Advanced Quantum Mechanics

Peter S. Riseborough

February 19, 2015

Contents

1 Introduction

Non-relativistic mechanics yields a reasonable approximate description of physical phenomena in the range where the particles' kinetic energies are small compared with their rest mass energies. However, it should be noted that when the relativistic invariant mass of a particle is expressed in terms of its energy E and momentum $\underline{p}$ via

$$E^2 \ - \ \underline{p}^2 \ c^2 \ = \ m^2 \ c^4 \tag{1}$$

it implies that its dispersion relation has two branches

$$E \ = \ \pm \, c \, \sqrt{\underline{p}^2 \ + \ m^2 \ c^2} \tag{2}$$

In relativistic classical mechanics, it is assumed expedient to neglect the negative-energy solutions. This assumption based on the expectation that, since energy
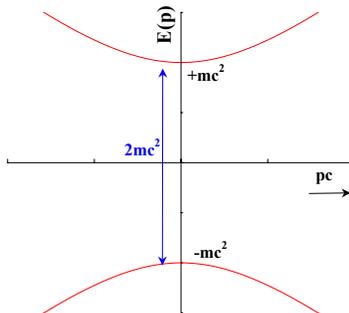
Figure 1: The positive and negative energy branches for a relativistic particle with rest mass m. The minimum separation between the positive-energy branch and the negative-energy branch is $2mc^2$.

can only change continuously, it is impossible that a particle with positive energy can make a transition from the positive to negative-energy states. However, in quantum mechanics, particles can make discontinuous transitions. Therefore, it is necessary to consider both the positive and negative-energy branches. These considerations naturally lead one to the concept of particles and anti-particles, and also to the realization that one must consider multi-particle quantum mechanics or field theory.

We shall take a look at the quantum mechanical description of the electromagnetic field, Dirac's relativistic theory of spin one-half fermions (such as leptons and quarks), and then look at the interaction between these particles and the electromagnetic field. The interaction between charged fermions and the electromagnetic field is known as Quantum Electrodynamics. Quantum Electrodynamics contains some surprises, namely that although the interaction

appears to be governed by a small coupling strength

$$\left(\frac{e^2}{\hbar\,c}\right) \sim \frac{1}{137.0359979} \qquad (3)$$

perturbation theory does not converge. In fact, straightforward perturbation theory is plagued by infinities. However, physics is a discipline which is aimed at uncovering the relationships between measured quantities. The quantities e and m which occur in quantum electrodynamics are theoretical constructs which, respectively, describe the bare charge of the electron and bare mass of the electron. This means one is assuming that e and m would be the results of measurements on a (fictional) electron which does not interact electromagnetically. That is, e and m are not physically measurable and their values are therefore unknown. What can be measured experimentally are the renormalized mass and the renormalized charge of the electron. The divergences found in quantum electrodynamics can be shown to cancel or drop out, when one relates different physically measurable quantities, as only the renormalized masses and energies enter the theory. Despite the existence of infinities, quantum electrodynamics is an extremely accurate theory. Experimentally determined quantities can be predicted to an extremely high degree of precision.

The framework of quantum electrodynamics can be extended to describe the unification of electrodynamics and the weak interaction via electro-weak theory, which is also well tested. The scalar ($U(1)$) gauge symmetry of the electromagnetic field is replaced by a matrix ($SU(2)$) symmetry of the combined electro-weak theory, in which the gauge field couples to the two components of the spinor wave functions of the fermions. The generalization of the gauge field necessitates the inclusion of additional components. Through symmetry breaking, some components of the field which mediates the electro-weak interaction become massive, i.e. have finite masses. The finite masses are responsible for the short range of the weak interaction. More tentatively, the gauge theory framework of quantum electrodynamics has also been extended to describe the interactions between quarks which is mediated by the gluon field. The gauge symmetry of the interaction is enlarged to an $SU(3)$ symmetry. However, unlike quantum electrodynamics where photons are uncharged and do not interact with themselves, the gluons do interact amongst themselves.

2   Quantum Mechanics of a Single Photon

Maxwell's equations were formulated to describe classical electromagnetism. In the quantum description, the classical electromagnetic field is described as being composed of a very large number of photons. Before one describes multi-photon quantum mechanics of the electromagnetic field, one should ascertain the form of the Schrödinger equation for a single photon. The photon is a massless, uncharged particle of spin-one.

