independent "generalized coordinates", and $\partial\Psi$ and $\partial\Psi^*$ as the corresponding "generalized momenta". The Lagrangian density $\mathcal{L}(\Psi, \Psi^*, \partial\Psi, \partial\Psi^*)$ is defined such that the action I is given by

$$I = \int d^4x\, \mathcal{L}(\Psi, \Psi^*, \partial\Psi, \partial\Psi^*). \tag{5.1}$$

The Euler-Lagrange equations are then

$$\partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi^*)} - \frac{\partial\mathcal{L}}{\partial\Psi^*} = 0, \qquad \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi)} - \frac{\partial\mathcal{L}}{\partial\Psi} = 0. \tag{5.2}$$

The Lagrangian density is to be chosen so that these equations reproduce the required field equations.

The energy momentum tensor for a field is determined by a procedure that is analogous to the derivation of a Hamiltonian from a Lagrangian. Specifically, one has $H = \Sigma p\dot{q} - L$ with L replaced by $\mathcal{L}$, two qs replaced by Ψ and Ψ^*, the corresponding $\dot{q}$s replaced by $\partial_0\Psi$ and $\partial_0\Psi^*$, the corresponding ps replaced by $\partial\mathcal{L}/\partial(\partial_0\Psi)$ and $\partial\mathcal{L}/\partial(\partial_0\Psi^*)$, and H replaced by T^{00}. Then replacing the time derivatives by covariant derivatives ($\partial^0 \to \partial^\mu$ or $\partial^0 \to \partial^\nu$), this prescription suggests the following identification of the energy-momentum tensor:

$$T^{\mu\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi^*)}\,\partial^\nu\Psi^* + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi)}\,\partial^\nu\Psi - g^{\mu\nu}\mathcal{L}. \tag{5.3}$$

The energy in the field is the spatial integral of T^{00}:

$$\int d^3\boldsymbol{x}\, T^{00} = \int d^3\boldsymbol{x} \left[\frac{\partial\mathcal{L}}{\partial(\partial\Psi^*/\partial t)}\,\frac{\partial\Psi^*}{\partial t} + \frac{\partial\mathcal{L}}{\partial(\partial\Psi/\partial t)}\,\frac{\partial\Psi}{\partial t} - \mathcal{L}\right]. \tag{5.4}$$

The Lagrangian may be modified to determine the 4-current by including an electromagnetic field. This involves making the minimal coupling replacement (3.3), *viz.* $\partial^\mu \to \partial^\mu + iqA^\mu(x)$. The resulting term in the Lagrangian that is proportional to $A^\mu(x)$ is written as $J_\mu(x)A^\mu(x)$, with $J^\mu(x)$ identified as the 4-current.

2. The Klein Gordon Lagrangian

The Klein Gordon Lagrangian is to be chosen such that the Euler-Lagrange equations (5.2) reproduce the Klein Gordon equation (3.2) and its hermitian conjugate, respectively. A choice of Lagrangian that achieves this is

$$\mathcal{L}(x) = (\partial_\mu \Psi^*)(\partial^\mu \Psi) - m^2 \Psi^* \Psi. \tag{5.5}$$

Given the form of the Lagrangian (5.5), the energy momentum tensor may be constructed using (5.3). For the Klein Gordon field one finds

$$T^{\mu\nu} = (\partial^\mu \Psi^*)(\partial^\nu \Psi) + (\partial^\nu \Psi^*)(\partial^\mu \Psi) - g^{\mu\nu}\mathcal{L}, \tag{5.6}$$

which is used in (4.4). The 4-current is identified as the coefficient of the terms proportional to A in $\mathcal{L} \to [(\partial_\mu - iqA_\mu)\Psi^*][(\partial^\mu + iqA^\mu)\Psi] - m^2\Psi^*\Psi$. This prescription gives

$$J^\mu(x) = iq[\Psi^*(\partial^\mu \Psi) - (\partial^\mu \Psi^*)\Psi]. \tag{5.7}$$

The continuity relation $\partial_\mu J^\mu(x) = 0$ is implied by the Klein Gordon equation.

3. Dirac Lagrangian

For the Dirac equation it is convenient to choose the independent fields as Ψ and $\overline{\Psi}$. Then the Euler-Lagrange equations (5.4) are replaced by

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \overline{\Psi})} - \frac{\partial \mathcal{L}}{\partial \overline{\Psi}} = 0, \qquad \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi)} - \frac{\partial \mathcal{L}}{\partial \Psi} = 0. \tag{5.8}$$

A choice of Lagrangian such that (5.8) reproduces the Dirac equation (3.4) and adjoint is

$$\mathcal{L} = \tfrac{1}{2} i \left(\overline{\Psi} \gamma^\mu \partial_\mu \Psi - (\partial_\mu \overline{\Psi}) \gamma^\mu \Psi \right) - m \overline{\Psi} \Psi. \tag{5.9}$$

The energy momentum tensor for the Dirac field is

$$\begin{aligned} T^{\mu\nu} &= \frac{\partial \mathcal{L}}{\partial(\partial_\mu \overline{\Psi})}\, \partial^\nu \overline{\Psi} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi)}\, \partial^\nu \Psi - g^{\mu\nu}\mathcal{L} \\ &= -\tfrac{1}{2} i\, (\partial^\nu \overline{\Psi}) \gamma^\mu \Psi + \tfrac{1}{2} i\, \overline{\Psi} \gamma^\mu \partial^\nu \Psi, \end{aligned} \tag{5.10}$$

which is used in (4.8). The 4-current is

$$J^\mu = q \overline{\Psi}\, \gamma^\mu\, \Psi. \tag{5.11}$$

4. The Lagrangian in Momentum Space

For some formal purposes it is useful to define a Lagrangian density in momentum space. This is achieved by writing the space time average of the action integral in terms of an integral over momentum space. One has

$$\frac{I}{VT} = \frac{1}{VT}\int d^4x\,\mathcal{L}(x) = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\,\mathcal{L}(p). \tag{5.12}$$

The Lagrangian density $\mathcal{L}(x)$ is bilinear in $\Psi(x)$ and its adjoint, and the integral over it in (5.12) may be written in terms of Fourier transforms by using the power theorem (A.4). The Fourier transforms of the wavefunctions are written in terms of $\varphi(P)$ using (5.5), specifically

$$\Phi(P) = \sum_{\epsilon=\pm} \Phi^{\epsilon}(\epsilon\boldsymbol{p})\,(2\pi)^4\,\delta^4(P-\epsilon p), \tag{5.13}$$

and the square of the δ function is rewritten using (A.9). This allows one to identify $\mathcal{L}(P)$.

For the Klein Gordon field the expression (5.5) for $\mathcal{L}(x)$ implies

$$\mathcal{L}(P) = \sum_{\epsilon=\pm}(p^2-m^2)\varphi^{\epsilon *}(\epsilon\boldsymbol{p})\varphi^{\epsilon}(\epsilon\boldsymbol{p})\,(2\pi)^3\,\delta^4(\boldsymbol{P}-\epsilon\boldsymbol{p}), \tag{5.14}$$

with $P^0 = \epsilon\varepsilon$. The independent variables may be identified as $\varphi^{\epsilon *}(\boldsymbol{p})$, $\varphi^{\epsilon}(\boldsymbol{p})$ and the phase. Variation of $\mathcal{L}(p)$ with respect to φ^* and φ leads to the wave equation in momentum space $(p^2-m^2)\varphi^{\epsilon}(\boldsymbol{p}) = 0$ and its adjoint respectively.

For the Dirac field the expression (5.9) for $\mathcal{L}(x)$ implies

$$\mathcal{L}(P) = \sum_{\epsilon=\pm}\overline{\varphi}^{\epsilon}(\boldsymbol{p})(\not{P}-m)\varphi^{\epsilon}(\boldsymbol{p})\,(2\pi)^3\,\delta^3(\boldsymbol{P}-\epsilon\boldsymbol{p}), \tag{5.15}$$

with $P^0 = \epsilon\varepsilon$. The independent variables may be identified as $\overline{\varphi}^{\epsilon}(\boldsymbol{p})$, $\varphi^{\epsilon}(\boldsymbol{p})$ and the phase. Variation of $\mathcal{L}(P)$ with respect to φ^* and φ leads to the wave equation in momentum space $(\not{p}-m)\varphi^{\epsilon}(\boldsymbol{p}) = 0$ and its adjoint respectively.

5. Particle Action and Occupation Numbers

The occupation numbers for particles may be derived from the Lagrangian in momentum space. The action is the variable conjugate to the phase, and on dividing the action by $\hbar$ (with $\hbar = 1$ here), the action is reinterpreted in terms of the occupation number, at least to within a sign. The essential idea is to regard $\mathcal{L}(p)$ as a function of the phase ψ of the field, with

$$p^{\mu} = \partial^{\mu}\psi. \tag{5.16}$$

The Lagrangian depends on $\partial\psi$ but not on ψ itself, implying the conservation law

$$\partial_{\mu}\left(\frac{\partial\mathcal{L}(\partial\psi)}{\partial(\partial_{\mu}\psi)}\right) = 0. \tag{5.17}$$

The conserved quantity is the action

$$\tilde{n}(\boldsymbol{P}) = \frac{\partial \mathcal{L}(P)}{\partial p^0}, \tag{5.18}$$

which, apart from a possible sign for antiparticles, is also the occupation number.

Let the occupation number be written as $n^\epsilon(\boldsymbol{p})$, with $\epsilon = 1$ for particles and with $\epsilon = -1$ for antiparticles. It is desirable to define the occupation numbers to be non-negative. Inspection of (5.14) and (5.15) shows that the particle action (5.18) is an odd function of P^0 for spin 0 and an even function of P^0 for spin $\frac{1}{2}$. Consequently for $P = \epsilon p$ the particle action is related the occupation numbers are identified as

$$n^\epsilon(\boldsymbol{p}) = \begin{cases} \epsilon\tilde{n}(\epsilon\boldsymbol{p}) & \text{for spin } 0, \\ \tilde{n}(\epsilon\boldsymbol{p}) & \text{for spin } \frac{1}{2}. \end{cases} \tag{5.19}$$

In this way the occupation numbers are identified as

$$n^\epsilon(\boldsymbol{p}) = \varphi^{\epsilon *}(\boldsymbol{P})\varphi^\epsilon(\boldsymbol{P})\,(2\pi)^3\,\delta^3(\boldsymbol{P} - \epsilon\boldsymbol{p}), \tag{5.20}$$

for spin 0, and as

$$n_s^\epsilon(\boldsymbol{p}) = \varphi_s^{\epsilon\dagger}(\boldsymbol{P})\varphi_s^\epsilon(\boldsymbol{P})\,(2\pi)^3\,\delta^3(\boldsymbol{P} - \epsilon\boldsymbol{p}), \tag{5.21}$$

for spin $\frac{1}{2}$.

The normalization condition to one particle in the normalization volume may be imposed trivially in terms of the occupation numbers. One requires

$$V\sum_s \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\, n_s^\epsilon(\boldsymbol{p}) = 1, \tag{5.22}$$

where the sum over s not relevant only for spin 0. On inserting the expressions (5.20) and (5.21) in (5.22) one finds

$$\varphi_s^{\epsilon *}(\epsilon\boldsymbol{p})\varphi_s^\epsilon(\epsilon\boldsymbol{p}) = \frac{1}{2\varepsilon V}. \tag{5.23}$$

for spin 0, and

$$\varphi_s^{\epsilon\dagger}(\epsilon\boldsymbol{p})\varphi_s^\epsilon(\epsilon\boldsymbol{p}) = \frac{1}{V}, \tag{5.24}$$

for spin $\frac{1}{2}$, respectively. The results (5.23) and (5.24) reproduce (4.5) and (4.14), respectively.

Lecture 6

Motion in a Magnetostatic Field

One can solve Dirac's equation exactly in the presence of a magnetostatic field $\boldsymbol{B}$. The form of the solution depends on a choice of gauge for the vector potential $\boldsymbol{A}(\boldsymbol{x})$ describing the field.

1. Dirac's Equation in the Landau Gauge

When an electromagnetic field with 4-potential A^μ is included, Dirac's equation is modified by the minimal coupling procedure of replacing $\hat{p}^\mu$ by $\hat{p}^\mu - q\hat{A}^\mu$, where q is the charge on the particle. One has $q = -e$ for an electron. The Dirac Hamiltonian is then replaced by

$$\hat{H} = \boldsymbol{\alpha} \cdot (\boldsymbol{p} + e\boldsymbol{A}) + \beta m - e\phi. \tag{6.1}$$

Here we set $\phi = 0$.

Let the magnetostatic field be along the z-axis. One choice of gauge is

$$\boldsymbol{A} = (0, Bx, 0), \tag{6.2}$$

which is called the *Landau gauge*. Other choices of gauge are related to this by adding the gradient of a scalar to $\boldsymbol{A}$. Other choices include

$$\boldsymbol{A} = (-By, 0, 0), \tag{6.3}$$

$$\boldsymbol{A} = \tfrac{1}{2}(-By, Bx, 0) = \tfrac{1}{2}B\varpi(-\sin\varphi, \cos\varphi, 0), \tag{6.4}$$

with $\varpi = (x^2 + y^2)^{1/2}$ and $x = \varpi\cos\varphi$, $y = \varpi\sin\varphi$.

An important point to note is that with the choice (6.2), on writing $\hat{p} = -i\,\mathrm{grad}$, Dirac's equation depends on only one coordinate, namely x. The other coordinates y and z are *ignorable* or *cyclic*. The corresponding momenta p_y and p_z are then constants of the motion, and may be chosen as (continuous) quantum numbers. Also the Hamiltonian does not depend on time and so the energy is also a constant of the motion. Dirac's equation may then be reduced to an ordinary differential equation in x. Similarly with the other two choices of gauge (6.3) and (6.4) Dirac's equation may be reduced to an ordinary differential equation, in y and ϖ respectively. The latter case is treated in Appendix B.

2. Construction of the Wave Functions

Let us assume a trial wave function of the form

$$\Psi(t, \boldsymbol{x}) = f(x)\,\exp(-i\epsilon\varepsilon t + i\epsilon p_y y + i\epsilon p_z z), \tag{6.5}$$

where $\epsilon = \pm$ is the sign of the energy whose magnitude is ε. The wave function Ψ is a column matrix, and so is $f(x)$. The components of the column matrix are denoted by $f_1 \ldots f_4$.

On inserting the trial solution (6.5) into Dirac's equation in the form

$$(i\frac{\partial}{\partial t} - \hat{H})\Psi(t, \boldsymbol{x}) = 0, \tag{6.6}$$

one requires

$$\begin{pmatrix} -\epsilon\varepsilon + m & 0 & \epsilon p_z & \hat{\mathcal{O}}_1 \\ 0 & -\epsilon\varepsilon + m & \hat{\mathcal{O}}_2 & -\epsilon p_z \\ \epsilon p_z & \hat{\mathcal{O}}_1 & -\epsilon\varepsilon - m & 0 \\ \hat{\mathcal{O}}_2 & -\epsilon p_z & 0 & -\epsilon\varepsilon - m \end{pmatrix} \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \\ f_4(x) \end{pmatrix} = 0, \tag{6.7}$$

with

$$\hat{\mathcal{O}}_1 := -i\left(\frac{\partial}{\partial x} + \epsilon p_y + eBx\right), \qquad \hat{\mathcal{O}}_2 := -i\left(\frac{\partial}{\partial x} - \epsilon p_y - eBx\right). \tag{6.8}$$

It is convenient to write

$$\xi := (eB)^{1/2}\left(x + \frac{\epsilon p_y}{eB}\right), \tag{6.9}$$

and then (6.7) reduces to

$$\begin{aligned} (-\epsilon\varepsilon + m)f_1 + \epsilon p_z f_3 - i(eB)^{1/2}(\xi + d/d\xi)f_4 = 0, \\ (-\epsilon\varepsilon + m)f_2 - \epsilon p_z f_4 + i(eB)^{1/2}(\xi - d/d\xi)f_3 = 0, \\ (-\epsilon\varepsilon - m)f_3 + \epsilon p_z f_1 - i(eB)^{1/2}(\xi + d/d\xi)f_2 = 0, \\ (-\epsilon\varepsilon - m)f_4 - \epsilon p_z f_2 + i(eB)^{1/2}(\xi - d/d\xi)f_1 = 0. \end{aligned} \tag{6.10}$$

Operating on the first of these equations with $(\xi - d/d\xi)$ and on the second with $(\xi + d/d\xi)$ gives

$$\begin{aligned} \left[\frac{d^2}{d\xi^2} + \frac{\varepsilon^2 - m^2 - p_z^2}{eB} - (\xi^2 + 1)\right] f_{1,3} = 0, \\ \left[\frac{d^2}{d\xi^2} + \frac{\varepsilon^2 - m^2 - p_z^2}{eB} - (\xi^2 - 1)\right] f_{2,4} = 0. \end{aligned} \tag{6.11}$$

Equations (6.11) are of the same form as Schrödinger's equation for a simple harmonic oscillator, the solution of which is given in any elementary textbook on quantum mechanics. The differential equations (6.11) have normalizable solutions only if the constant

$$n := \frac{\varepsilon^2 - m^2 - p_z^2}{eB} \tag{6.12}$$

in (6.11) has integral values. Specifically, let $\ell = 0, 1, 2, \ldots$ be the integer; then there are normalizable solutions only for

$$2n \mp 1 = 2\ell + 1. \tag{6.13}$$

The normalized solutions are the harmonic oscillator wave functions

$$v_\ell(\xi) = \frac{1}{(\sqrt{\pi}2^\ell l!)^{1/2}}\, H_\ell(\xi)\, e^{-\xi^2/2}, \tag{6.14}$$

where H_ℓ is a hermite polynomial. The solutions give

$$f(x) = \begin{pmatrix} C_1 v_{n-1}(\xi) \\ C_2 v_n(\xi) \\ C_3 v_{n-1}(\xi) \\ C_4 v_n(\xi) \end{pmatrix}, \tag{6.15}$$

where $C_1, \ldots, C_4$ are constants.

3. Spin Eigenfunctions

The constants $C_1, \ldots, C_4$ in (6.15) are to be determined by constructing eigenfunctions of specific spin operators, and normalizing the solutions. The procedure is closely analogous to that for the unmagnetized case. This follows from the properties

$$\begin{aligned} \left(\xi + \frac{d}{d\xi}\right) v_n(\xi) &= \sqrt{2n}\; v_{n-1}(\xi), \\ \left(\xi - \frac{d}{d\xi}\right) v_n(\xi) &= \sqrt{2(n+1)}\; v_{n+1}(\xi), \end{aligned} \tag{6.16}$$

of the harmonic oscillator wave functions implying that the wave functions may be eliminated when (6.15) is inserted in (6.11). Thus one requires

$$\begin{pmatrix} -\epsilon\varepsilon_q + m & 0 & \epsilon p_z & -ip_n \\ 0 & -\epsilon\varepsilon_q + m & ip_n & -\epsilon p_z \\ \epsilon p_z & -ip_n & -\epsilon\varepsilon_q - m & 0 \\ ip_n & -\epsilon p_z & 0 & -\epsilon\varepsilon_q - m \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} = 0, \tag{6.17}$$

with

$$p_n := (2neB)^{1/2} \qquad \varepsilon_q := (m^2 + p_z^2 + 2neB)^{1/2}, \tag{6.18}$$

where q denotes the quantum numbers collectively.

Now one may proceed as in the derivation of the solution (4.18) in the unmagnetized case. One finds four different solutions from the first two columns for $\epsilon = \pm$:

$$\begin{aligned} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} &= \frac{1}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} \epsilon\varepsilon_q + m \\ 0 \\ \epsilon p_z \\ ip_n \end{pmatrix}, \\ \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} &= \frac{1}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} 0 \\ \epsilon\varepsilon_q + m \\ -ip_n \\ -\epsilon p_z \end{pmatrix}, \end{aligned} \tag{6.19}$$

where the normalization is to one particle in the volume V. Let these be labeled as the $s = +1$ and $s = -1$ solutions respectively. Our four solutions (denoted by $\Psi^{\epsilon}_{q}(t, \boldsymbol{x})$ with q denoting p_z, n, s) are then

$$\Psi^{\epsilon}_{q+}(t, \boldsymbol{x}) = \frac{\exp[-i\epsilon\varepsilon t + i\epsilon p_y y + i\epsilon p_z z]}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} (\epsilon\varepsilon_q + m)v_{n-1}(\xi) \\ 0 \\ \epsilon p_z v_{n-1}(\xi) \\ ip_n v_n(\xi) \end{pmatrix},$$
$$\Psi^{\epsilon}_{q-}(t, \boldsymbol{x}) = \frac{\exp[-i\epsilon\varepsilon t + i\epsilon p_y y + i\epsilon p_z z]}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} 0 \\ (\epsilon\varepsilon_q + m)v_n(\xi) \\ -ip_n v_{n-1}(\xi) \\ -\epsilon p_z v_n(\xi) \end{pmatrix}, \tag{6.20}$$

where $q+$ and $q-$ denote the quantum numbers with $s = +1$ and $s = -1$ respectively.

4. The Normalization and Orthogonality Relations

The orthogonality relation between the states may be written in the general form

$$\int d^3\boldsymbol{x}\, [\Psi^{\epsilon}_{q}(t, \boldsymbol{x})]^{\dagger} \Psi^{\epsilon'}_{q'}(t, \boldsymbol{x}) = \delta^{\epsilon\epsilon'}\, \delta_{qq'}, \tag{6.21}$$

where q and q' include s and s'. The completeness relation for the states is

$$\sum_{q} \Psi_q(t, \boldsymbol{x})[\Psi_q(t, \boldsymbol{x}')]^{\dagger} = \delta^3(\boldsymbol{x} - \boldsymbol{x}'). \tag{6.22}$$

The explicit forms of these relations in the present case involves the interpretation of the sum and the δ-function involving the state q and q'.