A spin-one particle is described by a vector wave function. This can be heuristically motivated as follows:

A spin-zero particle has just one state and is uniquely described by a one-component field ψ.

A spin one-half particle has two independent states corresponding to the two allowed values of the z-component of the intrinsic angular momentum $S^z = \pm \frac{\hbar}{2}$. The wave function ψ of a spin one-half particle is a spinor which has two independent components

$$\psi(\underline{r}, t) = \left(\begin{array}{c} \psi^{(1)}(\underline{r}, t) \\ \psi^{(2)}(\underline{r}, t) \end{array} \right) \tag{4}$$

These two components can be used to represent two independent basis states.

We conjecture that since a particle with intrinsic spin S has $(2S + 1)$ independent basis states, then the wave function should have $(2S + 1)$ independent components.

A (non-relativistic) spin-one particle should have three independent states corresponding to the three possible values of the z-component of the intrinsic angular momentum. From the conjecture, one expects that the wave function ψ of a spin-one particle should have three components.

$$\underline{\psi}(\underline{r}, t) = \left(\begin{array}{c} \psi^{(1)}(\underline{r}, t) \\ \psi^{(3)}(\underline{r}, t) \\ \psi^{(3)}(\underline{r}, t) \end{array} \right) \tag{5}$$

This conjecture can be verified by examining the transformational properties of a vector field under rotations. Under a rotation of the field, the components of the field are transported in space, and also the direction of the vector field is rotated. This implies that the components of the transported field have to be rotated. The rotation of the direction of the field is generated by operators which turn out to be the intrinsic angular momentum operators. Specifically, the generators satisfy the commutation relations defining angular momentum, but also correspond to the subspace with angular momentum one.

2.1   Rotations and Intrinsic Spin

Under the transformation which takes $\underline{r} \rightarrow \underline{r}' = R\,\underline{r}$, the magnitude of the scalar field ψ at $\underline{r}$ is transferred to the point $\underline{r}'$. This defines the transformation $\psi \rightarrow \psi'$. The transformed scalar field ψ' is defined so that its value at $\underline{r}'$ has the same value as $\psi(\underline{r})$. That is

$$\psi'(\underline{r}') = \psi(\underline{r}) \tag{6}$$

or equivalently

$$\psi'(\hat{R}\,\underline{r}) \;=\; \psi(\underline{r}) \tag{7}$$

The above equation can be used to determine $\psi'(\underline{r})$ by using the substitution $\underline{r} \to \hat{R}^{-1}\underline{r}$ so

$$\psi'(\underline{r}) \;=\; \psi(\hat{R}^{-1}\underline{r}) \tag{8}$$

If $\hat{e}$ is a unit vector along the axis of rotation, the rotation of $\underline{r}$ through an
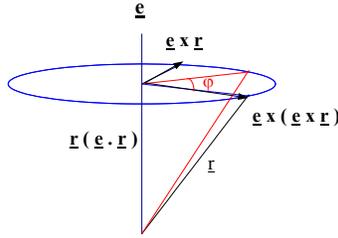
Figure 2: A rotation of the vector $\underline{r}$ through an angle φ about an axis $\hat{e}$.

infinitesimal angle $\delta\varphi$ is expressed as

$$\hat{R}\,\underline{r} \;=\; \underline{r} \;+\; \delta\varphi\,\hat{e}\,\wedge\,\underline{r}\;+\;\ldots \tag{9}$$

where terms of order $\delta\varphi^2$ have been neglected. Hence, under an infinitesimal
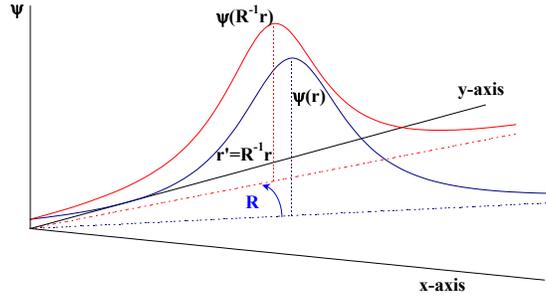
Figure 3: The effect of a rotation $\hat{R}$ on a scalar field $\psi(\underline{r})$.