The normalization condition may be understood by considering the case where the particle is confined to a large but finite box. Let the sides of the box be of length L_x, L_y, L_z in the x, y, z directions respectively. The eigenvalues p_y and p_z are then discrete rather than continuous, specifically one has $p_y = n_y 2\pi/L_y$ and $p_z = n_z 2\pi/L_z$ with $n_y, n_z = 0, \pm 1, \pm 2, \cdots$. The sum over states is then trivial, involving sums over these discrete eigenvalues (as well as over n and s). The sums may be approximated by integrals in the limit where the sides of the box becomes arbitrarily large. Then one has

$$\sum_{n_y, n_z = -\infty}^{\infty} \to L_y L_z \int \frac{dp_y}{2\pi} \int \frac{dp_z}{2\pi}.$$

Thus the sum over states becomes

$$\sum_{q} = \sum_{s=\pm} \sum_{n=0}^{\infty} L_y L_z \int \frac{dp_y}{2\pi} \int \frac{dp_z}{2\pi}. \tag{6.23}$$

Similarly the δ function expressing orthogonality of the states becomes

$$\delta_{qq'} = \delta_{ss'}\delta_{nn'} \frac{2\pi}{L_y}\, \delta(p_y - p'_y)\, \frac{2\pi}{L_z}\, \delta(p_z - p'_z). \tag{6.24}$$

5. The Center of Gyration

The variable p_y is related to the center of gyration of the particle through (6.9). The limits of integration of p_y are such that p_y is less than eBL_x, where L_x is the range of x. In many cases one is not interested in the spatial distributions of the centers of gyration, which are assumed homogeneous. One then wishes to perform the sum or integral explicitly. In such cases, (6.23) may be further replaced by

$$\sum_q = \sum_{s=\pm} \sum_{n=0}^{\infty} \frac{eBV}{2\pi} \int \frac{dp_z}{2\pi}, \tag{6.25}$$

with $V = L_x L_y L_z$.

Note that the range L_x of x can be fixed in a natural way (although one is not compelled to do so). This follows from the fact that eB has the dimensions (in natural units) of an inverse area, and hence $(eB)^{-1/2}$ is a natural length. Thus the natural normalization corresponds to

$$L_x = \frac{1}{(eB)^{1/2}}, \quad \text{or} \quad V = \frac{L_y L_z}{(eB)^{1/2}}. \tag{6.26}$$

The description of the center of gyration is gauge-dependent. If one chooses the gauge (6.3) in place of (6.2), then an analogous discussion to the one given here involves a quantum number p_x which is related to the y-coordinate of the center of gyration. In terms of the "cylindrical" gauge (6.4) there is a "radial" quantum number related to the position of the center of gyration.

6. The Magnetic Moment of the Electron

It is apparent from (6.13) and the solutions (6.20) that one has

$$n = \ell + \tfrac{1}{2}(1 + s). \tag{6.27}$$

The energy involves n in the form

$$\varepsilon_q = [m^2 + p_z^2 + 2neB]^{1/2}, \tag{6.28}$$

and separation (6.27) may be regarded as a separation into an orbital part described by ℓ and a spin part. The orbital part describes the perpendicular motion, which is simple harmonic motion. (Circular motion is simple harmonic motion.) The energy states of the simple harmonic oscillator correspond to $(\ell + \frac{1}{2})$ times $\hbar\Omega_0$, where Ω_0 ($\Omega_0 = eB/m$ here) is the frequency of the oscillator. The remaining part corresponds to an energy $\frac{1}{2}s\hbar\Omega_0$ which may be interpreted as a contribution $\boldsymbol{\mu} \cdot \boldsymbol{B}$, where $\boldsymbol{\mu}$ is the magnetic moment. Dirac's theory thus predicts a magnetic moment

$$\mu_B = \frac{\hbar e}{2m}. \tag{6.29}$$

In the Pauli-Schrödinger theory the gyromagnetic ratio of the electron, that is the ratio of the magnetic moment to the spin, has to be specified. Dirac's theory correctly predicts that the gyromagnetic ratio of the electron has the correct value of two, and this was one of the initial major successes of Dirac's theory. (Radiative corrections in QED correctly predict small difference from two.)

7. Magnetic Moment Tensor

A covariant version of the foregoing discussion enables one to identify a magnetic moment 4-tensor for the Dirac theory. This derivation proceeds as follows.

The minimal coupling replacement for including the effect of an electromagnetic field in *any* covariant wave equation is equivalent to the replacement

$$\partial^\mu \to D^\mu := \partial^\mu + iqA^\mu, \tag{6.30}$$

with $q = -e$ for the Dirac theory. Then Dirac's equation becomes

$$(i\not{\partial} + e\not{A} - m)\Psi = 0. \tag{6.31}$$

On introducing the ansatz

$$\Psi = (i\not{\partial} + e\not{A} + m)\chi, \tag{6.32}$$

Dirac's equation becomes

$$(i\not{D} - m)(i\not{D} + m)\chi = 0. \tag{6.33}$$

Then using

$$\gamma^\mu\gamma^\nu D_\mu D_\nu = D^\mu D_\mu + \frac{1}{4}[\gamma^\mu, \gamma^\nu]\,[D_\mu, D_\nu]$$

and

$$[D_\mu, D_\nu] = -ieF_{\mu\nu}, \qquad S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu],$$

one finds

$$\left(D^\mu D_\mu + m^2 - eS^{\mu\nu}F_{\mu\nu}\right)\chi = 0. \tag{6.34}$$

The term $-eS^{\mu\nu}F_{\mu\nu}$ in (6.34) contains the magnetic moment interaction term in covariant form. One has

$$S^{\mu\nu} = \frac{1}{2}\begin{pmatrix} 0 & i\alpha_x & i\alpha_y & i\alpha_z \\ -i\alpha_x & 0 & \sigma_z & -\sigma_y \\ -i\alpha_y & -\sigma_z & 0 & \sigma_x \\ -i\alpha_z & \sigma_y & -\sigma_x & 0 \end{pmatrix}, \tag{6.35}$$

and hence

$$S^{\mu\nu}F_{\mu\nu} = i\boldsymbol{\alpha}\cdot\boldsymbol{E} - \boldsymbol{\sigma}\cdot\boldsymbol{B}. \tag{6.36}$$

The $\boldsymbol{\sigma}\cdot\boldsymbol{B}$ term in (6.34) with (6.36) reproduces Pauli term (1.20) which is introduced into the Schrödinger theory in an *ad hoc* manner.

The actual definition of the magnetic moment operator involves further manipulations of the spin 4-tensor $S^{\mu\nu}$, as outlined in Appendix C.

Lecture 7

Dirac Theory for Hydrogen-like Atoms

Dirac's equation may be solved exactly in the case of a Coulomb field, leading to a relativistic theory for hydrogen-like atoms. The most important result is the so-called fine structure formula for the energy levels.

1. Dirac and Klein-Gordon Equations for a Coulomb Field

The Dirac Hamiltonian (6.1) simplifies to (in ordinary units)

$$\hat{H} = c\boldsymbol{\alpha} \cdot \boldsymbol{p} + \beta m_e c^2 - e\phi \tag{7.1}$$

for an electrostatic field in the Coulomb gauge. The potential is identified here as

$$\phi(r) = \frac{Ze}{4\pi\varepsilon_0 r}. \tag{7.2}$$

The case of the hydrogen atom corresponds to $Z = 1$, and it is of interest to consider $Z > 1$ corresponding to hydrogen-like atoms, e.g., for $Z = 26$ one has an Fe-atom stripped of all but one of its electrons.

The energy eigenvalues of the Klein-Gordon equation for a hydrogen-like atom are much simpler to find than are those of the Dirac equation. The eigenvalues of the Klein-Gordon equation may be used as an intermediate step in finding the eigenvalues of the Dirac equation. The Klein-Gordon equation is

$$\left[\left(\frac{E}{c} + \frac{Z\alpha\hbar}{r}\right)^2 + \hbar^2\nabla^2 - m_e^2 c^2\right]\psi(\boldsymbol{x}) = 0, \tag{7.3}$$

where the fine structure constant is $\alpha = r_0/r_c$, where $r_0 = e^2/4\pi\varepsilon_0 m_e c^2$ is the classical radius of the electron, and $r_c = \hbar/m_e c$ is its Compton wave length.

2. The Hydrogen-like Klein Gordon Atom

One may write doen the eigenvalues of the Klein Gordon equation (7.3) by noting that it can be forced into the same form as the Schrödinger equation, for which the solution is well known.

Recall from nonrelativistic quantum mechanics that in a central potential one writes

$$\hat{\boldsymbol{p}}^2 = -\hbar^2\nabla^2 = -\hbar^2\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right) + \frac{\hat{\boldsymbol{L}}^2}{r^2}, \tag{7.4}$$

so that Schrödinger's equation reduces to

$$\left[\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} - \frac{\hat{\boldsymbol{L}}^2}{r^2} + \frac{2m_e c\hbar Z\alpha}{r}\right]\psi(\boldsymbol{x}) = 2m_e E_{\mathrm{NR}}\psi(\boldsymbol{x}), \tag{7.5}$$

and $\hat{\boldsymbol{L}}^2$ has eigenvalues $\hbar^2 l(l+1)$. For a nonrelativistic hydrogen-like atom one has

$$(E_{\rm NR})_n = -\frac{Z^2\alpha^2 m_e c^2}{2n^2}, \qquad n-l = 1, 2, \ldots. \tag{7.6}$$

Now consider the Klein Gordon equation (7.3) for the hydrogen-like atom. Using (7.4), one has

$$\left[\left(\frac{E}{c} + \frac{Z\alpha\hbar}{r}\right)^2 + \frac{\hbar^2}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} - \frac{\hat{\boldsymbol{L}}^2}{r^2} - m_e^2c^2\right]\psi(\boldsymbol{x}) = 0. \tag{7.7}$$

One may rewrite (7.7) in the form

$$\left\{\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} - \frac{\hbar^2[l(l+1)-(Z\alpha)^2]}{r^2} + \frac{2EZ\hbar\alpha}{rc}\right\}\psi(\boldsymbol{x}) = \frac{E^2-m_e^2c^4}{c^2}\,\psi(\boldsymbol{x}). \tag{7.8}$$

This is is equivalent to the Schrödinger equation (7.5) if one makes the replacements

$$2m_ec^2E_{\rm NR} \to (E^2-m_e^2c^4), \quad m_ec^2\alpha \to E\alpha, \quad n \to n' = n-(l+\tfrac{1}{2})+[(l+\tfrac{1}{2})^2-(Z\alpha)^2]^{1/2}.$$

Then $E^2 - m_e^2c^4 = -E^2(Z\alpha)^2/n'^2$ implies

$$E_{nl} = m_ec^2\left(1+\frac{Z^2\alpha^2}{n'^2}\right)^{-1/2}, \qquad n' = n-(l+\tfrac{1}{2})+[(l+\tfrac{1}{2})^2-(Z\alpha)^2]^{1/2}. \tag{7.9}$$

with $n'-l' = 1,\, 2,\, \ldots,\, l'(l'+1) = l(l+1)-(Z\alpha)^2$.

According to (6.31)–(6.34), Dirac's equation, in natural units, in the presence of an electromagnetic field, $(i\partial\!\!\!/ + eA\!\!\!/ - m)\Psi = 0$, can be reduced to a Klein-Gordon like equation by introducing the ansatz $\Psi = (i\partial\!\!\!/ + eA\!\!\!/ + m)\chi$. Writing $D^\mu := \partial^\mu + iqA^\mu$ one finds

$$\left(D^\mu D_\mu + m^2 - eS^{\mu\nu}F_{\mu\nu}\right)\chi = 0. \tag{7.10}$$

One may apply (7.10) to discuss Dirac's equation for the hydrogen atom by identifying $F_{\mu\nu}$ as the Maxwell tensor for the electric field $\boldsymbol{E} = -\mathrm{grad}\,\phi$. From (6.34) we then have $S^{\mu\nu}F_{\mu\nu} = i\boldsymbol{\alpha}\cdot\boldsymbol{E}$. Now $\boldsymbol{\alpha}$ involves only the Pauli matrices, cf. (3.14), so that its eigenvalues must have the same values as those of the Pauli matrices. Moreover, the final term in (7.10) is proportional to r^{-2}, which is of the same form at the term $[l(l+1)-(Z\alpha)^2]/r^2$ in (7.8). In the presence of both orbital and spin angular momentum, only the total angular momentum is conserved, as discussed further below. The eigenvalues of the total angular momentum are $j = l \pm \frac{1}{2}$. It turns out that the only change that occurs in going from the eigenvalues of the Klein-Gordon equation to the Dirac equation involve the replacement of l in (7.9) by $j+\frac{1}{2}$.

The resulting expression for the energy levels of the hydrogen-like atom is called the *fine structure formula*. A standard form of it is

$$E_{nj} = m\left(1+\frac{Z^2\alpha^2}{n'^2}\right)^{-1/2}, \tag{7.11}$$

with the principal quantum number n related to n' by

$$n' = n - (j + \tfrac{1}{2}) + [(j + \tfrac{1}{2})^2 - (Z\alpha)^2]^{1/2}. \tag{7.12}$$

One expanding in powers of $Z^2\alpha^2$, (7.11) with (7.12) gives

$$\frac{E_{nj} - m_e c^2}{m_e c^2} = -\frac{Z^2\alpha^2}{2n^2}\left[1 + \frac{Z^2\alpha^2}{n}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right) + \cdots\right]. \tag{7.13}$$

The leading term in (7.13) is just the nonrelativistic formula for the energy of the hydrogen atom. This term is independent of j so that all energy levels with a given n and different j-values are degenerate. The allowed values correspond to $l = 0, 1, \ldots, n-1$ for each n, with $j = l \pm \frac{1}{2}$ except for $l = 0$ when only $j = \frac{1}{2}$ is possible. The next terms in (7.13) break the degeneracy so that there is a splitting between energy eigenvalues corresponding to different values of j. This splitting is the fine structure.

3. The Angular Momentum Eigenstates

Direct solution of the Dirac equation is required to construct the wave funtions. The following discussion is an outline of some of the steps involved, concentrating on how the direct approach leads to the fine structure formula (7.11).

Using the standard representation of the Dirac matrices, it is helpful to introduce 2-spinor wave functions φ and χ. These have the same character as a Schrödinger wave function for a spin-$\frac{1}{2}$ particle in the sense that they may be written as column matrices with two elements, and interpreted (somewhat loosely) as corresponding to spin up and spin down states. The Dirac wave function, which is a 4-spinor, is then written in the form

$$\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}. \tag{7.14}$$

Dirac's equation with the Hamiltonian (6.1) and the form (7.14) separates into two coupled equations that may be written in the form

$$(E - m_e c^2 + e\phi)\varphi - c\boldsymbol{\sigma}\cdot\boldsymbol{p}\chi = 0, \tag{7.15}$$

$$(E + m_e c^2 + e\phi)\chi - c\boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi = 0. \tag{7.16}$$

In the nonrelativistic limit one has $E \approx m_e c^2$ and (7.16) may be approximated by $2m_e c^2\chi - c\boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi = 0$; on solving for χ and substituting this into (7.15) one obtains the Schrödinger equation. The nonrelativistic limit is discussed explictly in the next lecture. Here we are concerned with finding the eigenvalues E implied by (7.15) and (7.16) in the case where no nonrelativistic assumption is made.

It is obvious on physical grounds that the angular momentum must be conserved in any central potential (because there can be no torque to change the angular momentum) and hence the total angular momentum $\boldsymbol{J}$ is a constant of the motion. Parity is also a constant of the motion for a spherically symmetric system. We look for energy eigenvalues,

that is, eigenvalues of the Hamiltonian $\hat{H}$, that are also simultaneous eigenvalues of $\hat{J}_z$, $\hat{J}^2$ and of parity. Let the eigenvalues of $\hat{J}_z$ and $\hat{J}^2$ be $\hbar m$ and $\hbar^2 j(j+1)$, respectively. (Note that I denote the mass of the electron by m_e in this lecture to avoid confusion with the eigenvalue m of $\hat{J}_z$). The total angular momentum $\boldsymbol{J}$ is the sum of the orbital angular momentum $\boldsymbol{L}$ and the spin angular momentum $\boldsymbol{S}$:

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}. \tag{7.17}$$

From the theory of the addition of angular momenta, we know that the allowed values of the orbital angular momentum l (so that the eigenvalues of $\boldsymbol{L}^2$ are $\hbar^2 l(l+1)$) are $l = j \pm \frac{1}{2}$, where we use the fact that $\boldsymbol{S} = \frac{1}{2}\hbar\boldsymbol{\sigma}$ has half-integral eigenvalues. Hence any relevant simultaneous eigenvalues of $\hat{J}_z$, and $\hat{J}^2$ involve the spherical harmonic $Y_{l'}^{m'}(\theta,\phi)$, with $m' = m \pm \frac{1}{2}$ and $l' = j \pm \frac{1}{2}$. With $j = l \pm \frac{1}{2}$ defining the sign $\pm$, let us define the 2-spinors

$$\phi_{j,m}^{(\pm)} = \frac{1}{\sqrt{2l+1}} \begin{pmatrix} \pm\sqrt{l \pm m + \frac{1}{2}}\, Y_l^{m-\frac{1}{2}} \\ \sqrt{l \mp m + \frac{1}{2}}\, Y_l^{m+\frac{1}{2}} \end{pmatrix}. \tag{7.18}$$

By construction these are eigenstates of $\hat{J}_z$, $\hat{J}^2$ with eigenvalues $m\hbar$, $j(j+1)\hbar^2$, respectively. They are also eigenstates of the parity operator with eigenvlues $(-1)^l$.

We now argue that the 2-spinors in (7.15) and (7.16) can be expressed in terms of the 2-spinor eigenstates (6.7). The argument involves several steps. First, let us write

$$\boldsymbol{J}^2 = \boldsymbol{L}^2 + \hbar\boldsymbol{L}\cdot\boldsymbol{\sigma} + \tfrac{3}{4}\hbar^2, \tag{7.19}$$

and assume that the operators act on $\phi_{j,m}^{(\pm)}$. Then we may solve

$$j(j+1)\hbar^2 = (j \pm \tfrac{1}{2})(j \pm \tfrac{1}{2} + 1)\hbar^2 + \hbar\boldsymbol{L}\cdot\boldsymbol{\sigma} + \tfrac{3}{4}\hbar^2, \tag{7.20}$$

to find that the allowed eigenvalues of $\boldsymbol{L}\cdot\boldsymbol{\sigma}$. The allowed values are $(j-\frac{1}{2})\hbar$ and $(-j-\frac{3}{2})\hbar$. Second, the identification of the eigenvalues of $\boldsymbol{L}\cdot\boldsymbol{\sigma}$ allows us to determine the eigenvalues of $\boldsymbol{\sigma}\cdot\boldsymbol{p}$ in (7.15) and (7.16): one uses the identity

$$\boldsymbol{\sigma}\cdot\boldsymbol{a}\,\boldsymbol{\sigma}\cdot\boldsymbol{b} = \boldsymbol{a}\cdot\boldsymbol{b} + i\boldsymbol{\sigma}\cdot(\boldsymbol{a}\times\boldsymbol{b}) \tag{7.21}$$

first with $\boldsymbol{a} = \boldsymbol{b} = \boldsymbol{x}$, and then with $\boldsymbol{a} = \boldsymbol{x}$ and $\boldsymbol{b} = \boldsymbol{p}$ and $\boldsymbol{L} = \boldsymbol{x}\times\boldsymbol{p}$ to show

$$\boldsymbol{\sigma}\cdot\boldsymbol{p} = \frac{1}{r^2}(\boldsymbol{\sigma}\cdot\boldsymbol{x})(\boldsymbol{\sigma}\cdot\boldsymbol{x})(\boldsymbol{\sigma}\cdot\boldsymbol{p}) = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{x}}{r^2}(\boldsymbol{x}\cdot\boldsymbol{p} + i\boldsymbol{\sigma}\cdot\boldsymbol{L}). \tag{7.22}$$

Thus the effect of the operator $\boldsymbol{\sigma}\cdot\boldsymbol{p}$, which appears in (7.15) and (7.16), on the eigenstates (7.18) may be determined in terms of the effects of the operators $\boldsymbol{\sigma}\cdot\boldsymbol{L}$, which is already determined, and of the two other operators that appear on the right hand side of (7.18). One of these is, in the coordinate representation,

$$\boldsymbol{x}\cdot\boldsymbol{p} = -i\hbar r\frac{\partial}{\partial r}. \tag{7.23}$$

Third, the effect of the remaining operator in (7.22) turns out to be simple for our choice of eigenstates. Specifically, one finds that the spinors (7.7) satisfy the identity

$$\boldsymbol{\sigma} \cdot \boldsymbol{x}\, \phi^{(\pm)}_{j,m} = -r\phi^{(\mp)}_{j,m}. \tag{7.24}$$

The proof of the identity (7.24) is given as an exercise.

The final step in the argument is to consider the effect of a parity transformation. Under $\boldsymbol{x} \to -\boldsymbol{x}$ the momentum changes sign, that is, $\boldsymbol{p} \to -\boldsymbol{p}$, and the other quantities (apart from ϕ and χ) in (7.15) and (7.16) are unchanged. It follows that if ϕ and χ are eigenstates of partiy then their eigenvalues have opposite signs. The only possibilities are that ϕ is proportional to $\phi^{(+)}_{j,m}$ and χ is proportional to $\phi^{(-)}_{j,m}$, or vice versa. The two choices are of the form

$$\psi = \begin{pmatrix} F(r)\phi^{(+)}_{j,m} \\ -if(r)\phi^{(-)}_{j,m} \end{pmatrix} \quad \text{or} \quad \psi = \begin{pmatrix} G(r)\phi^{(-)}_{j,m} \\ -ig(r)\phi^{(+)}_{j,m} \end{pmatrix}, \tag{7.25}$$

where $F(r)$, $f(r)$, $G(r)$, $g(r)$ are radial functions that are yet to be determined. (The factors of $-i$ are introduced in (7.25) for convenience so that some subsequent equations are real.)

4. The Radial Eigenvalues

Equations (7.15) and (7.16) with the first of (7.25) reduce to

$$(E - m_e c^2 + e\phi)F - \hbar c\left(\frac{d}{dr} + \frac{j+\frac{3}{2}}{r}\right) f = 0, \tag{7.26}$$

$$(E + m_e c^2 + e\phi)f + \hbar c\left(\frac{d}{dr} - \frac{j-\frac{1}{2}}{r}\right) F = 0. \tag{7.27}$$

The corresponding set of equations for the second choice in (7.25) is

$$(E + m_e c^2 + e\phi)G - \hbar c\left(\frac{d}{dr} + \frac{j-\frac{1}{2}}{r}\right) g = 0, \tag{7.28}$$

$$(E - m_e c^2 + e\phi)g + \hbar c\left(\frac{d}{dr} - \frac{j+\frac{3}{2}}{r}\right) G = 0. \tag{7.29}$$

On introducing the Coulomb potential (7.2) and writing

$$\lambda = j + \tfrac{1}{2}, \quad \varepsilon = E/m_e c^2, \quad x = r/r_c, \quad r_c = \hbar/m_e c, \quad \alpha = e^2/4\pi\varepsilon_0\hbar c, \tag{7.30}$$

(7.26) and (7.27) reduce to

$$\left(\varepsilon - 1 + \frac{Z\alpha}{x}\right) F - \left(\frac{d}{dx} + \frac{\lambda+1}{x}\right) f = 0, \tag{7.31}$$

$$\left(\varepsilon + 1 + \frac{Z\alpha}{x}\right) f + \left(\frac{d}{dx} - \frac{\lambda-1}{x}\right) F = 0, \tag{7.32}$$

respectively. The other pair of equations (7.28) and (7.29) reduces to a similar pair of equations obtained from (7.31) and (7.32) by the replacements

$$F \to G \quad f \to g \quad \lambda \to -\lambda. \tag{7.33}$$

As a result it suffices to consider only the pair (7.31) and (7.32) in detail.

To proceed formally, one eliminates either F or f from (7.31) and (7.32) and obtains a second order differential equation for the other. Then the conditions for normalizable solutions of the resulting equation are sought. This is closely analogous to the conventional procedure for constructing the radial wave functions for the hydrogen atom in nonrelativistic quantum mechanics, in which case the normalizable solutions are generalized Laguerre polynomials. As in the nonrelativistic case there are normalizable solutions only for specific values of the energy, and these specific values are the energy eigenvalues. In the relativistic case the normalizable solutions can be written in terms of the hypergeometric function. The detailed solutions are not derived here, and the following heuristic discussion is aimed at determining the energy eigenvalues without actually solving the second order differential equation.

A prelimiary point is that we are interested in bound states, which correspond to $E < m_e c^2$ and hence to $\varepsilon < 1$. For large x the asymptotic form of the second order differential equation obtained from (7.31) and (7.32) have solutions $\propto \exp[\pm(1-\varepsilon^2)^{1/2}x]$, and clearly only the $-$ sign is consistent with the solutions being normalizable. Hence we may seek solutions of the form

$$\begin{pmatrix} F \\ f \end{pmatrix} = e^{-(1-\varepsilon^2)^{1/2}x} x^\gamma \sum_{\nu=0} \begin{pmatrix} a_\nu \\ b_\nu \end{pmatrix} x^\nu. \tag{7.34}$$

On substituting (7.34) into (7.31) and (7.32) one obtains

$$(\varepsilon - 1)a_{\nu-1} + Z\alpha a_\nu + (1-\varepsilon^2)^{1/2} b_{\nu-1} - (\lambda + 1 + \gamma + \nu) b_\nu = 0, \tag{7.35}$$

$$(\varepsilon + 1)b_{\nu-1} + Z\alpha b_\nu - (1-\varepsilon^2)^{1/2} a_{\nu-1} - (\lambda - 1 - \gamma - \nu) a_\nu = 0, \tag{7.36}$$

The power of γ is determined by requiring that for $\nu = 0$ the terms a_{-1} and b_{-1} in (7.35) and (7.36) be zero. This gives $Z^2\alpha^2 = \lambda^2 - (\gamma+1)^2$, which has two solutions, one of which would make the wave function too singular at $r = 0$ to be normalizable. The remaining solution is

$$\gamma = -1 + [(j + \tfrac{1}{2})^2 - Z^2\alpha^2]^{1/2}. \tag{7.37}$$

Interestingly, real solutions are possible for $j = \frac{1}{2}$ only if the nucleus is not too highly charged, specifically, only for $Z < 1/\alpha \approx 137$.

As in nonrelativistic quantum mechanics, the next step in the argument is that normalizable solutions result only if the power series expansions (7.23) terminate, so that the function multiplying $\exp[\pm(1-\varepsilon^2)^{1/2}x]$ is a polynomial. Let us suppose that the highest term is $\nu = n'$, so that we have $a_{n'+1} = b_{n'+1} = 0$. Then for $\nu = n'$ in (7.35) and (7.36) one obtains

$$\frac{b_{n'}}{a_{n'}} = \left(\frac{1-\varepsilon}{1+\varepsilon}\right)^{1/2}. \tag{5.38}$$

By multiplying (7.35) by $(1+\varepsilon^2)^{1/2}$ and (7.36) by $(1-\varepsilon^2)^{1/2}$ and adding the resulting equations, one obtains

$$\frac{b_\nu}{a_\nu} = \frac{Z\alpha(1+\varepsilon^2)^{1/2} + (\lambda - 1 - \gamma - \nu)(1-\varepsilon^2)^{1/2}}{Z\alpha(1-\varepsilon^2)^{1/2} + (\lambda + 1 + \gamma + \nu)(1+\varepsilon^2)^{1/2}}. \tag{7.39}$$

On setting $\nu = n'$ in (7.39) and comparing with (7.38), one obtains an identity that determines the energy eigenvalues. These are given by the fine structure formula (7.11).

Lecture 8

Nonrelativistic Limit of the Dirac Equation

In the nonrelativistic limit the Dirac theory should reproduce the Schrödinger-Pauli theory. Furthermore the Dirac theory must enable one to derive corrections to the Schrödinger-Pauli theory in the limit where relativistic effects are weak.

1. The Large and Small Components

In the standard representation the upper two components of the Dirac wavefunction Ψ (which is a column matrix with four components) are much larger than the lower two components. This property enables one to derive the nonrelativistic limit in a simple way.

Reverting to normal units, the Dirac equation in the presence of an external field is

$$i\hbar \frac{\partial}{\partial t} \Psi = \left[c\boldsymbol{\alpha} \cdot (\hat{\boldsymbol{p}} + e\boldsymbol{A}) + \beta mc^2 - e\Phi \right] \Psi. \tag{8.1}$$

One now appeals to the fact that the energy ε is close to mc^2 in the nonrelativistic limit. The positive energy part of the wavefunction varies with time as $\exp[-i\varepsilon t/\hbar]$, and this may be approximated by $\exp[-imc^2 t/\hbar]$. The slower variation as $\exp[-i(\varepsilon - mc^2)t/\hbar]$ is retained in

$$\Psi' := \Psi \, \exp[imc^2 t/\hbar]. \tag{8.2}$$

Next one writes

$$\Psi' = \begin{pmatrix} \phi' \\ \chi' \end{pmatrix} \tag{8.3}$$

where ϕ' and χ' are 2-spinors, i.e., column matrices with two components. Then (8.1) reduces to

$$(i\hbar \frac{\partial}{\partial t} + e\Phi)\phi' = c\boldsymbol{\sigma} \cdot (\hat{\boldsymbol{p}} + e\boldsymbol{A})\chi', \tag{8.4}$$

$$(i\hbar \frac{\partial}{\partial t} + e\Phi + 2mc^2)\chi' = c\boldsymbol{\sigma} \cdot (\hat{\boldsymbol{p}} + e\boldsymbol{A})\phi'. \tag{8.5}$$

Henceforth the primes on ϕ and χ are omitted.

Consider the left hand side of (8.5). The term $2mc^2$ is much larger than the other two terms in the nonrelativistic limit. Making the approximation in which only this dominant term is retained, one has

$$\chi' \approx \frac{1}{2mc} \boldsymbol{\sigma} \cdot (\hat{\boldsymbol{p}} + e\boldsymbol{A})\phi'. \tag{8.6}$$

Substituting (8.6) into (8.4) gives

$$(i\hbar \frac{\partial}{\partial t} + e\Phi)\phi' \approx \frac{1}{2m} \left[\boldsymbol{\sigma} \cdot (\hat{\boldsymbol{p}} + e\boldsymbol{A})\right]^2 \phi'. \tag{8.7}$$

For arbitrary vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ one has

$$\boldsymbol{\sigma}\cdot\boldsymbol{a}\,\boldsymbol{\sigma}\cdot\boldsymbol{b}=\boldsymbol{a}\cdot\boldsymbol{b}+i\boldsymbol{\sigma}\cdot\boldsymbol{a}\times\boldsymbol{b}. \tag{8.8}$$

On applying this to the right hand side of (8.7) and using

$$[\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}}+e\boldsymbol{A})]\,[\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}}+e\boldsymbol{A})]=(\hat{\boldsymbol{p}}+e\boldsymbol{A})^2+ie\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}}\times\boldsymbol{A}+\boldsymbol{A}\times\hat{\boldsymbol{p}}), \tag{8.9}$$

the final term reduces to $-ie\hbar\,\mathrm{curl}\boldsymbol{A}=-ie\hbar\boldsymbol{B}$, and equation (8.7) reduces to

$$i\hbar\frac{\partial}{\partial t}\,\phi=\left[\frac{1}{2m}\,(\hat{\boldsymbol{p}}+e\boldsymbol{A})^2-e\Phi+\frac{e\hbar}{2m}\,\boldsymbol{\sigma}\cdot\boldsymbol{B}\right]\phi. \tag{8.10}$$

Equation (8.10) is just the Pauli-Schrödinger equation (1.20), as required.

2. The Foldy-Wouthuysen Transformation

There is a systematic way of generalizing this approach to derive a generalization of the Pauli-Schrödinger equation which includes relativistic corrections. The idea is to separate the Hamiltonian into *even* and *odd* components, denoted $\hat{\mathcal{O}}$ and $\hat{\mathcal{E}}$ respectively. These are defined to satisfy

$$\beta\hat{\mathcal{O}}=-\hat{\mathcal{O}}\beta,\qquad\beta\hat{\mathcal{E}}=\hat{\mathcal{E}}\beta. \tag{8.11}$$

One has

$$\hat{\mathcal{H}}=\beta mc^2+\hat{\mathcal{O}}+\hat{\mathcal{E}}, \tag{8.12}$$

with

$$\hat{\mathcal{O}}=c\boldsymbol{\alpha}\cdot(\hat{\boldsymbol{p}}+e\boldsymbol{A}),\quad\hat{\mathcal{E}}=-e\Phi, \tag{8.13}$$

for the Hamiltonian in (8.1). Suppose a change of representation is made so that in the new representation $\hat{\mathcal{O}}$ is zero. Then the equations for the large and small components are decoupled, and the equations for the large component is the desired generalization of the Pauli-Schrödinger equation. The appropriate change of representation is called the *Foldy-Wouthuysen* transformation.

Let the (unitary) transformation operator be written as $\exp[i\hat{S}]$, where $\hat{S}$ is hermitian. Our objective is to find the appropriate $\hat{S}$ and hence the generalization of the Pauli-Schrödinger equation. The new time-dependent Schrödinger equation is

$$i\hbar\frac{\partial}{\partial t}\,\Psi'=\hat{\mathcal{H}}'\Psi', \tag{8.14}$$

with

$$\Psi'=\exp(i\hat{S})\,\Psi,\quad\hat{\mathcal{H}}'=\exp(i\hat{S})\,\hat{\mathcal{H}}\,\exp(-i\hat{S})-i\hbar\,\exp(i\hat{S})\,\frac{\partial}{\partial t}\,\exp(-i\hat{S}). \tag{8.15}$$

The detailed evaluation of the new Hamiltonian is made using the identity

$$\exp(\hat{A})\,\hat{B}\,\exp(-\hat{A})=\hat{B}+[\hat{A},\hat{B}]+\tfrac{1}{2}[\hat{A},[\hat{A},\hat{B}]]+\cdots, \tag{8.16}$$

where the nth term contains n nested commutators and a numerical factor $(n!)^{-1}$. Thus one finds

$$\begin{aligned}\hat{\mathcal{H}}' = \hat{\mathcal{H}} + i[\hat{S},\hat{\mathcal{H}}] - \tfrac{1}{2}[\hat{S},[\hat{S},\hat{\mathcal{H}}]] + \cdots \\ - \hbar\,\frac{\partial}{\partial t}\,\hat{S} - i\tfrac{1}{2}\hbar\,[\hat{S},\frac{\partial}{\partial t}\,\hat{S}] + \cdots .\end{aligned} \tag{8.17}$$

Now suppose that $\hat{S}$ contains a small parameter. This is necessarily the case if the nonrelativistic effects are sufficiently weak. A perturbation expansion involves expanding in this parameter, i.e., in powers of $\hat{S}$. (It is assumed that $\partial\hat{S}/\partial t$ is of higher order than $\hat{S}$.) To zeroth order the Hamiltonian is βmc^2. To first order in $\hat{S}$ one finds

$$\hat{\mathcal{H}}' = \beta mc^2 + \hat{\mathcal{O}} + \hat{\mathcal{E}} + i[\hat{S},\beta mc^2]. \tag{8.18}$$

To this order the odd terms may be eliminated by requiring

$$\hat{\mathcal{O}} + i[\hat{S},\beta mc^2] = 0. \tag{8.19}$$

Using (8.11) the solution of (8.19) is

$$\hat{S} = -i\,\frac{\beta\hat{O}}{2mc^2}\,. \tag{8.20}$$

To find the Hamiltonian to this order one inserts (8.20) in (8.17) and retains only the second order terms. The result is the Pauli-Schrödinger equation.

3. Relativistic Corrections to the Pauli-Schrödinger Equation

The expansion (7.17) to next order contains odd terms which may be eliminated in the same way. That is one chooses a correction to (8.20) that causes the odd terms to also vanish to next highest order. This generalization of (8.20) is then used to evaluated the next highest order terms, and so find relativisitic corrections to the Pauli-Schrödinger equation.

Omitting the details of the calculation, the result is

$$\begin{aligned}\hat{\mathcal{H}} = \beta mc^2 + \frac{1}{2m}\,(\hat{\boldsymbol{p}} + e\boldsymbol{A})^2 - e\Phi + \frac{e\hbar}{2m}\,\boldsymbol{\sigma}\cdot\boldsymbol{B} - \frac{1}{8m^3c^2}\left\{(\hat{\boldsymbol{p}} + e\boldsymbol{A})^2 + e(\boldsymbol{\sigma}\cdot\boldsymbol{B})\right\}^2 \\ + \frac{ie\hbar^2}{8m^2c^2}\,\boldsymbol{\sigma}\cdot\mathrm{curl}\boldsymbol{E} + \frac{e\hbar}{4m^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{E}\times(\hat{\boldsymbol{p}} + e\boldsymbol{A}) + \frac{e\hbar^2}{8m^2c^2}\,\mathrm{div}\boldsymbol{E}.\end{aligned} \tag{8.21}$$

The first term on the right hand side of (8.21) corresponds to the rest energy for electrons, when only the leading 2×2 elements are retained. The second, third and fourth terms on the right hand side of (8.21) are just the Pauli-Schrödinger terms. The fourth term may be interpreted as a relativisitic correction from the expansion of $\sqrt{m^2c^4 + \hat{\boldsymbol{p}}^2c^2} - mc^2$ to second order in $\hat{\boldsymbol{p}}^2$. The fifth term is rarely important; Maxwell's equations allow one to re-express $\mathrm{curl}\boldsymbol{E}$ as $-\partial\boldsymbol{B}/\partial t$, and hence this term is important only when the rate of time derivative of the magnetic field is important. The sixth term is the spin orbit coupling

term which is introduced is an artificial way into the nonrelativistic theory. The final term is called the *Darwin* term.

The Darwin term may be attributed to the Zitterbewegung. An electron fluctuates in position very rapidly over a distance of order its Compton wavelength $\hbar/mc$. The effective potential that it experiences is the potential Φ at its mean position plus a correction due to the average over these rapid fluctuations. Let the correction be $\Delta\Phi$. On making a Taylor series expansion about the mean position $\boldsymbol{x}_0$, with the fluctuations over a distance $\Delta\boldsymbol{x}$, one has

$$\Phi(\boldsymbol{x}) = \Phi(\boldsymbol{x}_0) + \Delta\boldsymbol{x} \cdot \frac{\partial}{\partial \boldsymbol{x}_0} \Phi(\boldsymbol{x}_0) + \tfrac{1}{2}\Delta\boldsymbol{x}_i \Delta\boldsymbol{x}_j \frac{\partial^2}{\partial \boldsymbol{x}_{0i} \partial \boldsymbol{x}_{0j}} \Phi(\boldsymbol{x}_0) + \cdots . \tag{8.22}$$

On averaging over the position of the electron, the term linear in $\Delta\boldsymbol{x}$ averages to zero, and the second order term gives

$$\left\langle \tfrac{1}{2}\Delta\boldsymbol{x}_i \Delta\boldsymbol{x}_j \frac{\partial^2}{\partial \boldsymbol{x}_{0i} \partial \boldsymbol{x}_{0j}} \Phi(\boldsymbol{x}_0) \right\rangle \approx \left(\frac{\hbar}{mc} \right)^2 \nabla^2 \Phi, \tag{8.23}$$

where the angular brackets denote an average over position, and where $\langle \Delta\boldsymbol{x}^2 \rangle$ is assumed to be of order $(\hbar/mc)^2$. Apart from the numerical factor, (8.23) with $\nabla^2\Phi = -\mathrm{div}\boldsymbol{E}$ reproduces the Darwin term.

Lecture 9

Formal Aspects of the Dirac Equation

Certain properties of the Dirac matrices are needed both for formal purposes and for performing detailed calculations.

1. Independent 4×4 Matrices

The Dirac matrices γ^μ may be represented by 4×4 matrices which have either real or imaginary entries. Thus the matrices are not intrinsically complex, and the algebra of the matrices is equivalent to that of real 4×4 matrices. There are 16 independent such matrices.

A suitable choice for the 16 matrices involves

$$\gamma^5 := -i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{9.1}$$

Note that in the literature there several different definitions of γ^5, involving replacing $-i$ by i or ± 1 in (9.1). One finds that γ^5 satisfies the following relations:

$$\gamma^\mu\gamma^5 + \gamma^5\gamma^\mu = 0, \qquad (\gamma^5)^2 = 1, \qquad (\gamma^5)^\dagger = \gamma^5. \tag{9.2}$$

The standard representation of γ^5 is

$$\gamma^5 = \begin{pmatrix} \mathbf{0} & -\mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}. \tag{9.3}$$

A convenient choice of 16 independent matrices consists of the set

$$\gamma^A = \Big[1, \quad \gamma^\mu, \quad S^{\mu\nu}, \quad \gamma^\mu\gamma^5, \quad \gamma^5\Big], \tag{9.4}$$

where A runs from 1 to 16, and with

$$S^{\mu\nu} = \tfrac{1}{4}\, i[\gamma^\mu, \gamma^\nu]. \tag{9.5}$$

This choice involves a scalar and a pseudo scalar (1, γ^5), a 4-vector and a pseudo 4-vector (γ^μ, $\gamma^\mu\gamma^5$) and an antisymmetric second rank 4-tensor ($S^{\mu\nu}$), these having 1, 1, 4, 4, and 6 components respectively.

In principle any product or sum of products of γ-matrices may be re-expressed as a sum of terms involving only these 16 independent matrices. For example one finds

$$\begin{aligned}
\gamma^\mu\gamma^\nu &= g^{\mu\nu} - 2iS^{\mu\nu}, \\
\gamma^\mu\gamma^\nu\gamma^\rho &= g^{\mu\nu}\gamma^\rho - g^{\mu\rho}\gamma^\nu + g^{\nu\rho}\gamma^\mu - i\epsilon^{\mu\nu\rho\sigma}\gamma_\sigma\gamma^5, \\
\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma &= g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho} - 2i\Big\{g^{\mu\nu}S^{\rho\sigma} - g^{\mu\rho}S^{\nu\sigma} + g^{\mu\sigma}S^{\nu\rho} \\
&\qquad + S^{\mu\nu}g^{\rho\sigma} - S^{\mu\rho}g^{\nu\sigma} + S^{\mu\sigma}g^{\nu\rho}\Big\} + i\epsilon^{\mu\nu\rho\sigma}\gamma^5,
\end{aligned} \tag{9.6}$$

and so on. Also one has

$$\begin{aligned}
\epsilon^{\mu\nu\rho\sigma}S_{\rho\sigma} &= 2i\, S^{\mu\nu}\,\gamma^5, \\
\epsilon^{\mu\nu\rho\sigma}\gamma_\nu\gamma_\rho\gamma_\sigma &= -i\,3!\,\gamma^\mu\gamma^5 \\
\epsilon^{\mu\nu\rho\sigma}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma &= -i\,4!\,\gamma^5,
\end{aligned} \tag{9.7}$$

2. Traces of Products of γ-Matrices

The trace (Tr) of the unit 4×4 tensor is equal to four, and the traces of the remaining 15 matrices (9.4) are all zero. The traces of products of γs are important in detailed calculations in QED.