rotation, the transformation of a scalar wave function can be found from the Taylor expansion

$$
\begin{aligned}
\psi'(\underline{r}) \;&=\; \psi(\underline{r} - \delta\varphi\,\hat{e}\wedge\underline{r}) \\
&=\; \psi(\underline{r}) \;-\; \delta\varphi\,(\,\hat{e}\,\wedge\,\underline{r}\,)\,.\,\underline{\nabla}\,\psi(\underline{r}) \;+\;\ldots \\
&=\; \psi(\underline{r}) \;-\; \delta\varphi\,(\,\underline{r}\,\wedge\,\underline{\nabla}\,)\,.\,\hat{e}\,\psi(\underline{r}) \;+\;\ldots \\
&=\; \psi(\underline{r}) \;-\; \frac{i\,\delta\varphi}{\hbar}\,(\,\hat{e}\,.\,\underline{\hat{L}}\,)\,\psi(\underline{r}) \;+\;\ldots
\end{aligned}
$$

8

$$= \exp\left[- \frac{i\,\delta\varphi}{\hbar}\,(\,\hat{e}\,.\,\underline{\hat{L}}\,)\right]\psi(\underline{r}) \tag{10}$$

where the operator $\hat{L}$ has been defined as

$$\hat{L} = -\,i\,\hbar\,\underline{r}\,\wedge\,\underline{\nabla} \tag{11}$$

Therefore, locally, rotations of the scalar field are generated by the orbital angular momentum operator $\underline{\hat{L}}$.

Since the operation $\hat{R}$ is a rotation, it also rotates a vector field $\underline{\psi}(\underline{r})$. Not only does the rotation transfer the magnitude of $\psi(\underline{r})$ to the new point $\underline{r}'$ but it must also rotate the direction of the transferred vector $\underline{\psi}(\underline{r})$. The rotated vector $\underline{\psi}'(\underline{r}')$ is denoted by $\hat{\mathcal{R}}\,\underline{\psi}(r)$. Mathematically, the transformation is expressed
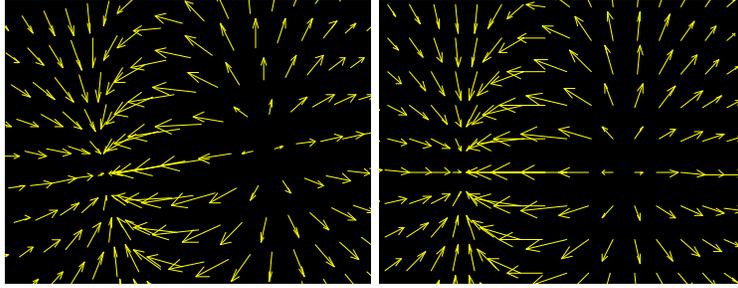
Figure 4: The effect of a rotation $\hat{R}$ on a vector field $\underline{\psi}(\underline{r})$. The rotation affects both the magnitude and direction of the vector.

as

$$\underline{\psi}'(\underline{r}') = \hat{\mathcal{R}}\,\underline{\psi}(\underline{r}) \tag{12}$$

or equivalently

$$\underline{\psi}'(\hat{R}\,\underline{r}) = \hat{\mathcal{R}}\,\underline{\psi}(\underline{r}) \tag{13}$$

The above equation can be used to determine $\underline{\psi}'(\underline{r})$ as

$$\underline{\psi}'(\underline{r}) = \hat{\mathcal{R}}\,\underline{\psi}(\hat{R}^{-1}\underline{r}) \tag{14}$$

The part of the rotational operator designated by $\hat{\mathcal{R}}$ does not affect the positional coordinates $(\underline{r})$ of the vector field, and so can be found by considering the rotation of the vector field $\underline{\psi}$ at the origin

$$\hat{\mathcal{R}}\,\underline{\psi} = \left(\,\hat{I}\,+\,\delta\varphi\,\hat{e}\,\wedge\,\right)\underline{\psi} \tag{15}$$