Consider

$$T^{\alpha_1\alpha_2\ldots\alpha_n} := \mathrm{Tr}\left(\gamma^{\alpha_1}\gamma^{\alpha_2}\ldots\gamma^{\alpha_n}\right). \tag{9.8}$$

One has the following results:

(i) The trace of an odd number of γs is zero,

$$T^{\alpha_1\alpha_2\ldots\alpha_n} = 0 \qquad \text{for } n \text{ odd.} \tag{9.9}$$

A derivation of this result involves premultiplying by $\gamma^5\gamma^5$ inside the trace in (9.8). According to (9.2) this is equivalent to multiplying by unity. One may move one γ^5 to the end of the other γs by consecutive interchanges; an odd number of changes in sign occurs in view of the anticommutation relations (9.2). The other γ^5 may then be moved directly to the end using the invariance of the trace under cyclic permutations. Then using $\gamma^5\gamma^5 = 1$, the trace is equal to minus itself, completing the proof of (9.9).

(ii) The trace of $\gamma^\mu\gamma^\nu$ follows directly from the definition (3.7) and the invariance of the trace under cyclic permutations:

$$T^{\mu\nu} = 4g^{\mu\nu}. \tag{9.10}$$

(iii) The values of $T^{\alpha_1\alpha_2\ldots\alpha_n}$ for $n = 4, 6\ldots$ may be derived sequentially as follows. For given n move γ^{α_1} in (9.8) to the end using the anticommutation relations for each interchange, and then use the invariance of the trace under cyclic permutations to derive an identity that involves only the traces of $n-2$ γ matrices. In this way one finds

$$\begin{aligned} T^{\mu\nu\rho\sigma} &= 4\Big[g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}\Big], \\ T^{\mu\nu\rho\sigma\alpha\beta} &= 4\Big[g^{\mu\nu}T^{\rho\sigma\alpha\beta} - g^{\mu\rho}T^{\nu\sigma\alpha\beta} + g^{\mu\sigma}T^{\nu\rho\alpha\beta} - g^{\mu\alpha}T^{\nu\rho\sigma\alpha} + g^{\mu\beta}T^{\nu\rho\sigma\alpha}\Big], \end{aligned} \tag{9.11}$$

and so on.

Another set of relations involves contractions over γ-matrices:

$$\begin{aligned} \gamma^\mu \not{a}\gamma_\mu &= -2\not{a}, \\ \gamma^\mu \not{a}\not{b}\gamma_\mu &= 4ab, \\ \gamma^\mu \not{a}\not{b}\not{c}\gamma_\mu &= -2\not{c}\not{b}\not{a}, \\ \gamma^\mu \not{a}\not{b}\not{c}\not{d}\gamma_\mu &= 2(\not{c}\not{d}\not{a}\not{b} + \not{c}\not{b}\not{a}\not{d}). \end{aligned} \tag{9.12}$$

3. Other Representations of the Dirac Matrices

So far only the standard representation has been mentioned. An arbitrary representation may be obtained by introducing a 4×4 transformation matrix S with unit determinant:

$$\det S = 1. \tag{9.13}$$

The wavefunctions and the γ matrices are transformed according to

$$\Psi'(x) = S\,\Psi(x), \qquad \Psi'^{\dagger}(x) = \Psi^{\dagger}(x)S^{-1}, \qquad \gamma'^{A} = S\gamma^{A}S^{-1}. \tag{9.14}$$

Besides allowing changes in representation, such transformations need to be considered when making Lorentz transformations and other transformations such as the parity, time reversal and charge conjugation transformations.

One alternative that is often used is the spinor representation. In this representation, the wavefunction is written

$$\Psi(x) = \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \tag{9.15}$$

where the dependence of the spinors ξ and η on x is implicit. The Dirac equation becomes the two coupled equations

$$(p^0 + \boldsymbol{p}\cdot\boldsymbol{\sigma})\eta = m\xi, \qquad (p^0 - \boldsymbol{p}\cdot\boldsymbol{\sigma})\xi = m\eta. \tag{9.16}$$

Thus in the *spinor* representation one has

$$\boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{\sigma} \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}. \tag{9.17}$$

A suitable choice for the transformation matrix is

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{1} & -\mathbf{1} \\ \mathbf{1} & \mathbf{1} \end{pmatrix}. \tag{9.18}$$

In the *Majorana* representation all the matrices are real. This may be obtained from the standard representation using the transformation matrix

$$S = \frac{1}{\sqrt{2}}\,(\alpha_y + \beta). \tag{9.19}$$

4. Lorentz Transformation of the Dirac Equation

The Klein Gordon equation (3.1) is in a manifestly covariant form because the wavefunction is a Lorentz invariant for spin 0. The Dirac equation (3.4) is in a covariant notation, but before the form can be said to be covariant it is necessary to specify how the wavefunction and the γ matrices transform.

An arbitrary Lorentz transformation is of the form

$$x^{\mu'} = L^{\mu'}{}_{\mu}\, x^{\mu} + a^{\mu'}. \tag{9.20}$$

One has

$$\partial_{\mu'} = L^{\mu}{}_{\mu'}\partial_{\mu}. \tag{9.21}$$

On applying the transformation (9.20) to the γ matrices, it is obvious that one has

$$\gamma^{\mu'}\partial_{\mu'} = L^{\mu'}{}_{\mu}\gamma^{\mu}L^{\nu}{}_{\mu'}\partial_{\nu} = \gamma^{\mu}\partial_{\mu}. \tag{9.22}$$

However the new γs are not in the same representation as the old γs.

To restore the original representation one needs to introduce a transformation matrix $S(L)$ in the 4-dimensional spin space. The transformed wavefunction is

$$\Psi'(x') = S(L)\,\Psi(L^{-1}x'), \tag{9.23}$$

where $x = L^{-1}x'$ denotes the inverse of (9.20). The transformation of the γ matrices is awkward to write. In words, $S(L)$ needs to be such that the new

$$S(L)\,\gamma^{\mu'}S^{-1}(L)$$

are term by term equal to the old γ^{μ}. That is, $S(L)$ must be such that the new $\gamma^{0'}$ transforms into the original γ^0, and so on.

Thus one requires two different sets of transformation matrices in discussing Lorentz transformations of the Dirac wavefunction and the Dirac matrices. One set consists of the the $L^{\mu'}{}_{\mu}$ that operate in the 4-dimensional space-time. The other set consists of the $S(L)$ that operate in the 4-dimensional spin space. Before considering the relation between these it is appropriate to discuss the generators of the transformations.

5. Explicit Forms for the Transformation Matrices $S(L)$

A Lorentz transformation may be described in terms of a 4-dimensional rotation with an antisymmetric tensor $\omega_{\alpha\beta}$. The space component of this tensor are related to ordinary rotation, with (to within a sign) ω_{12} being the angle through with the rotation about the 3-axis is made. The other components are related to boosts, as described below, with ω_{01} being related to a boost along the 1-axis.

The transformation matrix for a finite 4-dimensional rotation is

$$S = \exp\left[i\tfrac{1}{2}\omega_{\alpha\beta}S^{\alpha\beta}\right]. \tag{9.24}$$

For a boost along the i-axis by a speed v, all the components of $\omega_{\alpha\beta}$ are zero except $\omega_{0i} = -\omega_{0i} = \zeta$, with

$$v = \tanh\zeta, \qquad p = m\sinh\zeta, \qquad \gamma = \cosh\zeta, \tag{9.25}$$

and for a rotation through an angle ϕ about the i-axis, all the components of $\omega_{\alpha\beta}$ are zero except $\omega_{jk} = -\omega_{kj} = \phi$, where ijk is an even permutation of 123.

In the standard representation the explicit forms for a boost along a unit vector $\boldsymbol{n}$ and a rotation about a unit vector $\boldsymbol{n}$ are

$$S = \exp[-\tfrac{1}{2}\boldsymbol{\alpha}\cdot\boldsymbol{n}\zeta] = \cosh\tfrac{1}{2}\zeta - \boldsymbol{\alpha}\cdot\boldsymbol{n}\sinh\tfrac{1}{2}\zeta, \tag{9.26}$$

$$S = \exp[\tfrac{i}{2}\boldsymbol{\sigma}\cdot\boldsymbol{n}\phi] = \cos\tfrac{1}{2}\phi + i\boldsymbol{\sigma}\cdot\boldsymbol{n}\sin\tfrac{1}{2}\phi, \tag{9.27}$$

respectively.

7. The Parity, Time Reversal and Charge Conjugation Operators

Besides the continuous set of Lorentz transformations, there are also discrete transformations. These include the parity, time reversal and charge conjugation transformations. Parity corresponds to a reflection of the coordinate axes $\boldsymbol{x} \to -\boldsymbol{x}$, time reversal corresponds to $t \to -t$, and charge conjugation to the reversal of the roles of particles and antiparticles. It is convenient to define *classical* operators P_0 and T_0 which apply to any classical field:

$$P_0\psi(t,\boldsymbol{x}) = \psi(t,-\boldsymbol{x}), \quad T_0\psi(t,\boldsymbol{x}) = \psi(-t,\boldsymbol{x}). \tag{9.28}$$

The parity transformation for the Dirac equation involves a unitary operator U_P such that one has

$$\Psi_P(t,\boldsymbol{x}) = U_P\Psi(t,-\boldsymbol{x}). \tag{9.29}$$

To within an arbitrary phase factor η_P, one finds

$$U_P = i\eta_P P_0\gamma^0. \tag{9.30}$$

For both time reversal and charge conjugation, one needs to consider the complex conjugate $\Psi(x)^*$ of the wave function. If one writes

$$\Psi(x)^* = K\Psi(x), \tag{9.31}$$

then K is anti-unitary satisfying

$$K^2 = 1, \quad K^\dagger = -K^{-1} = -K. \tag{9.32}$$

The time reversal transformation is satisfied as follows:

$$\Psi_T(t,\boldsymbol{x}) = T\big(\Psi(-t,\boldsymbol{x})\big)^T, \quad T = -\gamma^1\gamma^3 K T_0. \tag{9.33}$$

Charge conjugation involves the transpose of the Dirac wavefunction. It is satisfied for

$$\Psi_C(x) = C\big(\Psi(x)\big)^T, \quad C = i\gamma^2 K. \tag{9.34}$$

Appendix A

Fourier Transforms

Fourier transformed quantities are used widely in the kinetic theory of plasmas. Here relevant properties of Fourier transforms are summarized with particular emphasis on result that are used frequently below.

1. 4-Dimensional Fourier Transform

The Fourier transform $\tilde{G}(k)$ of a function $G(x)$ is defined by

$$\tilde{G}(k) := \int d^4x\, e^{ikx}\, G(x), \tag{A.1}$$

with $d^4x = dx^0dx^1dx^2dx^3$. The inverse transform is

$$G(x) = \int \frac{d^4k}{(2\pi)^4}\, e^{-ikx}\tilde{G}(k). \tag{A.2}$$

Except in the remainder of this section, the tilde on $\tilde{G}(k)$ will be omitted.

Three elementary properties of Fourier transforms are the following.

The Reality Condition: If $G(x)$ is real then $\tilde{G}(k)$ satisfies

$$\tilde{G}^*(k) = \tilde{G}(-k), \tag{A.3}$$

where * denotes complex conjugation.

The Power Theorem: If $G_1(x)$ and $G_2(x)$ have Fourier transforms $\tilde{G}_1(k)$ and $\tilde{G}_2(k)$, respectively, then one has

$$\int d^4x\, G_1(x)\, G_2(x) = \int \frac{d^4k}{(2\pi)^4}\, \tilde{G}_1(k)\, \tilde{G}_2(-k)). \tag{A.4}$$

This result is referred to as the power theorem.

The Convolution Theorem: The Fourier transform $\tilde{H}(k)$ of the product

$$H(x) = G_1(x)\, G_2(x) \ldots G_n(x) \tag{A.5}$$

is the convolution of the Fourier transforms:

$$\tilde{H}(k) = \int d\lambda^{(n)}\, \tilde{G}_1(k_1)\, \tilde{G}_2(k_2) \ldots \tilde{G}_n(k_n)), \tag{A.6}$$

where

$$d\lambda^{(n)} = \frac{d^4k_1}{(2\pi)^4}\, \frac{d^4k_2}{(2\pi)^4}\, \cdots\, \frac{d^4k_n}{(2\pi)^4}\, (2\pi)^4\delta^4(k - k_1 - k_2 - \cdots - k_n) \tag{A.7}$$

denotes the n-fold convolution integral, and $\delta^4(k)$ denotes the 4-dimensional δ-function which is equal to $\delta(k^0)\delta(k^1)\delta(k^2)\delta(k^3)$, where k^0, k^1, k^2, k^3 are the four components of the 4-vector k.

2. Truncations and the Dirac δ-functions

The Fourier integral theorem requires that the function $G(x)$ be amplitude-integrable in order for its Fourier transform $\tilde{G}(k)$ to exist. In practice one is often concerned with idealized functions that do not vanish at infinity, and formally the Fourier transforms of such functions do not exist. However, the Fourier transform may still be defined as a generalized function, i.e., as a sequence of well-defined functions. One way of defining a sequence of functions each of whose Fourier transform exists is by replacing $G(x)$ by a truncated function equal to $G(x)$ inside a large space-time volume TV and zero outside this space-time volume, and then allowing TV to tend to infinity. One usually assumes implicitly that such truncations are performed, and ignores them. One situation where one needs to take account of the truncation explicitly is when the square of a δ-function arises.

The 4-dimensional Dirac δ-function is defined as the Fourier transform of unity, i.e.,

$$(2\pi)^4\,\delta^4(k) = \int d^4x\, e^{ikx}. \tag{A.8}$$

When the truncation is taken into account, in the limit of arbitrarily large TV, one has

$$[(2\pi)^4\,\delta^4(k)]^2 = TV\,(2\pi)^4\,\delta^4(k). \tag{A.9}$$

Similar relations apply independently to the temporal and spatial parts:

$$[2\pi\,\delta(\omega)]^2 = T\,2\pi\,\delta(\omega), \qquad [(2\pi)^3\,\delta^3(\boldsymbol{k})]^2 = V\,(2\pi)^3\,\delta^3(\boldsymbol{k}). \tag{A.10}$$

3. Fourier Transforms of the Step and Sign Function

Two other generalized functions which are important here may be defined as the Fourier transforms of the step and sign functions of time.

The step function $H(t)$ is defined by

$$H(t) := \begin{cases} 1 & \text{for} \quad t > 0, \\ 0 & \text{for} \quad t < 0. \end{cases} \tag{A.11}$$

One way of defining the generalized function is to replace unity in (A.11) by $e^{-\eta t}$ and allow $\eta > 0$ to tend to zero. The temporal Fourier transform is

$$\tilde{H}(\omega) = \lim_{\eta\to 0} \int_0^\infty dt\, \exp[i\omega t - \eta t] = \frac{i}{\omega + i0}, \tag{A.12}$$

which $i0$ denotes the limit of $i\eta$ as η tends to zero from above. The integral (A.12) defines the generalized function $i/(\omega + i0)$.

The Fourier transform of the sign function

$$S(t) := \frac{t}{|t|} \tag{A.13}$$

may be identified by replacing t by $t\,e^{-\eta t}$ for $t > 0$, and by $-t\,e^{\eta t}$ for $t < 0$. The resulting generalized function is

$$\tilde{S}(\omega) = \lim_{\eta\to 0}\int_0^\infty dt\,\exp[i\omega t - \eta t]) - \int_{-\infty}^0 dt\,\exp[i\omega t + \eta t]$$

$$= \lim_{\eta\to 0}\frac{2i\omega}{\omega^2+\eta^2} = 2i\wp\frac{1}{\omega}, \tag{A.14}$$

where

$$\wp\frac{1}{\omega} := \begin{cases} 1/\omega & \text{for} \quad \omega \neq 0, \\ 0 & \text{for} \quad \omega = 0, \end{cases} \tag{A.15}$$

denotes the operation of taking the Cauchy principal value of an ω-integral.

The Plemelj Formula

The identity

$$H(t) = \tfrac{1}{2}\,[1 + S(t)], \tag{A.16}$$

when Fourier transformed, implies the Plemelj formula

$$\frac{1}{\omega + i0} = \wp\frac{1}{\omega} - i\pi\,\delta(\omega). \tag{A.17}$$

This result, inside an integral, is important in allowing one to separate so-called resonant and nonresonant parts. The prescription for doing so is equivalent to one derived by Landau in 1946 using an argument based on Laplace transforming. The procedure of giving the frequency ω an infinitesimal positive imaginary part and interpreting poles in integrands according to (A.17) is often called the *Landau prescription.*

Confinement to the Forward Light Cone

The step function may be used to impose the causal condition. However, the special theory of relativity requires that an event depend only on events in its past light cone. To impose this stronger requirement we may use the function $H(t - \boldsymbol{\beta}\cdot\boldsymbol{x})$. The light cones may be spanned by allowing $\boldsymbol{\beta}$ to take on all values with $\boldsymbol{\beta}^2 < 1$. The Fourier transform in space and time of $H(t - \boldsymbol{\beta}\cdot\boldsymbol{x})$ is

$$\tilde{H}_{\boldsymbol{\beta}}(k) := \int d^4x\,e^{ikx}H(t - \boldsymbol{\beta}\cdot\boldsymbol{x}) = \frac{i}{\omega + i0}\,(2\pi)^3\,\delta^3(\boldsymbol{k} - \omega\boldsymbol{\beta}). \tag{A.18}$$

Appendix B

Specific Spin Eigenfunctions

In this appendix some specific solutions for wave functions are written down. These complement solutions given in the text.

1. Helicity States

The plane wave solutions when the spin operators is taken to be the helicity operator are

$$\boldsymbol{\Sigma}\cdot\hat{\boldsymbol{p}} = \begin{pmatrix} \epsilon p_z & \epsilon p_- & 0 & 0 \\ \epsilon p_+ & -\epsilon p_z & 0 & 0 \\ 0 & 0 & \epsilon p_z & \epsilon p_- \\ 0 & 0 & \epsilon p_+ & -\epsilon p_z \end{pmatrix}, \tag{B.1}$$

where $\epsilon\boldsymbol{p}$ is the eigenvalue of $\hat{\boldsymbol{p}}$. The identity $|\boldsymbol{\sigma}\cdot\boldsymbol{p}|^2 = |\boldsymbol{p}|^2$ implies that the eigenvalues of the helicity operator are $s|\boldsymbol{p}|$ with $s = \pm$.

The eigenfunctions may be found by operating on (4.12) with the helicity operator and determining C_s^ϵ and D_s^ϵ by the requirement that the solutions be eigenfunctions. Let the eigenvalues of the helicity be sh with $s = \pm$. The eigenvalues are found by setting the determinant of the matrix, cf. (B.1),

$$\begin{pmatrix} \epsilon p_z - sh & \epsilon p_- & 0 & 0 \\ \epsilon p_+ & -\epsilon p_z - sh & 0 & 0 \\ 0 & 0 & \epsilon p_z - sh & \epsilon p_- \\ 0 & 0 & \epsilon p_+ & -\epsilon p_z - sh \end{pmatrix}$$

equal to zero. This gives

$$h = (p_z^2 + p_+ p_-)^{1/2} = |\boldsymbol{p}|. \tag{B.2}$$

On writing

$$\begin{aligned} \boldsymbol{p} &= (p_\perp \cos\phi,\ p_\perp \sin\phi,\ p_z) \\ &= |\boldsymbol{p}|\ (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta), \end{aligned} \tag{B.3}$$

one finds that the ϕ-dependence can be satisfied by writing the solution in the form

$$\varphi_s^\epsilon(\epsilon\boldsymbol{p}) = \begin{pmatrix} C_1 e^{-i\phi/2} \\ C_2 e^{i\phi/2} \\ C_3 e^{-i\phi/2} \\ C_4 e^{i\phi/2} \end{pmatrix}. \tag{B.4}$$

Then the coefficients C_s^ϵ and D_s^ϵ in (4.12) satisfy

$$(\epsilon p_z - sh)C_s^\epsilon + \epsilon p_\perp D_s^\epsilon = 0, \tag{B.5a}$$

$$(\epsilon p_z - sh)(\epsilon p_z C_s^\epsilon + \epsilon p_\perp D_s^\epsilon) + \epsilon p_\perp(\epsilon p_\perp C_s^\epsilon - \epsilon p_z D_s^\epsilon) = 0. \tag{B.5b}$$

One may then solve for C_s^ϵ and D_s^ϵ, using the normalization condition (4.13), viz.

$$|C_s^\epsilon|^2 + |D_s^\epsilon|^2 = 1. \tag{B.6}$$

A common phase factor for C_s^ϵ and D_s^ϵ is arbitrary, and it is chosen for convenience. After using the identity

$$p_\perp = \sqrt{(h + \epsilon s p_z)(h - \epsilon s p_z)}, \tag{B.7}$$

a specific solution is

$$C_s^\epsilon = \epsilon\sqrt{\frac{h + \epsilon s p_z}{2h}}, \qquad D_s^\epsilon = s\sqrt{\frac{h - \epsilon s p_z}{2h}}. \tag{B.8}$$

A further identity used in simplifying the result is

$$h = \sqrt{(\varepsilon + \epsilon m)(\varepsilon - \epsilon m)}. \tag{B.9}$$

One then finds

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} = \frac{1}{\sqrt{2h2\varepsilon V}} \begin{pmatrix} \sqrt{\varepsilon + \epsilon m}\,\sqrt{h + \epsilon s p_z} \\ s\epsilon\sqrt{\varepsilon + \epsilon m}\,\sqrt{h - \epsilon s p_z} \\ s\epsilon\sqrt{\varepsilon - \epsilon m}\,\sqrt{h + \epsilon s p_z} \\ \sqrt{\varepsilon - \epsilon m}\,\sqrt{h - \epsilon s p_z} \end{pmatrix}. \tag{B.10}$$

Thus one obtains the following eigenfunctions:

$$\varphi_s^\epsilon(\epsilon \boldsymbol{p}) = \frac{1}{\sqrt{2h2\varepsilon V}} \begin{pmatrix} \sqrt{\varepsilon + \epsilon m}\,\sqrt{h + \epsilon s p_z}\, e^{-i\phi/2} \\ s\epsilon\sqrt{\varepsilon + \epsilon m}\,\sqrt{h - \epsilon s p_z}\, e^{i\phi/2} \\ s\epsilon\sqrt{\varepsilon - \epsilon m}\,\sqrt{h + \epsilon s p_z}\, e^{-i\phi/2} \\ \sqrt{\varepsilon - \epsilon m}\,\sqrt{h - \epsilon s p_z}\, e^{i\phi/2} \end{pmatrix}. \tag{B.11}$$

One may rewrite the solutions in terms of polar rather then cylindrical coordinates using, cf. (B.3),

$$\frac{\sqrt{h + s p_z}}{\sqrt{2h}} = \cos\tfrac{1}{2}\theta, \qquad \frac{\sqrt{h - s p_z}}{\sqrt{2h}} = \sin\tfrac{1}{2}\theta. \tag{B.12}$$

2. Eigenstates of the Magnetic Moment Operator

In practice electrons are usually polarized by a magnetic field. The polarization states then correspond to eigenfunctions of the component of the magnetic moment operator $\hat{\boldsymbol{\mu}}$ along $\boldsymbol{B}$. It is shown in Appendix C that this operator is given by

$$\hat{\boldsymbol{\mu}} = m\boldsymbol{\Sigma} + \rho_y\, \boldsymbol{\Sigma} \times \hat{\boldsymbol{p}} = \begin{pmatrix} m & 0 & 0 & \epsilon p_- \\ 0 & -m & -\epsilon p_+ & 0 \\ 0 & -\epsilon p_- & m & 0 \\ \epsilon p_+ & 0 & 0 & -m \end{pmatrix}. \tag{B.13}$$

Let the eigenvalues of μ_z may be written as $s\lambda$ with $s = \pm$. The matrix used to determine the eigenvalues in this case is

$$\begin{pmatrix} m - s\lambda & 0 & 0 & \epsilon p_- \\ 0 & -m - s\lambda & -\epsilon p_+ & 0 \\ 0 & -\epsilon p_- & m - s\lambda & 0 \\ \epsilon p_+ & 0 & \epsilon p_+ & -m - s\lambda \end{pmatrix}.$$

The eigenvalues have

$$\lambda = (m^2 + p_\perp^2)^{1/2}. \tag{B.14}$$

As in the case of the helicity eigenfunctions, one requires that the solutions (4.12) be eigenfunctions of μ_z, and this determined C_s^ϵ and D_s^ϵ. Again the ϕ-dependence is satisfied by writing the solution in the form (B.4). In place of (B.5) one requires

$$(m - s\lambda)(\epsilon\varepsilon + m)C_s^\epsilon + \epsilon p_\perp(\epsilon p_\perp C_s^\epsilon - \epsilon p_z D_s^\epsilon) = 0, \tag{B.15a}$$

$$-(m + s\lambda)(\epsilon\varepsilon + m)D_s^\epsilon - \epsilon p_\perp(\epsilon p_z C_s^\epsilon - \epsilon p_\perp D_s^\epsilon) = 0. \tag{B.15b}$$

Using the identities

$$p_z = \sqrt{(\varepsilon + \epsilon s\lambda)(\varepsilon - \epsilon s\lambda)}, \qquad p_\perp = \sqrt{(\lambda + sm)(\lambda - sm)}, \tag{B.16}$$

Specific solutions are

$$C_s^\epsilon = \frac{\epsilon}{\sqrt{2\varepsilon(\varepsilon + \epsilon m)V}} \frac{\sqrt{\varepsilon + \epsilon s\lambda}\,\sqrt{\lambda + sm}}{\sqrt{2\lambda(\varepsilon + \epsilon m)}}, \qquad D_s^\epsilon = \frac{-s}{\sqrt{2\varepsilon(\varepsilon + \epsilon m)V}} \frac{\sqrt{\varepsilon + \epsilon s\lambda}\,\sqrt{\lambda + sm}}{\sqrt{2\lambda(\varepsilon + \epsilon m)}}. \tag{B.17}$$

The resulting eigenfunctions are

$$\varphi_s^\epsilon(\epsilon\boldsymbol{p}) = \frac{1}{\sqrt{2\epsilon\varepsilon\, 2s\lambda V}} \begin{pmatrix} \sqrt{(\varepsilon + \epsilon s\lambda)(\lambda + sm)}\, e^{-i\phi/2} \\ -\epsilon s\sqrt{(\varepsilon - \epsilon s\lambda)(\lambda - sm)}\, e^{i\phi/2} \\ \sqrt{(\varepsilon - \epsilon s\lambda)(\lambda + sm)}\, e^{-i\phi/2} \\ \epsilon s\sqrt{(\varepsilon + \epsilon s\lambda)(\lambda - sm)}\, e^{-i\phi/2} \end{pmatrix}. \tag{B.18}$$

This solutions apply for $p_z = |\boldsymbol{p}|\cos\theta = \sqrt{\varepsilon - \lambda} > 0$, and the obvious changes of sign are to be made for $p_z < 0$.

3. Eigenstates in Cylindrical Coordinates in a Magnetic Field

Suppose that in place of the Landau gauge (6.3) one chooses the gauge (6.4), viz.

$$\boldsymbol{A} = \tfrac{1}{2}(-By, Bx, 0) = \tfrac{1}{2}B\varpi(-\sin\varphi, \cos\varphi, 0), \tag{B.19}$$

with

$$\varpi := (x^2 + y^2)^{1/2} \quad \text{and} \quad x = \varpi\cos\varphi,\ y = \varpi\sin\varphi. \tag{B.20}$$

Then z is the only cyclic coordinate, and in place of (6.5) an appropriate trial wave function is

$$\Psi(t, \boldsymbol{x}) = g(\varpi, \phi)\,\exp(-i\epsilon\varepsilon t + i\epsilon p_z z). \tag{B.21}$$

In place of (6.7), Dirac's equation gives

$$\begin{pmatrix} -\epsilon\varepsilon + m & 0 & \epsilon p_z & \hat{\mathcal{D}}_1 \\ 0 & -\epsilon\varepsilon + m & \hat{\mathcal{D}}_2 & -\epsilon p_z \\ \epsilon p_z & \hat{\mathcal{D}}_1 & -\epsilon\varepsilon - m & 0 \\ \hat{\mathcal{D}}_2 & -\epsilon p_z & 0 & -\epsilon\varepsilon - m \end{pmatrix} \begin{pmatrix} g_1(\varpi, \phi) \\ g_2(\varpi, \phi) \\ g_3(\varpi, \phi) \\ g_4(\varpi, \phi) \end{pmatrix} = 0, \tag{B.22}$$

with

$$\hat{\mathcal{D}}_1 := -ie^{-i\phi}\left(\frac{\partial}{\partial\varpi} - \frac{i}{\varpi}\frac{\partial}{\partial\phi} + \tfrac{1}{2}eB\varpi\right), \tag{B.23a}$$

$$\hat{\mathcal{D}}_2 := -ie^{i\phi}\left(\frac{\partial}{\partial\varpi} + \frac{i}{\varpi}\frac{\partial}{\partial\phi} - \tfrac{1}{2}eB\varpi\right). \tag{B.23b}$$

The dependence on ϕ is satisfied by the choice

$$g_1(\varpi, \phi) = g_1(\varpi)\exp[i(m-1)\phi], \qquad g_2(\varpi, \phi) = g_2(\varpi)\exp[im\phi], \tag{B.24a}$$

$$g_3(\varpi, \phi) = g_3(\varpi)\exp[i(m-1)\phi], \qquad g_4(\varpi, \phi) = g_4(\varpi)\exp[im\phi], \tag{B.24b}$$

with $m = 0, \pm1, \pm2, \ldots$.

In place of (6.10) one then finds

$$(-\epsilon\varepsilon + m)g_1 + \epsilon p_z g_3 - i\left(\frac{d}{d\varpi} + \frac{m}{\varpi} + \tfrac{1}{2}eB\varpi\right)g_4 = 0, \tag{B.25a}$$

$$(-\epsilon\varepsilon + m)g_2 - \epsilon p_z g_4 - i\left(\frac{d}{d\varpi} + \frac{m-1}{\varpi} + \tfrac{1}{2}eB\varpi\right)g_3 = 0, \tag{B.25b}$$

$$(-\epsilon\varepsilon - m)g_3 + \epsilon p_z g_1 - i\left(\frac{d}{d\varpi} + \frac{m}{\varpi} + \tfrac{1}{2}eB\varpi\right)g_2 = 0, \tag{B.25c}$$

$$(-\epsilon\varepsilon - m)g_4 - \epsilon p_z g_2 - i\left(\frac{d}{d\varpi} + \frac{m-1}{\varpi} + \tfrac{1}{2}eB\varpi\right)g_1 = 0. \tag{B.25d}$$

In place of (6.11) one finds

$$\left[\frac{d^2}{d\varpi^2}+\frac{1}{\varpi}\,\frac{d}{d\varpi}-\frac{(m-1)^2}{\varpi^2}+eB(2n-m)-\tfrac{1}{4}e^2B^2\varpi^2\right]f_{1,3}=0, \tag{B.26a}$$

$$\left[\frac{d^2}{d\varpi^2}+\frac{1}{\varpi}\,\frac{d}{d\varpi}-\frac{m^2}{\varpi^2}+eB(2n-m+1)-\tfrac{1}{4}e^2B^2\varpi^2\right]f_{2,4}=0. \tag{B.26b}$$

The normalizable solutions of (B.26) are generalized Laguerre polynomials $L_n^\nu(x)$, specifically the functions

$$J_\nu^n(x):=[n!/(n+\nu)!]^{1/2}\,\exp(-\tfrac{1}{2}x)\,x^{\frac{1}{2}\nu}\,L_n^\nu(x). \tag{B.27}$$

In place of (6.15) one obtains the solutions

$$g(\varpi,\phi)=\begin{pmatrix} C_1J^s_{n-r-1}(\frac{1}{2}eB\varpi^2)\exp[i(n-r-1)\phi] \\ C_2J^s_{n-r}(\frac{1}{2}eB\varpi^2)\exp[i(n-r)\phi] \\ C_3J^s_{n-r-1}(\frac{1}{2}eB\varpi^2)\exp[i(n-r-1)\phi] \\ C_4J^s_{n-r}(\frac{1}{2}eB\varpi^2)\exp[i(n-r)\phi] \end{pmatrix}, \tag{B.28}$$

where

$$r:=n-m \tag{B.29}$$

is referred to as the *radial* quantum number.

The construction of specific spin eigenstates involves only the determination of the coefficients C_1 to C_4 and is the same in all gauges, including the Landau gauge and the "cylindrical" gauge chosen here.

4. Helicity Eigenstates in a Magnetic Field

The helicity operator in a magnetic field is $\boldsymbol{\sigma}\cdot(\hat{\boldsymbol{p}}+e\boldsymbol{A})$. Evaluating this in the Landau gauge gives

$$\boldsymbol{\sigma}\cdot[\hat{\boldsymbol{p}}+e\boldsymbol{A}]=\begin{pmatrix} \epsilon p_z & \hat{\mathcal{X}}_+ & 0 & 0 \\ \hat{\mathcal{X}}_- & -\epsilon p_z & 0 & 0 \\ 0 & 0 & \epsilon p_z & \hat{\mathcal{X}}_+ \\ 0 & 0 & \hat{\mathcal{X}}_- & -\epsilon p_z \end{pmatrix}, \tag{B.30}$$

with

$$\hat{\mathcal{X}}_\pm=-ieB\left(\frac{\partial}{\partial\xi}\pm\xi\right). \tag{B.31}$$

Then operating on the solutions (6.15) one seeks eigenfunctions of the form

$$\boldsymbol{\sigma}\cdot[\hat{\boldsymbol{p}}+e\boldsymbol{A}]\Psi^\epsilon_{qs}(t,\boldsymbol{x})=sh\Psi^\epsilon_{qs}(t,\boldsymbol{x}), \tag{B.32}$$

with $s=\pm$, and h to be determined. One finds the eigenvalue equation

$$\begin{pmatrix} -\epsilon p_z-sh & -ip_n & 0 & 0 \\ -ip_n & -\epsilon p_z-sh & 0 & 0 \\ 0 & 0 & \epsilon p_z-sh & -ip_n \\ 0 & 0 & -ip_n & -\epsilon p_z-sh \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix}=0. \tag{B.33}$$

The eigenvalues are doubly degenerate and have

$$h = (p_n^2 + p_z^2)^{1/2} = (\varepsilon_q^2 - m^2)^{1/2}. \tag{B.34}$$

If the eigenvalues were not degenerate, the eigenfunctions could be constructed from the inverse of the square matrix in (B.33). However, the inverse is singular here, and this procedure cannot be used.

One way to proceed is to use the solutions (6.20) already obtained. The required solutions must be linear combinations of the solutions (6.20). Hence we may write

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} = \sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V} \left[\begin{pmatrix} (\epsilon\varepsilon_q + m)v_{n-1}(\xi) \\ 0 \\ \epsilon p_z v_{n-1}(\xi) \\ ip_n v_n(\xi) \end{pmatrix} D_1 + \begin{pmatrix} 0 \\ (\epsilon\varepsilon_q + m)v_n(\xi) \\ -ip_n v_{n-1}(\xi) \\ -\epsilon p_z v_n(\xi) \end{pmatrix} D_2 \right], \tag{B.35}$$

with

$$|D_1|^2 + |D_2|^2 = 1, \quad \text{and} \quad \sum_{i=1}^{4} |C_i|^2 = 1. \tag{B.36}$$

On operating with (B.30) one requires

$$(\epsilon p_z - sh)D_1 - ip_n D_2 = 0, \tag{B.37a}$$

$$(\epsilon p_z - sh)(\epsilon p_z D_1 - ip_n D_2) - ip_n(ip_n D_1 - \epsilon p_z D_2) = 0. \tag{B.37b}$$

The solution is determined only to within an arbitrary phase factor for each eigenfunction. One choice of phase factors corresponds to

$$D_1 = \left(\frac{h + s\epsilon p_z}{2h}\right)^{1/2}, \qquad D_2 = is\left(\frac{h - s\epsilon p_z}{2h}\right)^{1/2}. \tag{B.38}$$

Using the identity

$$p_n = \sqrt{(h + \epsilon s p_z)(h - \epsilon s p_z)}, \tag{B.39}$$

one may write the final result in the form

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} = \frac{1}{\sqrt{2h2\varepsilon V}} \begin{pmatrix} \sqrt{\varepsilon + \epsilon m}\,\sqrt{h + \epsilon s p_z} \\ -is\sqrt{\varepsilon - \epsilon m}\,\sqrt{h - \epsilon s p_z} \\ s\sqrt{\varepsilon - \epsilon m}\,\sqrt{h + \epsilon s p_z} \\ i\sqrt{\varepsilon + \epsilon m}\,\sqrt{h - \epsilon s p_z} \end{pmatrix}, \tag{B.40}$$

with $\varepsilon = \varepsilon_q$ given by (6.18). As already remarked, this result also applies to the cylindrical coordinate case discussed above.

5. Magnetic Moment Eigenstates in a Magnetic Field

The magnetic moment operator in the presence of a magnetic field is

$$\hat{\boldsymbol{\mu}} = m\boldsymbol{\Sigma} + \rho_y \boldsymbol{\Sigma} \times (\hat{\boldsymbol{p}} + e\boldsymbol{A}). \tag{B.41}$$

The z-component of this operator in the Landau gauge has the matrix representation

$$\mu_z = \begin{pmatrix} m & 0 & 0 & \hat{\mathcal{X}}_+ \\ 0 & -m & \hat{\mathcal{X}}_- & 0 \\ 0 & -\hat{\mathcal{X}}_+ & m & 0 \\ -\hat{\mathcal{X}}_- & 0 & 0 & -m \end{pmatrix}, \tag{B.42}$$

with $\hat{\mathcal{X}}_\pm$ defined by (B.31). We seek eigenfunctions of the operator (B.42).

Let the eigenvalues be $s\lambda$. One finds

$$\lambda = (m^2 + p_n^2)^{1/2} = (\varepsilon_q^2 - p_z^2)^{1/2}. \tag{B.43}$$

Using the identities

$$p_z^2 = (\varepsilon + \epsilon s\lambda)(\varepsilon - \epsilon s\lambda), \qquad p_n^2 = (\lambda + sm)(\lambda - sm), \tag{B.44}$$

one finds

$$D_1 = \left(\frac{(\varepsilon + \epsilon s\lambda)(\lambda + sm)}{2\lambda(\varepsilon + \epsilon m)} \right)^{1/2}, \qquad D_2 = -is \left(\frac{(\varepsilon - \epsilon s\lambda)(\lambda - sm)}{2\lambda(\varepsilon + \epsilon m)} \right)^{1/2}. \tag{B.45}$$

The solution is

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} = \frac{1}{2\lambda 2\varepsilon_q} \begin{pmatrix} \sqrt{\varepsilon + \epsilon s\lambda}\,\sqrt{\lambda + sm} \\ -is\sqrt{\varepsilon + \epsilon s\lambda}\,\sqrt{\lambda - sm} \\ s\sqrt{\varepsilon - \epsilon s\lambda}\,\sqrt{\lambda + sm} \\ i\sqrt{\varepsilon - \epsilon s\lambda}\,\sqrt{\lambda - sm} \end{pmatrix}. \tag{B.46}$$

Appendix C

The Poincaré Group and Spin Operators

It is desirable that a spin operator be chosen that (a) commutes with the Hamiltonian so that it is a constant of the motion, and (b) is a component of a 4-vector or 4-tensor. In this Appendix a general procedure for constructing such spin operators is outlined, and the specific operators used in Appendix A are identified. As a prelude the formal aspects of Lorentz transformations, generalized to the Poincaré group, are discussed.

1. Generators of the Poincaré Group

Lorentz transformation may be regarded as rotations in 4-dimensional space-time, and when an arbitrary translation of the space-time origin is added, one has the *Poincaré group*. There are ten generators for this group, four representing the translations and six representing the rotations. The six rotations can be further separated into three representing rotations of the coordinate axes, and three boosts.

In the case of an infinitesimal transformation, cf. (2.11), $x^{\mu'}$ and x^μ differ by an infinitesimal amount that is denoted dx^μ. Note however the awkwardness in notation: dx^μ is not the difference between two 4-vectors, but the difference between the same 4-vector in two frame, and thus dx^μ is not itself a 4-vector. An infinitesimal transformation can be described by the 4-vector da_α representing the displacement of the origins of the two frames, and an antisymmetric tensor $d\omega_{\alpha\beta}$ representing the rotation. Then one may write

$$dx^\mu = id\omega_{\alpha\beta}(M^{\alpha\beta})^\mu{}_\nu x^\nu + ida_\alpha(P^\alpha)^\mu, \tag{C.1}$$

where M and P are the generators of the transformations and where the inclusion of the factor i is conventional. The six components $\alpha, \beta = 0, 1, 2, 3$ of $(M^{\alpha\beta})^{\mu\nu} = -(M^{\beta\alpha})^{\mu\nu}$ and the four components $\alpha = 0, 1, 2, 3$ of $(P^\alpha)^\mu$ are the ten generators of the Poincaré group. The other indices μ and ν relate the components of a 4-vector in the new frame to those in the old frame. The explicit forms for the generators are readily identified. They are

$$(M^{\alpha\beta})^{\mu\nu} = -i\Big(g^{\alpha\mu}g^{\beta\nu} - g^{\alpha\nu}g^{\beta\mu}\Big), \qquad (P^\alpha)^\mu = -i\,g^{\alpha\mu}. \tag{C.2}$$

These generators are more easily understood in terms of their forms in the coordinate representation. Consider the Klein Gordon wave function, which is an invariant so that one has $\Psi'(x') = \Psi(x)$. From Nöther's theorem one knows that the generator of a translation is related to the relevant component of the momentum, and that the generator of a rotation is related to the relevant component of the angular momentum. Thus by considering infinitesimal translations and rotations in space-time one can identify the 4-momentum operator $\hat{p}^\mu$ and the 4-dimensional angular momentum operator $\hat{J}^{\mu\nu}$ respectively.

To be more specific, in quantum mechanics Nöther's theorem leads to the identification of the relevant operators as follow. The transformation matrix $S(L)$ must be unitary, and a unitary operator $\hat{U} = \hat{U}^*$ that is written $\hat{U} = \exp[i\hat{u}]$ defines an hermitian operator $\hat{u}$. The relevant momentum operator $\hat{p}^\alpha$ and angular momentum operator $\hat{J}^{\alpha\beta}$ are most

easily identified in terms of infinitesimal transformations, when one has $\hat{u} = da_\alpha \hat{p}^\alpha$ or $\hat{u} = d\omega_{\alpha\beta}\hat{J}^{\alpha\beta}$.