That is, the operator $\hat{\mathcal{R}}$ only produces a mixing of the components of $\underline{\psi}$. Hence, the complete rotational transformation of the vector field can be represented as

$$
\begin{aligned}
\underline{\psi}'(\underline{r}) &= \hat{\mathcal{R}} \; \underline{\psi}(\underline{r} - \delta\varphi \; \hat{e} \wedge \underline{r}) \\
&= \underline{\psi}(\underline{r} - \delta\varphi \; \hat{e} \wedge \underline{r}) \; + \; \delta\varphi \; \hat{e} \wedge \underline{\psi}(\underline{r} - \delta\varphi \; \hat{e} \wedge \underline{r}) \\
&= \underline{\psi}(\underline{r} - \delta\varphi \; \hat{e} \wedge \underline{r}) \; + \; \delta\varphi \; \hat{e} \wedge \underline{\psi}(\underline{r}) \; + \; \dots \\
&= \underline{\psi}(\underline{r}) \; - \; \frac{i \; \delta\varphi}{\hbar} \; (\; \hat{e} \; . \; \underline{\hat{L}} \;) \; \underline{\psi}(\underline{r}) \; + \; \delta\varphi \; \hat{e} \wedge \underline{\psi}(\underline{r}) \; + \; \dots \\
&= \underline{\psi}(\underline{r}) \; - \; \frac{i \; \delta\varphi}{\hbar} \; (\; \hat{e} \; . \; \underline{\hat{L}} \;) \; \underline{\psi}(\underline{r}) \; - \; \frac{i \; \delta\varphi}{\hbar} \; (\; \hat{e} \; . \; \underline{\hat{S}} \;) \; \underline{\psi}(\underline{r}) \qquad (16)
\end{aligned}
$$

where the terms of order $(\delta\varphi)^2$ have been neglected and a vector operator $\underline{\hat{S}}$ has been introduced. The operator $\underline{\hat{S}}$ only admixes the components of ψ^μ, unlike $\underline{\hat{L}}$ which only acts on the $\underline{r}$ dependence of the components. The components of the three-dimensional vector operator $\underline{S}$ are expressed as 3×3 matrices[1], with matrix elements

$$
(\hat{S}^{(i)})^{j,k} \; = \; - \; i \; \hbar \; \xi^{i,j,k} \qquad (18)
$$

where $\xi^{i,j,k}$ is the antisymmetric Levi-Civita symbol. The Levi-Civita symbol is defined by $\xi^{i,j,k} = 1$ if the ordered set (i, j, k) is obtained by an even number of permutations of $(1, 2, 3)$ and is -1 if it is obtained by an odd number of permutations, and is zero if two or more indices are repeated. Specifically, the antisymmetric matrices are given by

$$
\hat{S}^{(1)} \; = \; \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad (19)
$$

and by

$$
\hat{S}^{(2)} \; = \; \hbar \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \qquad (20)
$$

and finally by

$$
\hat{S}^{(3)} \; = \; \hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad (21)
$$

By using a unitary transform, the above three operators can be transformed into the standard representation of spin-one operators where $S^{(3)}$ is chosen to be diagonal. It is easily shown that the components of the matrix operators $\underline{\hat{L}}$ and $\underline{\hat{S}}$ satisfy the same type of commutation relations

$$
[\; \hat{L}^{(i)} \; , \; \hat{L}^{(j)} \;] \; = \; i \; \hbar \; \xi^{i,j,k} \; \hat{L}^{(k)} \qquad (22)
$$

---

[1] The component of the matrix denoted by

$$
(\underline{\hat{S}})^{j,k} \qquad (17)
$$

denotes the element of $\underline{\hat{S}}$ in the j-th row and k-th column.

and

$$[\ \hat{S}^{(i)}\ ,\ \hat{S}^{(j)}\ ]\ =\ i\ \hbar\ \xi^{i,j,k}\ \hat{S}^{(k)} \tag{23}$$

where the repeated index (k) is summed over. The above set of operators form a Lie algebra associated with the corresponding Lie group of continuous rotations. Thus, it is natural to identify these operators which arise in the analysis of transformations in classical physics with the angular momentum operators of quantum mechanics. In terms of these operators, the infinitesimal transformation has the form

$$\underline{\psi}'(\underline{r})\ \approx\ \underline{\psi}(\underline{r})\ -\ \frac{i\ \delta\varphi}{\hbar}\ \hat{e}\ .\ (\ \underline{\hat{L}}\ +\ \underline{\hat{S}}\ )\ \underline{\psi}(\underline{r})\ +\ \dots \tag{24}$$

or

$$\underline{\psi}'(\underline{r})\ =\ \exp\left[\ -\ \frac{i\ \delta\varphi}{\hbar}\ \hat{e}\ .\ (\ \underline{\hat{L}}\ +\ \underline{\hat{S}}\ )\ \right]\ \underline{\psi}(\underline{r}) \tag{25}$$