Under an infinitesimal translation, $x' = x - da$, $\Psi'(x)$ differs from $\Psi(x)$ by an infinitesimal amount $d\Psi(x)$ given by

$$d\Psi(x) = (1 - da^\alpha \partial_\alpha)\Psi(x) = (1 + ida_\alpha \hat{p}^\alpha)\Psi(x), \tag{C.3}$$

leading to the identification

$$\hat{p}^\alpha = i(P^\alpha)^\mu \partial_\mu = i\partial^\alpha. \tag{C.4}$$

For a 4-dimensional rotation (C.1) the total angular momentum may be separated into orbital and spin parts:

$$\hat{J}^{\mu\nu} = \hat{M}^{\mu\nu} + \hat{S}^{\mu\nu}, \tag{C.5}$$

For the spin 0 case the spin part is zero, and hence the total angular momentum is equal to the orbital part, which is thereby identified as

$$\hat{M}^{\mu\nu} = i\Big(x^\mu \partial^\nu - x^\nu \partial^\mu\Big). \tag{C.6}$$

For the Dirac field, the spin part may be identified by considering the transformation of the γ matrices. In this case the orbital part does not contribute. One requires

$$\Big(g^{\mu\nu} - id\omega_{\alpha\beta}(M^{\alpha\beta})^{\mu\nu}\Big)\gamma_\nu = \Big(1 - id\omega_{\alpha\beta}\hat{S}^{\alpha\beta}\Big)\gamma^\mu\Big(1 + id\omega_{\alpha\beta}\hat{S}^{\alpha\beta}\Big), \tag{C.7}$$

where $\hat{J}^{\alpha\beta}$ has been replaced by $\hat{S}^{\alpha\beta}$ on the right hand side. It follows that $\hat{S}^{\alpha\beta}$ satisfies the commutation relation

$$[\gamma^\mu, \hat{S}^{\alpha\beta}] = -(M^{\alpha\beta})^{\mu\nu}\gamma_\nu. \tag{C.8}$$

With the explicit form (C.2) for the generator, the solution of (C.8) is

$$\hat{S}^{\alpha\beta} = \tfrac{i}{4}\,[\gamma^\alpha, \gamma^\beta], \tag{C.9}$$

as anticipated in (9.5).

2. Invariants of the Poincaré Group

The generators of the Poincaré group form a Lie algebra. It is convenient to write the algebra in terms of the momentum and angular momentum operators $\hat{p}^\alpha$ and $\hat{J}^{\mu\nu}$. These satisfy

$$[\hat{J}^{\mu\nu}, \hat{J}^{\rho\sigma}] = -i(g^{\mu\rho}\hat{J}^{\nu\sigma} + g^{\nu\sigma}\hat{J}^{\mu\rho} - g^{\nu\rho}\hat{J}^{\mu\sigma} - g^{\mu\sigma}\hat{J}^{\nu\rho}) \tag{C.10}$$

$$[\hat{J}^{\mu\nu}, \hat{p}^\alpha] = i(g^{\nu\alpha}\hat{p}^\mu - g^{\mu\alpha}\hat{p}^\nu) \tag{C.11}$$

$$[\hat{p}^\alpha, \hat{p}^\beta] = 0 \tag{C.12}$$

It is straighforward to show that $\hat{p}^2$ commutes with all ten generators. It follows that $\hat{p}^2$ is an invariant of the group. However, this is already obvious in that the wave equations imply $\hat{p}^2 = m^2$. There is another invariant that involves the spin and is not so obvious.

Consider the operator

$$\hat{W}^\mu := -\tfrac{1}{2}\epsilon^{\mu\alpha\beta\gamma}\hat{J}_{\alpha\beta}\hat{p}_\gamma. \tag{C.13}$$

One finds

$$[\hat{J}^{\mu\nu}, \hat{W}^\alpha] = -i(g^{\nu\alpha}\hat{W}^\mu - g^{\mu\alpha}\hat{W}^\nu) \tag{C.14}$$

$$[\hat{p}^\alpha, \hat{W}^\beta] = 0, \tag{C.15}$$

$$[\hat{W}^\alpha, \hat{W}^\beta] = -i\epsilon^{\alpha\beta\gamma\delta}\hat{W}_\gamma\hat{p}_\delta. \tag{C.16}$$

It follows that $\hat{W}^2$ also commutes with all ten generators. In the rest frame on has $\hat{W}^3$ is numerically equal to $-m\hat{J}_{12}$ plus cyclic permutations. It follows that in this frame one has $\hat{W}^2 = -m^2|\hat{\boldsymbol{J}}|^2 = -m^2S(S+1)$, where S denotes the spin ($S = 0, \frac{1}{2}, 1, \ldots$).

For the Dirac theory one has $S = \frac{1}{2}$. A formal proof involves constructing $\hat{W}$ explicitly in the rest frame in the standard representation. One finds $\hat{\boldsymbol{W}} = -\frac{1}{2}m\boldsymbol{\sigma}$, confirming that S is indeed equal to $\frac{1}{2}$.

3. Conserved Quantities and Constants of the Motion

The spin operators $\hat{S}^{\mu\nu}$ and $\hat{W}^\mu$ satisfy one for the desirable properties for acceptable spin operators in that one is a 4-tensor and the other is a 4-vector. However they do not necessarily commute with the Hamiltonian, and hence their eigenvalues are not necessarily constants of the motion.

Consider an arbitrary operator $\hat{F}^{\alpha\beta\cdots}$ and the corresponding density

$$F^{\alpha\beta\cdots}(x) = \Psi^\dagger\hat{F}^{\alpha\beta\cdots}\,\Psi(x). \tag{C.17}$$

For example, the energy momentum tensor may be written

$$T^{\mu\nu}(x) = \Psi^\dagger\,\hat{T}^{\mu\nu}\,\Psi(x), \tag{C.18}$$

with

$$\hat{T}^{\mu\nu} = \gamma^0\Big(\gamma^\mu\hat{p}^\nu\Big). \tag{C.19}$$

The continuity equation $\partial_\mu T^{\mu\nu}(x) = 0$ implies a conservation law. Thus the 4-momentum in the field

$$P^\nu(x) := T^{0\nu}(x) = \Psi^\dagger\,\hat{p}^\nu\,\Psi(x) \tag{C.20}$$

is a constant. The operator $\hat{p}^\nu$ commutes with the Hamiltonian and hence is a constant of the motion.

From any given operator $\hat{F}^{\alpha\beta\cdots}$ one may construct a conserved quantity and a corresponding constant of the motion as follows. Consider

$$f^{\mu\alpha\beta\cdots}(x) := \frac{-i}{2m}\left[\Big(\partial^\mu\Psi^\dagger(x)\Big)\hat{F}^{\alpha\beta\cdots}\,\Psi(x) - \Psi^\dagger(x)\hat{F}^{\alpha\beta\cdots}\Big(\partial^\mu\Psi(x)\Big)\right]. \tag{C.21}$$

The Klein Gordon equation or the Dirac equation, and their adjoints, imply $\partial^\mu\partial_\mu\Psi(x) = -m^2\Psi(x)$ and $\partial^\mu\partial_\mu\Psi^\dagger(x) = -m^2\Psi^\dagger(x)$. These then imply

$$\partial_\mu f^{\mu\alpha\beta\cdots}(x) = 0. \tag{C.22}$$

It follows that $f^{0\alpha\beta\cdots}(x)$ is a conserved quantity. Moreover, suppose one writes (C.21) in the form

$$f^{\mu\alpha\beta\cdots}(x) = \Psi^\dagger(x)\hat{f}^{\mu\alpha\beta\cdots}\,\Psi(x), \tag{C.23}$$

with

$$\hat{f}^{\mu\alpha\beta\cdots} = \frac{1}{2m}\Big[\hat{p}^\mu\hat{F}^{\alpha\beta\cdots} + \hat{F}^{\alpha\beta\cdots}\hat{p}^\mu\Big]. \tag{C.24}$$

Then one expects the $\mu = 0$ component of this operator to be a constant of the motion.

Thus for any arbitrary tensorial operator $\hat{F}^{\alpha\beta\cdots}$ one may construct another related tensorial operator

$$\hat{\tilde{F}}^{\alpha\beta\cdots} := \frac{1}{2m}\Big[\hat{p}^0\hat{F}^{\alpha\beta\cdots} + \hat{F}^{\alpha\beta\cdots}\hat{p}^0\Big] \tag{C.25}$$

that is a constant of the motion. The corresponding density

$$\tilde{F}^{\alpha\beta\cdots}(x) = \Psi^\dagger(x)\,\hat{\tilde{F}}^{\alpha\beta\cdots}\,\Psi(x) \tag{C.26}$$

is a conserved quantity.

In the case of the Dirac field for the density to be of the form (C.17), one needs to take the factor γ^0 in the Dirac adjoint $\overline{\Psi}(x) = \Psi^\dagger(x)\gamma^0$ into account. This implies that any identified $\hat{F}^{\mu\alpha\beta\cdots}$ must have a factor γ^0 premultiplying it.

4. Spin Operators that Commute with the Hamiltonian

The two spin operators identified so far are $\hat{S}^{\mu\nu}$ and $\hat{W}^\mu$. These are defined in (9.5) and (C.13) respectively. The corresponding constants of the motion are

$$\hat{\tilde{S}}^{\mu\nu} = \frac{1}{2m}\Big[\hat{H}\gamma^0\hat{S}^{\mu\nu} + \gamma^0\hat{S}^{\mu\nu}\hat{H}\Big] \tag{C.27}$$

and

$$\hat{\tilde{W}}^{\mu} = \frac{1}{2m}\Big[\hat{H}\gamma^0\hat{W}^{\mu} + \gamma^0\hat{W}^{\mu}\hat{H}\Big] \tag{C.28}$$

respectively, where $\hat{p}^0$ is now identified as the Hamiltonian $\hat{H}$.

The explicit forms of these operators in the standard representation correspond to the replacements

$$\gamma^0 \hat{S}^{01} \to -\tfrac{1}{2}\rho_y\sigma_x \qquad \gamma^0 \hat{S}^{12} \to \tfrac{1}{2}\rho_z\sigma_z,$$

plus cyclic permutations, where "$\to$" denotes a numerical equality between quantities in different notations. The Dirac Hamiltonian is $\hat{H} = \rho_y \boldsymbol{\sigma}\cdot\boldsymbol{p} + \rho_z\, m$. The tensor $\hat{\tilde{S}}^{\mu\nu}$ includes the electric moment operator $\hat{\boldsymbol{d}}$ and the magnetic moment operator $\hat{\boldsymbol{\mu}}$. One has

$$\hat{\tilde{S}}^{\mu\nu} = \frac{1}{2m}\begin{pmatrix} 0 & -\hat{d}_x & -\hat{d}_y & -\hat{d}_z \\ \hat{d}_x & 0 & \hat{\mu}_z & -\hat{\mu}_y \\ \hat{d}_y & -\hat{\mu}_z & 0 & \hat{\mu}_x \\ \hat{d}_z & \hat{\mu}_y & -\hat{\mu}_x & 0 \end{pmatrix}, \tag{C.29}$$

with

$$\hat{\boldsymbol{d}} = -\rho_z\, \boldsymbol{\sigma} \times \hat{\boldsymbol{p}}, \tag{C.30}$$

and

$$\hat{\boldsymbol{\mu}} = m\boldsymbol{\sigma} + \rho_y\, \boldsymbol{\sigma} \times \hat{\boldsymbol{p}}. \tag{C.31}$$

For the 4-vector operator $\hat{\tilde{W}}^{\mu}$ I find the following explicit form

$$\hat{\tilde{W}}^{\mu} = [\tfrac{1}{2}\boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}}, \varepsilon\boldsymbol{\sigma} - \varepsilon\rho_x\boldsymbol{\sigma}\times\hat{\boldsymbol{p}}/2m]. \tag{C.32}$$

The spin operators used in Appendix B in writing down spin eigenfunctions are the helicity operator $\boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}}$, that comes from the 0-component of (C.32), and $\hat{\mu}_z$, that comes from the 12-component of (C.29). By construction these are constants of the motion and are known components of a 4-vector and a 4-tensor respectively. Clearly there are several other possible choices. One notable one is the operator $\hat{d}_3$; its eigenfunctions would correspond to the spin states obtained by polarizing the electrons using an electric field directed along the 3-axis.

Appendix D

Wave Dispersion

The dispersive properties of a plasma (or other medium) may be described in terms of a response tensor, which is introduced here in terms of a relation between the induced 4-current $J^\mu(k)$ and the 4-potential $A_\nu(k)$ in k-space. This relation defines the linear response tensor $\alpha^{\mu\nu}(k)$ for the plasma. The properties of waves in the plasma may be derived from the wave equation which includes this response tensor. The following is an outline of a general procedure for deriving the wave properties.

1. The Linear Response Tensor

The formal introduction of the response tensor involves starting from the Fourier transformed form of Maxwell's equation (2.18) written in terms of the 4-potential using (2.31), which gives

$$F^{\mu\nu}(k) = i\,[k^\mu A^\nu(k) - k^\nu A^\mu(k)]. \tag{D.1}$$

Then (2.18) reduces to

$$[k^2\, g^{\mu\nu} - k^\mu k^\nu]\, A_\nu(k) = -\mu_0\, J^\mu(k). \tag{D.2}$$

The medium is introduced by separating $J(k)$ into an induced (ind) part that describes the response of the medium and an extraneous (ext) part that acts as a source term:

$$J^\mu(k) = J^\mu_{\rm ind}(k) + J^\mu_{\rm ext}(k). \tag{D.3}$$

The induced current is then written in the form

$$J^\mu_{\rm ind}(k) = \alpha^\mu{}_\nu(k)\, A^\nu(k). \tag{D.4}$$

The response tensor necessarily satisfies several conditions. One is the *reality condition*

$$\alpha^{\mu\nu}(k) = [\alpha^{\mu\nu}(-k)]^* \tag{D.5}$$

A second condition follows from the charge continuity relation $k_\mu J^\mu(k) = 0$. This requires that the response tensor satisfy the first of the following relations:

$$k_\mu\, \alpha^{\mu\nu}(k) = 0, \qquad k_\nu\, \alpha^{\mu\nu}(k) = 0. \tag{D.6}$$

The second of these is imposed to ensure that the response is independent of the choice of gauge. A gauge transformation is of the form

$$A'^\mu(k) = A^\mu(k) + i\, k^\mu\, \psi(k), \tag{D.7}$$

and (D.6) ensures that the induced current is independent of the value of the arbitrary function $\psi(k)$. The other general relations that the linear response tensor must satisfy are

the Onsager relations, and the causal (or Kramers Kronig) relations; these are important but are not discussed here.

Another general description of the response of the medium gives the polarization $\boldsymbol{P}$ and the magnetization $\boldsymbol{M}$ as functions of the electric field $\boldsymbol{E}$ and the magnetic induction $\boldsymbol{B}$. The most general form of this description of the response (for an anisotropic, dispersive and spatially dispersive medium with magneto-electric responses) involves relations of the form

$$\begin{aligned} \boldsymbol{P} &= \boldsymbol{\chi}^e \cdot \boldsymbol{E} + \boldsymbol{\chi}^{em} \cdot \boldsymbol{B}, \\ \boldsymbol{M} &= \boldsymbol{\chi}^{me} \cdot \boldsymbol{E} + \boldsymbol{\chi}^m \cdot \boldsymbol{B}, \end{aligned} \tag{D.8}$$

where all quantities are function of ω and $\boldsymbol{k}$, where $\boldsymbol{\chi}^e$ and $\boldsymbol{\chi}^m$ are the electric and magnetic susceptibility tensors, respectively, and where the other two tensors describe magneto-electric responses (which are often ignored). The complicated description (D.8) is no more general than (D.4). Indeed the two are equivalent, and one may express the components of the tensors in (D.8) in terms of the components of $\alpha^{\mu\nu}(k)$, and vice versa.

2. The Wave Equation

The relation obtained from Maxwell's equations then reduces to the *wave equation*

$$\Lambda^{\mu\nu}(k) A_\nu(k) = -\mu_0 {J_{\text{ext}}}^\mu(k), \tag{D.9}$$

with

$$\Lambda^{\mu\nu}(k) = k^2 g^{\mu\nu} - k^\mu k^\nu + \mu_0\, \alpha^{\mu\nu}(k). \tag{D.10}$$

The extraneous current in (D.9) is a source term. The homogeneous wave equation is obtained from (D.9) by omitting this term:

$$\Lambda^{\mu\nu}(k) A_\nu(k) = 0. \tag{D.11}$$

The homogeneous wave equation (D.11) may be regarded as four simultaneous linear algebraic equations for the four components of $A^\mu(k)$. The condition for a solution of a set of simultaneous linear equation to exist is that the determinant of the coefficients vanish. However, the determinant of $\Lambda^{\mu\nu}(k)$ vanishes identically because it satisfies the relations (D.6), viz.

$$k_\mu\, \Lambda^{\mu\nu}(k) = 0, \qquad k_\nu\, \Lambda^{\mu\nu}(k) = 0. \tag{D.12}$$

It follows that $A^\mu(k) \propto k^\mu$ is a solution, and this merely reflects the gauge independence of the theory, cf. (D.7). To find a nontrivial solution some other procedure is required.

A general method is to construct the matrix of cofactors of $\Lambda^{\mu\nu}(k)$. Let $\lambda^{\mu\nu}(k)$ be the (signed) cofactor of $\Lambda^{\nu\mu}(k)$; note the inversion of the indices. The fact that k^μ is always a solution of the wave equation implies that $\lambda^{\mu\nu}(k)$ must be proportional to $k^\mu k^\nu$. Thus if one writes

$$\lambda^{\mu\nu}(k) = k^\mu k^\nu\, \lambda(k), \tag{D.13}$$

then it follows that nontrivial solutions satisfy

$$\lambda(k) = 0. \tag{D.14}$$

The condition (D.14) is a covariant form of the *dispersion equation.*

One may solve (D.14) for ω as a function of $\boldsymbol{k}$. The solution may be written

$$\omega = \omega_M(\boldsymbol{k}) = -\omega_M(-\boldsymbol{k}), \tag{D.15}$$

where the negative-frequency solution is included explicitly. Then one has

$$k_M = [\omega_M(\boldsymbol{k}), \boldsymbol{k}]. \tag{D.16}$$

(Sometimes confusion can arise concerning the transformation $k \to -k$. As in (D.15) it can correspond to a change from a positive-frequency branch of a solution of the dispersion equation to a separate, although related, negative-frequency branch. However it may also correspond to a change from forward to backward propagating waves. When any possible confusion might arise, it may be avoided by regarding forward and backward propagating waves as defining two separate wave modes.)

Amongst the many alternative ways of writing a dispersion relation, that in terms of the refractive index is the most favored. In place of (D.15) one has

$$\mu := |\boldsymbol{k}|/\omega = \mu_M(\omega, \boldsymbol{k}), \tag{D.17}$$

where

$$\boldsymbol{\kappa} := \boldsymbol{k}/|\boldsymbol{k}| \tag{D.18}$$

is a unit vector along the direction of the wavevector. With this form for the dispersion relation one has $k_M = [\omega, \omega\mu_M\boldsymbol{\kappa}]$.

3. The Polarization 4-vector

Given a solution of the dispersion equation, a solution of the homogeneous wave equation (D.11) exists. Let $A_M(k_M)$ denote a solution. The phase, magnitude and gauge of $A_M(k_M)$ remain arbitrary. The general form of $A_M(k_M)$ may be obtained from the second order cofactors $\lambda^{\mu\alpha\beta\gamma}(k_M)$ evaluated at $k = k_M$. Algebraically this procedure appears to be quite complicated. However, apart from notation, it is nothing more than the standard procedure for solving a set of coupled equations by inverting the matrix of coefficients and identifying the solution with any of the columns of the inverted matrix. This procedure gives

$$\begin{aligned}\lambda^{\mu\nu\alpha\beta}(k_M) \propto [&A_M{}^{\mu}(k_M)k_M{}^{\nu} - A_M{}^{\nu}(k_M)k_M{}^{\mu}] \\ &[A_M{}^{\alpha}(k_M)k_M{}^{\beta} - A_M{}^{\beta}(k_M)k_M{}^{\alpha}]^*.\end{aligned} \tag{D.19}$$

To fix the normalization of the polarization 4-vector the temporal gauge is chosen. In the temporal gauge one has $A_M{}^0(k_M) = 0$, and hence $[A_M(k_M)]^2 = -|A_M(k_M)|^2$ is strictly negative.

The definition of the polarization 4-vector includes the following stages.

1) A particular inertial frame is chosen to solve the dispersion equation. This frame is usually the rest frame of the medium.

2) In this frame the temporal gauge is chosen, and the polarization 4-vector is written ${e_M}^{\mu}(k)$, with $k = k_M$. With the dispersion relation in the form $\omega = \omega_M(\boldsymbol{k})$ in this frame, one may write ${e_M}^{\mu}(k) = [0, \boldsymbol{e}_M(\boldsymbol{k})]$, where $\boldsymbol{e}_M(\boldsymbol{k})$ is the polarization 3-vector. The normalization $\boldsymbol{e}_M(\boldsymbol{k}) \cdot {\boldsymbol{e}_M}^*(\boldsymbol{k}) = 1$ then implies

$$[e_M(k)]^{\mu}[{e_M}^*(k)]_{\mu} = -1. \tag{D.20}$$

3) The polarization in any other gauge is constructed by applying a gauge transformation to $e_M(k)$ in the temporal gauge without any change in normalization.

With this choice of normalization, The constant of proportionality in (D.19) is determined:

$$\begin{aligned}\lambda^{\mu\nu\alpha\beta}(k_M) = -\frac{{\lambda^{0\sigma}}_{0\sigma}(k_M)}{[\omega_M(\boldsymbol{k})]^2}\,[{e_M}^{\mu}(k_M){k_M}^{\nu} - {e_M}^{\nu}(k_M){k_M}^{\mu}] \\ [{e_M}^{\alpha}(k_M){k_M}^{\beta} - {e_M}^{\beta}(k_M){k_M}^{\alpha}]^*.\end{aligned} \tag{D.21}$$

4. Ratio of Electric to Total Energy

The parameter ${\lambda^{0\sigma}}_{0\sigma}(k_M)$ in (D.21) may be interpreted in terms of the ratio $R_M(k)$ of electric to total energy in the waves. To see this one needs first to construct the electric energy in the waves, and then to identify an equation that may be interpreted as expressing continuity of wave energy, thereby allowing identification of the total wave energy. The ratio is found by comparing these two quantities.

To construct the electric energy one must first define the wave amplitude. In a particular frame and in the temporal gauge, the solution ${A_M}^{\mu}(k_M)$ for waves in a wave mode M may be written in the form

$$\begin{aligned}{A_M}^{\mu}(k_M) =& a_M(\boldsymbol{k})\,{e_M}^{\mu}(k)\,2\pi\,\delta(\omega - \omega_M(\boldsymbol{k}))+ \\ & a_M(-\boldsymbol{k})\,{e_M}^{\mu}(-k)\,2\pi\,\delta(\omega + |\omega_M(-\boldsymbol{k})|),\end{aligned} \tag{D.22}$$

which defines the amplitude $a_M(\boldsymbol{k})$.

The electrical energy density in waves may be evaluated by averaging the energy density $\frac{1}{2}\,\varepsilon_0|E(x)|^2$ over all space and time. For this purpose it is important to truncate the integrals to a finite volume V and a finite time T, which are then allowed to approach infinity. Using the temporal gauge, the power theorem implies

$$\int d^4x\,\tfrac{1}{2}\,\varepsilon_0|E(x)|^2 = \int \frac{d^4k}{(2\pi)^4}\,\tfrac{1}{2}\,\varepsilon_0|\omega A(k)|^2 \tag{D.23}$$

The average electrical energy density in waves in the mode M is then given by

$$\frac{1}{VT}\int \frac{d^4k}{(2\pi)^4}\,\tfrac{1}{2}\,\varepsilon_0|\omega A_M(k_M)|^2 = \frac{1}{V}\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\,\varepsilon_0|\omega_M(\boldsymbol{k})a_M(\boldsymbol{k})|^2, \tag{D.24}$$

where the positive and negative frequency solutions contribute equally to cancel the factor $\frac{1}{2}$, and where the square of the δ-functions is evaluated using (A.10). Equation (D.24) effectively identifies the electric energy.

To identify the total energy in the waves let us use a Lagrangian approach. The Lagrangian must be constructed such that the Euler-Lagrange equation give the wave equation and its hermitian conjugate. This determines $\mathcal{L}$ to within a normalization factor that may be fixed by applying the theory to a vacuum and ensuring that the energy-momentum tensor reproduces the known value. A suitable Lagrangian is in momentum space and is a the function of the amplitude and its complex conjugate, expressed as 4-vectors incorporating the polarization 4-vector, i.e., $a_{M\mu} := a_M\, e_{M\mu}$ and ${a_M}^{*\mu}$, and of the phase Ψ_M of the wave. The Lagrangian does not involve the derivatives of the the amplitudes explicitly. Also for random phase waves, the phase itself does not appear, but its derivative does, with ${k_M}^\mu = \partial^\mu \Psi_M$. (This formalism is a covariant form of the *averaged-Lagrangian* approach introduced by Whitham*.)

A suitable Lagrangian is

$$\mathcal{L}(a_M, a_M*, \partial\Psi_M) = -\varepsilon_0 {a_M}^*_\mu a_{M\nu} \Lambda^{\mu\nu}(\partial\Psi_M). \tag{D.25}$$

The Euler-Lagrange equations associated with the "generalized coordinates" ${a_M}^*_\mu$ and $a_{M\nu}$ give the wave equation and its hermitian conjugate respectively. The derivatives of these quantities do not appear in the Lagrangian (D.25), which is henceforth denoted by $\mathcal{L}_M$ for simplicity. Thus the Euler-Lagrange equations give

$$-\frac{\partial \mathcal{L}_M}{\partial {a_M}^*_\mu} = \varepsilon_0\, \Lambda^{H\mu\nu}(k_M) a_{M\nu} = 0, \qquad -\frac{\partial \mathcal{L}_M}{\partial a_{M\mu}} = \varepsilon_0\, {a_M}^*_\nu \Lambda^{H\mu\nu}(k_M) = 0. \tag{D.26}$$

The phase itself does not appear in $\mathcal{L}_M$, and so the remaining Euler-Lagrange equation reduces to

$$\partial_\mu \left(\frac{\partial \mathcal{L}_M}{\partial(\partial_\mu \Psi_M)} \right) = 0. \tag{D.27}$$

Equation (D.27) is of the form of a continuity equation. It is a continuity equation for the wave action. One has

$$\frac{\partial \mathcal{L}_M}{\partial(\partial_\mu \Psi_M)} = N_M(\boldsymbol{k}) {v_{gM}}^\mu(\boldsymbol{k}) \tag{D.28}$$

where

$$N_M(\boldsymbol{k}) = \frac{\partial \mathcal{L}_M}{\partial \omega} = -\varepsilon_0\, |a_M|^2 \left[\frac{\partial \mathcal{L}_M(k)}{\partial \omega} \right]_{\omega\, =\, \omega_M(\boldsymbol{k})} \tag{D.29}$$

is the *wave action*, and where

$$v_{gM}(\boldsymbol{k}) := \frac{\partial {k_M}^\mu}{\partial \omega} = [1, \boldsymbol{v}_{gM}(\boldsymbol{k})] \tag{D.30}$$

* G.B. Whitham, *J. Fluid Mech.* **22**, 273 (1965).

is related to the group velocity

$$\boldsymbol{v}_{gM}(\boldsymbol{k}) := \frac{\partial \omega_M(\boldsymbol{k})}{\partial \boldsymbol{k}}. \tag{D.31}$$

(Note that $v_{gM}(\boldsymbol{k})$ is *not* a 4-vector.) The wave action (D.29) may be interpreted as the occupation number for photons; in ordinary units, the wave action is $\hbar$ times the occupation number.

The energy-momentum tensor is constructed using (8.11). One finds

$$T_M{}^{\mu\nu}(\boldsymbol{k}) = \frac{\partial \mathcal{L}_M}{\partial(\partial_\mu \Psi_M)}\, \partial^\nu \Psi_M = N_M(\boldsymbol{k}) v_{gM}{}^\mu(\boldsymbol{k}) k_M{}^\nu. \tag{D.32}$$

The components of $T_M{}^{\mu\nu}(\boldsymbol{k})$ are

$$\begin{aligned} T_M{}^{00}(\boldsymbol{k}) &= W_M(\boldsymbol{k}) = \omega_M(\boldsymbol{k}) N_M(\boldsymbol{k}), \qquad & T_M{}^{0j}(\boldsymbol{k}) &= [\boldsymbol{P}_M(\boldsymbol{k})]^j = [k N_M(\boldsymbol{k})]^j, \\ T_M{}^{i0}(\boldsymbol{k}) &= [\boldsymbol{F}_M(\boldsymbol{k})]^i = [\boldsymbol{v}_{gM}(\boldsymbol{k})(\boldsymbol{k}) N_M(\boldsymbol{k})]^i, \qquad & T_M{}^{ij}(\boldsymbol{k}) &= [\boldsymbol{v}_{gM}(\boldsymbol{k})(\boldsymbol{k})]^i [\boldsymbol{P}_M(\boldsymbol{k})]^j. \end{aligned} \tag{D.33}$$

These are the energy density, the momentum density, the energy flux and the stress 3-tensor, respectively.

On dividing the expression for the electrical energy in the waves implied by (D.24) by the expression implied by the first of (D.33) with (D.32) for the total energy $W_M(\boldsymbol{k})$, one identifies the ratio $R_M(\boldsymbol{k})$ of electric to total energy:

$$R_M(\boldsymbol{k}) = -\left[\frac{1}{\omega}\,\frac{\partial}{\partial \omega}\,\Lambda_M(k)\right]^{-1}_{\omega\,=\,\omega_M(\boldsymbol{k})}, \tag{D.34}$$

with

$$\Lambda_M(k) := e_M{}^*_\mu(\boldsymbol{k}) e_{M\nu}(\boldsymbol{k}) \Lambda^{\mu\nu}(k_M). \tag{D.35}$$

An alternative form is

$$R_M(\boldsymbol{k}) = \Big[\frac{\lambda^{0\sigma}{}_{0\sigma}(k)}{\omega \partial\lambda(k)/\partial\omega}\Big]_{\omega\,=\,\omega_M(\boldsymbol{k})}, \tag{D.36}$$

and another alternative form is

$$R_M(\boldsymbol{k}) = \Big[2 - \frac{\mu_0}{\omega}\,\frac{\partial}{\partial \omega}\,\alpha_M(k)\Big]^{-1}_{\omega\,=\,\omega_M(\boldsymbol{k})}, \tag{D.37}$$

with

$$\alpha_M(k) := e_M{}^*_\mu(\boldsymbol{k}) e_{M\nu}(\boldsymbol{k}) \alpha^{\mu\nu} k_M. \tag{D.38}$$

The derivations of these alternative forms starts from the defining relation

$$\Lambda^\mu{}_\rho(k) \lambda^{\rho\nu\alpha\beta}(k) = \lambda(k)\,[g^{\mu\alpha} k^\nu k^\beta - g^{\mu\beta} k^\nu k^\alpha] \tag{D.39}$$

for the second order cofactors. A contraction gives

$$\Lambda^{\mu}{}_{\rho}(k)\lambda^{\rho\sigma}{}_{\mu\beta}(k) = -3\,k^{\sigma}k_{\beta}\lambda(k), \tag{D.40}$$

and then differentiating with respect to ω and setting $\lambda(k) = 0$ gives

$$\Big[\Big\{\frac{\partial}{\partial\omega}\,\Lambda^{\mu}{}_{\rho}(k)\Big\}\,\lambda^{\rho\sigma}{}_{\mu\beta}(k)\Big]_{\lambda(k)\,=\,0} = -\Big[k^{\sigma}k_{\beta}\,\frac{\partial}{\partial\omega}\,\lambda(k)\Big]_{\lambda(k)\,=\,0}. \tag{D.41}$$

The equivalence of the forms (D.35) to (D.36) may then be deduced by setting $\sigma = 0 = \beta$ in (D.41). The remaining form (D.37) follows by using the explicit expression (D.10) for $\Lambda^{\mu\nu}(k)$, together with the normalization and gauge conditions for $e_M{}^{\mu}(\boldsymbol{k})$.

5. The Photon Propagator

The inhomogeneous wave equation (D.9) may be solved by introducing the Green's function or photon propagator $D(k)$. Then the solution is

$$A^{\mu}(k) = D^{\mu}{}_{\nu}(k)J_{\rm ext}{}^{\nu}(k). \tag{D.42}$$

As discussed in Lecture 11, the photon propagator is defined only to within a transformation of the form

$$D'^{\mu\nu}(k) = D^{\mu\nu}(k) + \chi^{\mu}(k)k^{\nu} + k^{\mu}\zeta^{\nu}(k), \tag{D.43}$$

where $\chi(k)$ and $\zeta(k)$ are arbitrary.

A formal definition of the propagator is as a solution of

$$\Lambda^{\mu}{}_{\nu}(k)D^{\nu\rho}(k) = -\mu_0 g^{\mu\rho}. \tag{D.44}$$

Then (D.42) is obviously a solution of (D.9). However, it is convenient to use the charge continuity relation to replace (D.44) by

$$\Lambda^{\mu}{}_{\nu}(k)D^{\nu\rho}(k) = -\mu_0[g^{\mu\rho} - k^{\mu}k^{\rho}/k^2]. \tag{D.45}$$

A practical difficulty in constructing the photon propagator is that $\Lambda(k)$, regarded as a matrix, has no inverse. This may be seen by noting that the relations

$$k_{\mu}\Lambda^{\mu\nu}(k) = 0, \qquad k_{\nu}\Lambda^{\mu\nu}(k) = 0. \tag{D.46}$$

imply that k is an eigenfunction of the matrix $\Lambda(k)$ with zero eigenvalue. The fact that $\Lambda(k)$ has one zero eigenvalue implies that its determinant vanishes and that its inverse does not exist. There are two ways of overcoming this difficulty. One is to choose a particular gauge, and to replace (D.44) by a gauge-dependent form whose solution is straightforward. Two examples of this approach are illustrated below, for the temporal and Lorentz gauges respectively. The other approach involves solving (D.45) directly by projecting onto the 3-dimensional matrix subspace orthogonal to k.

Temporal Gauge

In the temporal gauge $A^0(k)$ is zero by hypothesis. One may then replace (D.44) by an equation that involves only the space indices:

$$\Lambda^i{}_r(k) D^r{}_j(k) = \mu_0\, \delta^i{}_j. \tag{D.47}$$

Let the determinant of $\Lambda^i{}_j(k)$ be $\lambda_T(k)$, and let $\lambda_T{}^i{}_j(k)$ be the signed cofactor of $\Lambda^j{}_i(k)$. Then by construction on has

$$\Lambda^i{}_r(k) \lambda_T{}^r{}_j(k) = \lambda_T(k)\, \delta^i{}_j \tag{D.48}$$

The solution of (D.47) is then

$$D^i{}_j(k) = \frac{\lambda_T{}^i{}_j(k)}{\lambda_T(k)} \tag{D.49}$$

This gives an acceptable form for the propagator $D^\mu{}_\nu(k)$ by inserting zeros in the $\mu = 0$ and $\nu = 0$ elements.

Lorentz Gauge

In the Lorentz gauge one has $kA(k) = 0$, and then one is free to replace $\Lambda(k)$, as given by (1.2.2), by $\Lambda_L(k)$, defined by

$$\Lambda_L{}^{\mu\nu}(k) = k^2 g^{\mu\nu} + \mu_0\, \alpha^{\mu\nu}(k). \tag{D.50}$$

The matrix $\Lambda_L(k)$ has a nonzero determinant, $\lambda_L(k)$ say, and its signed cofactors are denoted $\lambda_L{}^\mu{}_\nu(k)$. The propagator in this case is given by

$$D^\mu{}_\nu(k) = \frac{\lambda_L{}^\mu{}_\nu(k)}{\lambda_L(k)}. \tag{D.51}$$

Although this is a covariant form of the propagator it is explicitly gauge-dependent. The general form of the propagator includes both the forms (D.49) and (D.50) as special cases.

Gauge-Independent Form of the Propagator

A general solution of (D.45) may be obtained by considering the relation (D.39) satisfied by the second order cofactors. After contracting this expression with $k_\nu k_\beta$, comparison with (D.45) leads to identification of an expression for the propagator:

$$D^{\mu\nu}(k) = -\mu_0\, \frac{k_\alpha k_\beta}{k^4}\, \frac{\lambda^{\mu\alpha\nu\beta}(k)}{\lambda(k)}. \tag{D.52}$$

More generally one may contract (D.39) with arbitrary quantities G_α and G'_β,and then (D.52) is replaced by

$$D^{\mu\nu}(k) = -\mu_0\, \frac{G_\alpha G'_\beta}{(Gk)(G'k)}\, \frac{\lambda^{\mu\alpha\nu\beta}(k)}{\lambda(k)}. \tag{D.53}$$

If one identifies G in (D.53) with the quantities introduced in (2.34) to (2.36), then $A(k)$ has the appropriate gauge. Specifically, after Fourier transforming, (2.34) to (2.36) give

$$\text{Lorentz gauge :} \qquad \hat{G}_\alpha = k_\alpha, \tag{D.54}$$

$$\text{Coulomb gauge :} \qquad \hat{G}_\alpha = [0, \boldsymbol{k}], \tag{D.55}$$

$$\text{temporal gauge :} \qquad \hat{G}_\alpha = [1, \mathbf{0}]. \tag{D.56}$$

respectively. The expressions (D.49) for the temporal gauge and (D.51) for the Lorentz gauge are equivalent to the expressions obtained from (D.53) by setting $G = G'$ and equating them to the expressions given for the temporal gauge (D.56) and the Lorentz gauge (D.54) respectively.

It is useful to refer to an arbitrary gauge as the "G-gauge", with the gauge condition being

$$G_\alpha A^\alpha(k) = 0. \tag{D.57}$$

Appendix E

The Vertex Function in a Magnetic Field

The vertex function for an electron in a magnetic field is defined for the Landau gauge by (18.4). In this Appendix the vertex function is evaluated explicitly in both the Landau gauge and in the cylindrical gauge, as discussed in Appendix B, and it is shown that a gauge-independent vertex function may be identified. The properties of the functions $J^n_\nu(x)$, which are closely related to generalized Laguerre polynomials are then reviewed and are used to derive approximate forms for the vertex function.

1. The Vertex Function in the Landau Gauge

The definition (18.4) of the vertex function is

$$[\gamma^{\epsilon' \epsilon}_{q' q}(\boldsymbol{k})]^\mu := V \int d^3\boldsymbol{x}\, \exp[-i\boldsymbol{k}\cdot\boldsymbol{x}]\, \overline{\Psi}^{\epsilon'}_{q'}(\boldsymbol{x}) \gamma^\mu \Psi^\epsilon_q(\boldsymbol{x}), \tag{E.1}$$

with the wavefunctions given by (6.20), viz.

$$\begin{aligned}
\Psi^\epsilon_{q+}(t,\boldsymbol{x}) &= \frac{\exp[-i\epsilon\varepsilon t + i\epsilon p_y y + i\epsilon p_z z]}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} (\epsilon\varepsilon_q + m)v_{n-1}(\xi) \\ 0 \\ \epsilon p_z v_{n-1}(\xi) \\ i p_n v_n(\xi) \end{pmatrix}, \\
\Psi^\epsilon_{q-}(t,\boldsymbol{x}) &= \frac{\exp[-i\epsilon\varepsilon t + i\epsilon p_y y + i\epsilon p_z z]}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} 0 \\ (\epsilon\varepsilon_q + m)v_n(\xi) \\ -i p_n v_{n-1}(\xi) \\ -\epsilon p_z v_n(\xi) \end{pmatrix},
\end{aligned} \tag{E.2}$$

Let the components of $\boldsymbol{k}$ be written

$$\boldsymbol{k} = (k_\perp \cos\psi, k_\perp \sin\psi, k_z). \tag{E.3}$$

In (E.1) the integrals over y and z are trivial. The integral over x may be reduced to the standard integral (Gradshteyn and Ryzhik 1965, 7.377)*

$$\begin{aligned}
\int_{-\infty}^{\infty} dx\, \exp[-ik_x x]\, v_{n'}(\xi') v_n(\xi) = (eB)^{-1/2} \exp[ik_x(\epsilon p_y + \epsilon' p'_y)/2eB] \\
\{-i\exp(-i\psi)\}^{n'-n}\, J^n_{n'-n}(k_\perp^2/2eB),
\end{aligned} \tag{E.4}$$

where the definition (18.6) has been used, viz.

$$J^n_\nu(x) = \left(\frac{n!}{(n+\nu)!}\right)^{1/2} e^{-x/2} x^{\nu/2}\, L^\nu_n(x). \tag{E.5}$$

* I.S. Gradshteyn and I.M. Ryzhik, *Tables of Integrals, Sums and Products*, Academic Press.

where $L_n^\nu(x)$ is a generalized Laguerre polynomial. Then (E.1) gives

$$[\gamma_{q'q}^{\epsilon'\epsilon}(\boldsymbol{k})]^\mu = \frac{2\pi}{L_y}\,\delta(\epsilon p_y - \epsilon' p'_y - k_y)\,\frac{2\pi}{L_z}\,\delta(\epsilon p_z - \epsilon' p'_z - k_z)\,[\Gamma_{q'q}^{\epsilon'\epsilon}(\boldsymbol{k})]^\mu. \tag{E.6}$$

It is convenient to write the gauge-independent vertex function $[\Gamma_{q'q}^{\epsilon'\epsilon}(\boldsymbol{k})]^\mu$ in terms of the $C_1,\ldots,C_4$ introduced in (6.19), viz.

$$\begin{aligned}
\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} &= \frac{1}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} \epsilon\varepsilon_q + m \\ 0 \\ \epsilon p_z \\ ip_n \end{pmatrix}, \\
\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} &= \frac{1}{\sqrt{2\epsilon\varepsilon_q(\epsilon\varepsilon_q + m)V}} \begin{pmatrix} 0 \\ \epsilon\varepsilon_q + m \\ -ip_n \\ -\epsilon p_z \end{pmatrix},
\end{aligned} \tag{E.7}$$

with, cf. (6.26),

$$V = \frac{L_y L_z}{(eB)^{1/2}}. \tag{E.8}$$

One finds

$$\begin{aligned}
[\Gamma_{q'q}^{\epsilon'\epsilon}(\boldsymbol{k})]^\mu =& \{-i\exp(-i\psi)\}^{n'-n} \\
&\Big([C_1'^* C_1 + C_3'^* C_3]J_{n'-n}^{n-1} + [C_2'^* C_2 + C_4'^* C_4]J_{n'-n}^{n}, \\
&i[C_1'^* C_4 + C_3'^* C_2]e^{i\psi}J_{n'-n-1}^{n} - i[C_2'^* C_3 + C_4'^* C_1]e^{-i\psi}J_{n'-n+1}^{n-1}, \\
&[C_1'^* C_4 + C_3'^* C_2]e^{i\psi}J_{n'-n-1}^{n} + [C_2'^* C_3 + C_4'^* C_1]e^{-i\psi}J_{n'-n+1}^{n-1}, \\
&[C_1'^* C_3 + C_3'^* C_1]J_{n'-n}^{n-1} - [C_2'^* C_4 + C_4'^* C_2]J_{n'-n}^{n}\Big),
\end{aligned} \tag{E.9}$$

where the argument of the functions $J_{n'-n}^n$ is $k_\perp^2/2eB$. The form (E.9) is general in the sense that different choices of the spin eigenfunctions correspond to different values of $C_1,\ldots,C_4$. The values for the eigenfunctions (E.2) are given by (E.7) and the value for the other eigenfunctions may be written down by inspection of the results given in Appendix B, specifically (B.40) or (B.46).