Thus, the transformation is locally accomplished by

$$\underline{\psi}'(\underline{r})\ =\ \exp\left[\ -\ \frac{i\ \delta\varphi}{\hbar}\ (\ \hat{e}\ .\ \underline{\hat{J}}\ )\ \right]\ \underline{\psi}(\underline{r}) \tag{26}$$

where

$$\underline{\hat{J}}\ =\ \underline{\hat{L}}\ +\ \underline{\hat{S}} \tag{27}$$

is the total angular momentum. The operator $\underline{\hat{S}}$ is the intrinsic angular momentum of the vector field $\underline{\psi}$. The magnitude of S is found from

$$\underline{\hat{S}}^2\ =\ (\hat{S}^{(1)})^2\ +\ (\hat{S}^{(2)})^2\ +\ (\hat{S}^{(3)})^2 \tag{28}$$

which is evaluated as

$$\underline{\hat{S}}^2\ =\ 2\ \hbar^2\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{29}$$

and is the Casimir operator. It is seen that a vector field has intrinsic angular momentum, with a magnitude given by the eigenvalue of $\underline{\hat{S}}^2$ which is

$$S\ (\ S\ +\ 1\ )\ \hbar^2\ =\ 2\ \hbar^2 \tag{30}$$

hence $S\ =\ 1$. Thus, it is seen that a vector field is associated with an intrinsic angular momentum of spin one.

2.2 Massless Particles with Spin Zero

First, we shall try and construct the Schrödinger equation describing a massless uncharged spinless particle. A spinless particle is described by a scalar wave function, and an uncharged particle is described by a real wave function. The derivation is based on the energy-momentum relation for a massless particle

$$E^2\ -\ \underline{p}^2\ c^2\ =\ 0 \tag{31}$$

11

which is quantized by using the substitutions

$$
\begin{aligned}
E &\rightarrow i\,\hbar\,\frac{\partial}{\partial t} \\
\underline{p} &\rightarrow \underline{\hat{p}} = -\,i\,\hbar\,\underline{\nabla}
\end{aligned}
\tag{32}
$$

One finds that the real scalar wave function $\psi(\underline{r}, t)$ satisfies the wave equation

$$
\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \underline{\nabla}^2 \right]\psi = 0
\tag{33}
$$

since $\hbar$ drops out. This is not a very useful result, since it is a second-order differential equation in time, and the solution of a second-order differential equation can only be determined if two initial conditions are given. Usually, the initial conditions are given by

$$
\begin{aligned}
\psi(\underline{r}, 0) &= f(\underline{r}) \\
\left.\frac{\partial \psi}{\partial t}\right|_{t=0} &= g(\underline{r})
\end{aligned}
\tag{34}
$$

In quantum mechanics, measurements disturb the state of the system and so it becomes impossible to design two independent measurements which can uniquely specify two initial conditions for one state. Hence, one has reached an impasse. Due to this difficulty and since there are no known examples of massless spinless particles found in nature, this theory is not very useful.

2.3   Massless Particles with Spin One

The wave function of an uncharged spin-one particle is expected to be represented by a real vector function.

We shall try and factorize the wave equation for the vector $\underline{E}$ into two first-order differential equations, each of which requires one boundary condition. This requires one to specify six quantities. Therefore, one needs to postulate the existence of two independently measurable fields, $\underline{E}$ and $\underline{B}$. Each of these fields should satisfy the two wave equations

$$
\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \underline{\nabla}^2 \right]\underline{E} = 0
\tag{35}
$$

and

$$
\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \underline{\nabla}^2 \right]\underline{B} = 0
\tag{36}
$$

The first-order equations must have the form

$$i \hbar \frac{\partial \underline{E}}{\partial t} = c \left(a \, \hat{\underline{p}} \wedge \underline{E} + b \, \hat{\underline{p}} \wedge \underline{B} \right)$$

$$i \hbar \frac{\partial \underline{B}}{\partial t} = c \left(d \, \hat{\underline{p}} \wedge \underline{B} + e \, \hat{\underline{p}} \wedge \underline{E} \right) \tag{37}$$

since the left-hand side is a vector, the right-hand side must also be a vector composed of the operator $\hat{p}$ and the wave functions. Like Newton's laws, these equations must be invariant under time-reversal invariance, $t \rightarrow -t$. The transformation leads to the identification

$$a = d = 0 \tag{38}$$

and

$$b = -e \tag{39}$$

if one also requires that one of the two fields changes sign under time reversal[2]. We shall adopt the convention that the field $\underline{E}$ retains its sign, so

$$\underline{E} \rightarrow \underline{E}$$

$$\underline{B} \rightarrow -\underline{B} \tag{40}$$

under time-reversal invariance. On taking the time derivative of the first equation, one obtains

$$-\hbar^2 \frac{\partial^2 \underline{E}}{\partial t^2} = -c^2 \, b^2 \, \hat{\underline{p}} \wedge \left(\hat{\underline{p}} \wedge \underline{E} \right) \tag{41}$$

Likewise, the $\underline{B}$ field is found to satisfy

$$-\hbar^2 \frac{\partial^2 \underline{B}}{\partial t^2} = -c^2 \, b^2 \, \hat{\underline{p}} \wedge \left(\hat{\underline{p}} \wedge \underline{B} \right) \tag{42}$$

Thus, one has found the two equations

$$-\hbar^2 \frac{\partial^2 \underline{E}}{\partial t^2} = -c^2 \, b^2 \left[-\hat{\underline{p}}^2 \, \underline{E} + \hat{\underline{p}} \left(\hat{\underline{p}} \cdot \underline{E} \right) \right] \tag{43}$$

and

$$-\hbar^2 \frac{\partial^2 \underline{B}}{\partial t^2} = -c^2 \, b^2 \left[-\hat{\underline{p}}^2 \, \underline{B} + \hat{\underline{p}} \left(\hat{\underline{p}} \cdot \underline{B} \right) \right] \tag{44}$$

On substituting the operator $\hat{\underline{p}} = -i \hbar \underline{\nabla}$, one obtains

$$\frac{\partial^2 \underline{E}}{\partial t^2} = -c^2 \, b^2 \left[-\underline{\nabla}^2 \, \underline{E} + \underline{\nabla} \left(\underline{\nabla} \cdot \underline{E} \right) \right] \tag{45}$$

---

[2]For the non-relativistic Schrödinger equation, time-reversal invariance implies that $t \rightarrow t' = -t$ and $\psi \rightarrow \psi' = \psi^*$.

13

and

$$\frac{\partial^2 \underline{B}}{\partial t^2} \;=\; -\,c^2\,b^2\left[\,-\,\underline{\nabla}^2\,\underline{B}\,+\,\underline{\nabla}\left(\,\underline{\nabla}\cdot\underline{B}\,\right)\right] \tag{46}$$

so $\hbar$ drops out. To reduce these equations to the form of wave equations, one needs to impose the conditions

$$\underline{\nabla}\cdot\underline{E} \;=\; 0 \tag{47}$$

and

$$\underline{\nabla}\cdot\underline{B} \;=\; 0 \tag{48}$$

On identifying the coefficients with those of the wave equation, one requires that

$$b^2 \;=\; 1 \tag{49}$$

Thus, one has arrived at the set of the source-free Maxwell's equations

$$
\begin{aligned}
\frac{1}{c}\,\frac{\partial \underline{E}}{\partial t} &\;=\; \underline{\nabla}\wedge\underline{B} \\
-\frac{1}{c}\,\frac{\partial \underline{B}}{\partial t} &\;=\; \underline{\nabla}\wedge\underline{E} \\
\underline{\nabla}\cdot\underline{E} &\;=\; 0 \\
\underline{\nabla}\cdot\underline{B} &\;=\; 0
\end{aligned}
\tag{50}
$$

which describe the one-particle Schrödinger equation for a massless spin-one particle, with the wave function $(\underline{E},\underline{B})$. These have a form which appears to be completely classical, since $\hbar$ has dropped out. Furthermore, in the absence of sources, Maxwell's equations are invariant under the symmetry transformation $(\underline{E},\underline{B}) \to (-\underline{B},\underline{E})$.