2. The Vertex Function in the Cylindrical Gauge

The solutions of Dirac's equation for an electron in a magnetostatic field is derived in the cylindrical gauge in Appendix B. Combining (B.21) and (B.28) one has

$$\Psi(t,\boldsymbol{x}) = \exp[-i\epsilon\varepsilon t + i\epsilon p_z z] \begin{pmatrix} C_1 J_{n-r-1}^s(\frac{1}{2}eB\varpi^2)\exp[i(n-r-1)\phi] \\ C_2 J_{n-r}^s(\frac{1}{2}eB\varpi^2)\exp[i(n-r)\phi] \\ C_3 J_{n-r-1}^s(\frac{1}{2}eB\varpi^2)\exp[i(n-r-1)\phi] \\ C_4 J_{n-r}^s(\frac{1}{2}eB\varpi^2)\exp[i(n-r)\phi] \end{pmatrix} \tag{E.10}$$

where r is the radial quantum number. In this case it the normalization volume $V = L_x L_y L_z$ may be identified as

$$V = \frac{L_z}{eB}. \tag{E.11}$$

The argument for the choice (E.11) is similar to the argument for (E.8). Specifically, on taking the discrete limit only $p_z = n_z 2\pi/L_z$ is discrete, and only the length L_z appears. The area $L_x L_y$ may thus be set equal to the natural area $1/eB$ that is available on dimensional grounds.

The definition (E.1) of the vertex function unchanged, but is to be evaluated in cylindrical polar coordinates ϖ, ϕ, z, rather than in Cartesian coordinates. The standard integral required is

$$\begin{aligned}\int_0^{2\pi} d\phi \int_0^\infty d\varpi\, \varpi\, e^{-ik_\perp \varpi \cos(\psi-\phi)} e^{-i(n'-r')\phi} J^{r'}_{n'-r'}(\tfrac{1}{2}eB\varpi^2)\, e^{i(n-r)\phi} J^{r}_{n-r}(\tfrac{1}{2}eB\varpi^2) \\ = \frac{2\pi}{eB} \left\{-i\exp(-i\psi)\right\}^{r-r'} J^r_{r'-r}(k_\perp^2/2eB) \left\{-i\exp(-i\psi)\right\}^{n'-n} J^n_{n'-n}(k_\perp^2/2eB).\end{aligned} \tag{E.12}$$

The resulting expression is

$$[\gamma^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})]^\mu = \frac{2\pi}{L_z}\, \delta(\epsilon p_z - \epsilon' p'_z - k_z) \left\{-i\exp(-i\psi)\right\}^{r-r'} J^r_{r'-r}(k_\perp^2/2eB)\, [\Gamma^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})]^\mu. \tag{E.13}$$

Quite generally the vertex function may be written in the form

$$[\gamma^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})]^\mu = d^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})\, [\Gamma^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})]^\mu, \tag{E.14}$$

where $d^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})$ is a gauge-dependent factor, and where $[\Gamma^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})]^\mu$ is the gauge-independent vertex function. The gauge-dependent factors are

$$d^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k}) = \frac{2\pi}{L_y}\, \delta(\epsilon p_y - \epsilon' p'_y - k_y)\, \frac{2\pi}{L_z}\, \delta(\epsilon p_z - \epsilon' p'_z - k_z) \tag{E.15}$$

in the Landau gauge and

$$d^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k}) = \frac{2\pi}{L_z}\, \delta(\epsilon p_z - \epsilon' p'_z - k_z) \left\{-i\exp(-i\psi)\right\}^{r-r'} J^r_{r'-r}(k_\perp^2/2eB) \tag{E.16}$$

in the cylindrical gauge.

3. Average over the Position of the Gyrocenter

The quantum number p_y in the Landau gauge and r in the cylindrical gauge are interpreted as specifying the position of the center of gyration of the particle. This may be seen by considering mean values.

The mean value of the x-coordinate in the Landau gauge is the x-coordinate of the center of gyration in the non-quantum limit. The correspondence principle allows us to identify the x-coordinate of the center of gyration in the quantum case as this mean value, which is given by

$$\langle x\rangle = \int d^3\boldsymbol{x}\, x\, [\Psi_q^\epsilon(t,\boldsymbol{x})]^\dagger \Psi_q^\epsilon(t,\boldsymbol{x}). \tag{E.17}$$

On inserting the explict wavefunctions (E.2), the integrals over y and z are trivial and the integral over x reduces to

$$\langle x\rangle = \int_{-\infty}^{\infty} d\xi\, x\, v_n^2(\xi) = -\frac{\epsilon p_y}{eB}. \tag{E.18}$$

Thus $\langle x\rangle$ is proportional to p_y, justifying the interpretation of p_y as specifying the x-coordinate of the center of gyration.

In the cylindrical gauge the mean value of the square of the radial distance ($\varpi^2 = x^2 + y^2$) may be evaluated (Sokolov and Ternov 1968, p.73)*:

$$\langle \varpi^2\rangle = \frac{1}{eB}\,(2n + 2r + 1). \tag{E.19}$$

In the non-quantum limit $\langle \varpi^2\rangle$ is equal to $R^2 + a^2$, where R is the radius of gyration and a is the radial distance of the center of gyration from the z-axis. The radius of gyration in the non-quantum limit is $R = (2n/eB)^{1/2}$, and hence the radial distance of the center of gyration from the z-axis is related to $(2r/eB)^{1/2}$. This justifies the interpretation of r as specifying the radial position of the center of gyration.

For most purposes the position of the center of gyration is irrelevant and one averages over it. In the Landau gauge with $L_x = 1/(eB)^{1/2}$, cf. (6.26), this average involves the operation

$$\hat{O}_{\rm av} := \frac{1}{L_x}\int_{-L_x/2}^{L_x/2} = \frac{2\pi}{(eB)^{1/2}}\int \frac{dp_y}{2\pi} \tag{E.20}$$

where (E.17) has been used. Similarly, in the cylindrical gauge the average involves the operation

$$\hat{O}_{\rm av} := \frac{1}{\pi a^2}\int_0^{2\pi} d\phi \int_0^a d\varpi\,\varpi = 2\pi \sum_r, \tag{E.21}$$

where the choice of normalization area corresponds to $\pi a^2 = 1/eB$. The averaging operation is to be applied to the transition rate, which involves the square of the S-matrix element.

---

* A.A. Sokolov and I.M. Ternov, *Synchrotron Radiation*, Pergamon Press.

For gyromagnetic emission, there is an initial electron and a final electron. The average over the position (along the x-axis in the Landau gauge or in the x, y-plane in the cylindrical gauge) reduces to

$$\hat{O}_{\rm av}|d^{\epsilon'\epsilon}_{q'q}(\boldsymbol{k})|^2 = \frac{1}{eBV}\,2\pi\,\delta(\epsilon p_z - \epsilon' {p'}_z - k_z). \tag{E.22}$$

The transition rate is then independent of the quantum numbers that specify the position of the final electron. The density of states for the final electron is

$$D_f = \frac{L_y L_z}{(\pi)^2}\,dp_y dp_z \quad \text{or} \quad \frac{L_z}{2\pi}\,dp_z \tag{E.23}$$

in the Landau or cylindrical gauges, respectively. The modified density of final states, after performing the integral over ${p'}_y$ over the sum over r' is

$$D_f = \frac{eBV}{2\pi}\,\frac{dp_z}{2\pi}. \tag{E.24}$$

These procedures have been used in deriving the probability (18.7).

4. Properties of the Functions $J^n_\nu(x)$

The functions $J^n_\nu(x)$ have the following properties.

Definition

$$J^n_\nu(x) == (-)^\nu J^{n+\nu}_{-\nu}(x)\left(\frac{n!}{(n+\nu)!}\right)^{1/2} e^{-x/2}x^{\nu/2}\,L^\nu_n(x). \tag{E.25}$$

Recursion Formulas

$$\begin{aligned} J^{n+1}_\nu(x) &= \left(\frac{n+\nu+1}{n+1}\right)^{1/2} J^n_\nu(x) - \left(\frac{x}{n+1}\right)^{1/2} J^n_{\nu+1}(x) \qquad \text{(E.26)}\\ &= \frac{-x+n+1}{(n+1)^{1/2}(n+\nu+1)^{1/2}} J^n_\nu(x) + \left(\frac{x(n+\nu)}{(n+1)(n+\nu+1)}\right)^{1/2} J^n_{\nu-1}(x), \\ J^{n-1}_\nu(x) &= \left(\frac{n+\nu}{n}\right)^{1/2} J^n_\nu(x) - \left(\frac{x}{n}\right)^{1/2} J^n_{\nu-1}(x) \qquad \text{(E.27)}\\ &= \frac{-x+n}{(n)^{1/2}(n+\nu)^{1/2}} J^n_\nu(x) + \left(\frac{x(n+\nu+1)}{n(n+\nu)}\right)^{1/2} J^n_{\nu+1}(x), \end{aligned}$$

$$(x+\nu)J^n_\nu(x) = [x(n+\nu)]^{1/2}J^n_{\nu-1}(x) + [x(n+\nu+1)]^{1/2}J^n_{\nu+1}(x), \tag{E.28}$$

$$2x\frac{d}{dx}\,J^n_\nu(x) = [x(n+\nu)]^{1/2}J^n_{\nu-1}(x) - [x(n+\nu+1)]^{1/2}J^n_{\nu+1}(x). \tag{E.29}$$

Sum Rules

$$\sum_{n'=0}^{\infty} = J^{n}_{n'-n}(x)J^{n''}_{n'-n''}(x) = \delta_{nn''}, \tag{E.30}$$

$$\sum_{n'=0}^{\infty} = (n'-n)[J^{n}_{n'-n}(x)]^2 = x. \tag{E.31}$$

Integral Identities

$$\int_0^{\infty} dx\, J^{n}_{\nu}(x)J^{n'}_{\nu}(x) == \delta_{nn'}, \tag{E.32}$$

$$\int_0^{\infty} dx\, x^{1/2}[J^{n}_{\nu}(x)]^2 = (n+\nu+1)^{1/2}\left(1+\frac{n+\frac{1}{2}}{4(n+\nu+1)}\right), \tag{E.33}$$

$$\sum_{n'=0}^{\infty} x[J^{n}_{\nu}(x)]^2 = 2n+\nu+1. \tag{E.34}$$

Closure Relation for Laguerre Polynomials

$$\sum_{n=0}^{\infty} J^{n}_{0}(x)J^{n}_{0}(x') = \delta(x-x'). \tag{E.35}$$

Particular Values

$$J^{0}_{\nu}(x) = (-)^{\nu}J^{\nu}_{-\nu}(x) = \left(\frac{x^{\nu}e^{-x}}{\nu!}\right)^{1/2}, \tag{E.36}$$

$$J^{1}_{\nu}(x) = (-)^{\nu}J^{\nu+1}_{-\nu}(x) = \left(\frac{x^{\nu}e^{-x}}{(\nu+1)!}\right)^{1/2}(\nu+1-x), \tag{E.37}$$

$$J^{2}_{\nu}(x) = (-)^{\nu}J^{\nu+2}_{-\nu}(x) = \left(\frac{x^{\nu}e^{-x}}{(\nu+2)!}\right)^{1/2}[(\nu+1)(\nu+2)-2(\nu+2)x+x^2], \tag{E.38}$$

$$J^{3}_{\nu}(x) = (-)^{\nu}J^{\nu}_{-\nu}(x) = \left(\frac{x^{\nu}e^{-x}}{(\nu+3)!}\right)^{1/2}[(\nu+1)(\nu+2)(\nu+3) - 3(\nu+2)(\nu+3)x+3(\nu+3)x^2-x^3]. \tag{E.38}$$

Limiting Values

(i) Small x ($x \ll 1$)

$$J^{n}_{\nu}(x) = \begin{cases} [(n+\nu)!x^{\nu}/n!(\nu!)^2]^{1/2}, & \text{for } \nu > 0; \\ (-)^{\nu}[n!x^{|}\nu|/(n-|\nu|)!(|\nu|!)^2]^{1/2}, & \text{for } \nu < 0. \end{cases} \tag{E.39}$$

(ii) Classical Limit ($n \to \infty$)

$$J_\nu^n(z^2/4n) = \left(\frac{(n+\nu)!}{n!n^\nu}\right)^{1/2} \sum_{s=0}^{\infty} C_s\,(z/2n)^s\,J_{\nu+s}(z); \tag{E.40}$$

$$C_0 = 1, \quad C_2 = -\tfrac{1}{2}(\nu+1), \quad C_2 = \tfrac{1}{8}(\nu+1)(\nu+2),$$
$$(s+1)C_{s+1} = -\tfrac{1}{2}(\nu+1)C_s + \tfrac{1}{4}(\nu+s)C_{s-1} - \tfrac{1}{2}nC_{s-2}.$$

Airy Integral Approximation

(i) $0 < x < x_0 := (\sqrt{n} - \sqrt{n'})^2$

$$J_{n-n'}^{n'}(x) = \frac{1}{\pi\sqrt{3}}\,[1 - x/x_0]^{1/2}\,K_{1/3}(\tfrac{2}{3}[nn'x_0^2]^{1/4}[1-x/x_0]^{3/2}); \tag{E.41}$$

(i) $x > x'_0 := (\sqrt{n} + \sqrt{n'})^2$

$$J_{n-n'}^{n'}(x) = \frac{1}{\pi\sqrt{3}}\,[x/x'_0 - 1]^{1/2}\,K_{1/3}(\tfrac{2}{3}[nn'x'^2_0]^{1/4}[x/x'_0-1]^{3/2}); \tag{E.42}$$

Addition Theorem

$$\{-i\exp(-i\Phi)\}^{n''-n}\,J_{n''-n}^{n}(\tfrac{1}{2}R^2) = \sum_{n'=0}^{\infty} \exp[-\tfrac{1}{2}irr'\sin(\phi-\phi')]\{-i\exp(-i\phi)\}^{n'-n}$$
$$J_{n'-n}^{n}(\tfrac{1}{2}r^2)\{-i\exp(-i\phi')\}^{n''-n'}\,J_{n''-n'}^{n'}(\tfrac{1}{2}r'^2); \tag{E.43}$$

$$R\cos\Phi = r\cos\phi + r'\cos\phi', \quad R\sin\Phi = r\sin\phi + r'\sin\phi'.$$

5. Approximate Forms for the Vertex Function

Four approximations to the vertex function are relevant. One is the limit of small to moderate n which is relevant for nonrelativistic particles in strong magnetic fields. The second is the classical (i.e. non-quantum) limit. The third is the ultrarelativistic limit for gyromagnetic emission, which is the "synchrotron" limit. The fourth is the ultrarelativistic limit for pair creation. These correspond to the approximations (E.39), (E.40), (E.41), and(E.42), respectively. The first three approximations apply to emission. The formulas given below are for emission by an electron, and the corresponding formulas for emission by a positron follow from

$$[\Gamma^{--}_{q'q}(\boldsymbol{k})]^\mu = (-)^{l'-l}[\Gamma^{++}_{qq'}(\boldsymbol{k})]^{*\mu}. \tag{E.44}$$

In the nonrelativistic, strong-field limit the approximations (E.39) are made in (E.9) together with the approximations

$$p_z^2/m^2 \ll 1, \quad p_n^2/m^2 = 2nB/B_c \ll 1, \quad x = k_\perp^2/2eB \ll 1. \tag{E.45}$$

In addition the dispersion is assumed weak (that is, the refractive index is assumed close to unity), so that one has

$$k_z \approx \omega\cos\theta, \quad x \approx j^2(B/2B_c)\sin^2\theta, \tag{E.46}$$

with $j = n - n'$ so that the emission is at $\omega \approx j\Omega$ with $\Omega = eB/m$. The relevant approximation to (E.9) is then given by (18.9), viz.

$$\begin{aligned}[\Gamma^{++}_{q'q}(\boldsymbol{k})]^\mu =& \delta_{s's}\left[(-)^j\left(\frac{B}{2B_c}\right)^{1/2}\left(\frac{\ell!}{(\ell-j)!}\right)^{1/2}\frac{x^{\frac{1}{2}(j-1)}}{(j-1)!}\right](1,i,0)\\ &+\delta_{s'-s}\left[(-)^(j-s)\frac{B\cos\theta}{2B_c}\left(\frac{\ell!}{(\ell-j+s)!}\right)^{1/2}\frac{x^{\frac{1}{2}(j-s)}}{(j-s)!}(j,isj,-\tan\theta)\right].\end{aligned} \tag{E.47}$$

In the classical limit one has $p_n = p_\perp = \gamma m v_\perp$ and it is convenient to write $p_z = \gamma m v_z$. The (E.9) with the approximation (E.40) gives

$$[\Gamma^{++}_{q'q}(\boldsymbol{k})]^\mu = U^\mu(s,k), \tag{E.48}$$

with

$$\begin{aligned}U^\mu(s,k) &= [\gamma J_s(z), \gamma \boldsymbol{V}(s,k)],\\ \boldsymbol{V}(s,k) &= \Big\{\tfrac{1}{2}v_\perp\,[e^{i\psi}J_{s-1}(z) + e^{-i\psi}J_{s+1}(z)],\\ &\qquad i\tfrac{1}{2}v_\perp\,[e^{i\psi}J_{s-1}(z) - e^{-i\psi}J_{s+1}(z)], v_z J_s(z)\Big\},\end{aligned} \tag{E.49}$$

and with $z = k_\perp R$, where $R = p_\perp/eB$ is the radius of gyration.

The remaining two approximations apply in the ultrarelativistic limit and involve the Airy-integral approximations (E.41) and (E.42). In these cases n and n' are large and are assumed continuous; the sum over n is replaced by and integral. The derivations of the probabilities in these cases involve more than just approximations to the vertex function, and these derivations are not given here. The form (18.22) arises after changing the variables of integration from p_z and n to u and w by writing

$$\sinh u := p_z/m, \quad \sinh w := [\sqrt{n'} - \sqrt{n}]/2\sqrt{n'n}. \tag{E.50}$$

The actual approximations used are

$$J^n_{n'-n}(x) \approx (1/\pi\sqrt{3})(2m/\omega)\cosh u\,\cosh w\,K_{1/3}(z) \tag{E.51}$$

$$\frac{d}{dx}J^n_{n'-n}(x) \approx (1/\pi\sqrt{3})(2m^2/\omega^2)\cosh^2 u\,\cosh w\,K_{2/3}(z) \tag{E.52}$$

with

$$z := (2/3\chi)\,\cosh^2 u\cosh^2 w, \qquad \chi := \omega eB/2m^3 = \omega B/mB_c. \tag{E.53}$$

Exercise Set 1

Relativistic Quantum Mechanics Set 1

1. Maxwell's equations may be written in the 4-vector form

$$\partial^\mu F^{\nu\rho}(x) + \partial^\rho F^{\mu\nu}(x) + \partial^\nu F^{\rho\mu}(x) = 0,$$

$$\partial_\mu F^{\mu\nu}(x) = \mu_0 J^\nu(x).$$

Show explicitly how these equations reproduce the 3-vector familiar form of Maxwell's equation.

2. Show how

$$F^{\mu\nu}(x) = \partial^\mu A^\nu(x) - \partial^\nu A^\mu(x)$$

reduces to the familiar relation between $\boldsymbol{E}$, $\boldsymbol{B}$ and the vector and scalar potentials $\boldsymbol{A}$, ϕ:

$$\boldsymbol{E} = -\text{grad}\,\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \quad \boldsymbol{B} = \text{curl}\,\boldsymbol{A}.$$

3. The permutation symbol $\epsilon^{\mu\nu\rho\sigma}$ is completely antisymmetric in all its indices, and ϵ^{0123} is chosen equal to unity.

(a) Why is ϵ_{0123} necessarily equal to minus ϵ^{0123}?

(b) What is the numerical value of the invariant $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu\nu\rho\sigma}$?

4. Show that the Dirac matrices in the standard representation

$$\gamma^0 = \rho_z, \qquad \boldsymbol{\gamma} = i\rho_y\boldsymbol{\sigma}$$

satisfy

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}.$$

5. Construct the matrices $\not{P} \pm m$ in the standard representation, and show that the matrices satisfy

$$(\not{P} + m)(\not{P} - m) = P^2 - m^2.$$

Hints: One has

$$x^\mu = [t, \boldsymbol{x}], \quad J^\mu = [\rho, \boldsymbol{J}], \quad A^\mu = [\phi, \boldsymbol{A}],$$

$$\partial_\mu = [\partial/\partial t, \partial/\partial \boldsymbol{x}], \quad F^{\mu\nu}(x) = -F^{\nu\mu}(x),$$

$$F^{\mu\nu}(x) := [\boldsymbol{E}, \boldsymbol{B}] = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}.$$

6. How many independent invariants can be constructed from
   (a) two different 4-vectors p and k,
   (b) the Maxwell tensor alone,
   (c) the Maxwell tensor and one 4-vector,
   (d) the Maxwell tensor and two different 4-vectors?

Hint: The answer to (d) is 14; can you explain why?

Exercise Set 2

Relativistic Quantum Mechanics Set 2

1. Show that the Dirac matrices in the standard representation

$$\gamma^0 = \rho_z, \qquad \boldsymbol{\gamma} = i\rho_y \boldsymbol{\sigma}$$

satisfy

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}.$$

2. Construct the matrices $\not{P} \pm m$ in the standard representation, and show that the matrices satisfy

$$(\not{P} + m)(\not{P} - m) = P^2 - m^2.$$

3. In the *spinor representation* the Dirac matrices have the following forms:

$$\boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{\sigma} \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}.$$

(a) Construct γ^μ in this representation and show that it satisfies

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}.$$

(b) Show that if the Dirac wavefunction is written in the form

$$\Psi(x) = \begin{pmatrix} \xi \\ \eta \end{pmatrix},$$

where the ξ and η are spinors in two dimensions, then the Dirac equation becomes the two coupled equations

$$(p^0 + \boldsymbol{p} \cdot \boldsymbol{\sigma})\eta = m\xi, \qquad (p^0 - \boldsymbol{p} \cdot \boldsymbol{\sigma})\xi = m\eta.$$

(c) Show that both ξ and η satisfy the Klein-Gordon equation.