3 Maxwell's Equations

Classical Field Theories describe systems in which a very large number of particles are present. Measurements on systems containing very large numbers of particles are expected to result in average values, with only very small deviations. Hence, we expect that the subtleties of quantum measurements should be completely absent in systems that can be described as quantum fields. Classical Electromagnetism is an example of such a quantum field, in which an infinitely large number of photons are present.

In the presence of a current density $\underline{j}$ and a charge density ρ, Maxwell's equations assume the forms

$$\underline{\nabla}\wedge\underline{B} \;-\; \frac{1}{c}\,\frac{\partial \underline{E}}{\partial t} \;=\; \frac{4\,\pi}{c}\,\underline{j}$$

$$\underline{\nabla} \wedge \underline{E} + \frac{1}{c} \frac{\partial \underline{B}}{\partial t} = 0$$
$$\underline{\nabla} \cdot \underline{E} = 4 \pi \rho$$
$$\underline{\nabla} \cdot \underline{B} = 0 \tag{51}$$

The field equations ensure that the sources $\underline{j}$ and ρ satisfy a continuity equation.

Taking the divergence of the first equation and combining it with the time derivative of the third, one obtains

$$\underline{\nabla} \cdot \left(\underline{\nabla} \wedge \underline{B} \right) - \frac{1}{c} \left(\frac{\partial}{\partial t} \underline{\nabla} \cdot \underline{E} \right) = \frac{4 \pi}{c} \underline{\nabla} \cdot \underline{j}$$
$$- \frac{1}{c} \left(\frac{\partial}{\partial t} \underline{\nabla} \cdot \underline{E} \right) = \frac{4 \pi}{c} \underline{\nabla} \cdot \underline{j}$$
$$- \frac{4 \pi}{c} \frac{\partial \rho}{\partial t} = \frac{4 \pi}{c} \underline{\nabla} \cdot \underline{j} \tag{52}$$

Hence, one has derived the continuity equation

$$\frac{\partial \rho}{\partial t} + \underline{\nabla} \cdot \underline{j} = 0 \tag{53}$$

which shows that charge is conserved.

3.1 Vector and Scalar Potentials

Counting each component of Maxwell's equations separately, one arrives at eight equations for the six components of the unknown fields $\underline{E}$ and $\underline{B}$. As the equations are linear, this would over-determine the fields. Two of the eight equations must be regarded as self-consistency equations for the initial conditions on the fields.

One can solve the two source-free Maxwell equations, by expressing the electric $\underline{E}$ and magnetic fields $\underline{B}$ in terms of the vector $\underline{A}$ and scalar ϕ potentials, via

$$\underline{E} = -\frac{1}{c} \frac{\partial \underline{A}}{\partial t} - \underline{\nabla} \phi \tag{54}$$

and

$$\underline{B} = \underline{\nabla} \wedge \underline{A} \tag{55}$$

The expressions for $\underline{B}$ and $\underline{E}$ automatically satisfy the two source-free Maxwell's equations. This can be seen by examining

$$\underline{\nabla} \wedge \underline{E} + \frac{1}{c} \frac{\partial \underline{B}}{\partial t} = 0 \tag{56}$$

15

which, on substituting the expressions for the electromagnetic fields in terms of the vector and scalar potentials, becomes

$$\underline{\nabla} \wedge \left(-\frac{1}{c}\frac{\partial \underline{A}}{\partial t} - \underline{\nabla}\phi \right) + \frac{1}{c}\frac{\partial}{\partial t}\left(\underline{\nabla} \wedge \underline{A} \right) = 0 \qquad (57)$$

which is automatically satisfied since

$$\underline{\nabla} \wedge \left(\underline{\nabla}\phi \right) = 0 \qquad (58)$$

and the terms involving $\underline{A}$ cancel since $\underline{A}$ is analytic. The remaining source-free Maxwell equation is satisfied, since it has the form

$$\underline{\nabla} \cdot \underline{B} = 0 \qquad (59)$$

which reduces to

$$\underline{\nabla} \cdot \left(\underline{\nabla} \wedge \underline{A} \right) = 0 \qquad (60)$$

which is identically zero.

Therefore, the six components of $\underline{E}$ and $\underline{B}$ have been replaced by the four quantities $\underline{A}$ and ϕ. These four quantities are determined by the Maxwell equations which involve the sources, which are four in number.

The fields are governed by the set of non-trivial equations which relate $\underline{A}$ and ϕ to the sources $\underline{j}$ and ρ. When expressed in terms of $\underline{A}$ and ϕ, the remaining non-trivial Maxwell equations become

$$\underline{\nabla} \wedge \left(\underline{\nabla} \wedge \underline{A} \right) + \frac{1}{c}\frac{\partial}{\partial t}\left(\frac{1}{c}\frac{\partial \underline{A}}{\partial t} + \underline{\nabla}\phi \right) = \frac{4\pi}{c}\underline{j}$$

$$-\underline{\nabla} \cdot \left(\frac{1}{c}\frac{\partial \underline{A}}{\partial t} + \underline{\nabla}\phi \right) = 4\pi\rho \qquad (61)$$

but since

$$\underline{\nabla} \wedge \left(\underline{\nabla} \wedge \underline{A} \right) = \underline{\nabla}\left(\underline{\nabla} \cdot \underline{A} \right) - \underline{\nabla}^2 \underline{A} \qquad (62)$$

the pair of equations can be written as

$$\left(-\underline{\nabla}^2 \underline{A} + \frac{1}{c^2}\frac{\partial^2 \underline{A}}{\partial t^2} \right) + \underline{\nabla}\left(\underline{\nabla} \cdot \underline{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} \right) = \frac{4\pi}{c}\underline{j}$$

$$-\underline{\nabla}^2 \phi - \frac{1}{c}\frac{\partial}{\partial t}\left(\underline{\nabla} \cdot \underline{A} \right) = 4\pi\rho \qquad (63)$$

We shall make use of gauge invariance to simplify these equations.

3.2 Gauge Invariance

The vector and scalar potentials are defined as the solutions of the coupled partial differential equations describing the electric and magnetic fields

$$\underline{E} = -\frac{1}{c}\frac{\partial \underline{A}}{\partial t} - \underline{\nabla}\,\phi \tag{64}$$

and

$$\underline{B} = \underline{\nabla}\,\wedge\,\underline{A} \tag{65}$$

Hence, one expects that the solutions are only determined up to functions of integration. That is the vector and scalar potentials are not completely determined, even if the electric and magnetic fields are known precisely. It is possible to transform the vector and scalar potentials, in a way such that the $\underline{E}$ and $\underline{B}$ fields remain invariant. These transformations are known as gauge transformations of the second kind[3].

In particular, one can perform the transform

$$
\begin{aligned}
\underline{A} &\rightarrow \underline{A}' = \underline{A} - \underline{\nabla}\,\Lambda \\
\phi &\rightarrow \phi' = \phi + \frac{1}{c}\frac{\partial \Lambda}{\partial t}
\end{aligned} \tag{66}
$$

where Λ is an arbitrary analytic function and this transformation leaves the $\underline{E}$ and $\underline{B}$ fields invariant. The magnetic field is seen to be invariant since

$$
\begin{aligned}
\underline{B}' &= \underline{\nabla}\,\wedge\,\underline{A}' \\
&= \underline{\nabla}\,\wedge\,\left(\underline{A} - \underline{\nabla}\,\Lambda\right) \\
&= \underline{\nabla}\,\wedge\,\underline{A} \\
&= \underline{B}
\end{aligned} \tag{67}
$$

where the identity

$$\underline{\nabla}\,\wedge\,\left(\underline{\nabla}\,\Lambda\right) = 0 \tag{68}$$

valid for any scalar function Λ has been used. The electric field is invariant, since the transformed electric field is given by

$$\underline{E}' = -\frac{1}{c}\frac{\partial \underline{A}'}{\partial t} - \underline{\nabla}\,\phi'$$

---

[3]The transformation

$$
\begin{aligned}
\psi &\rightarrow \psi' = \psi\,\exp\left[\frac{i\,\chi}{\hbar}\right] \\
\underline{p} &\rightarrow \underline{\hat{p}}' = -\,i\,\hbar\,\underline{\nabla} - \underline{\nabla}\,\chi
\end{aligned}
$$

used in quantum mechanics is known as a gauge transformation of the first kind.

$$
\begin{aligned}
&= -\frac{1}{c}\frac{\partial}{\partial t}\left(\underline{A} - \underline{\nabla}\,\Lambda\right) - \underline{\nabla}\left(\phi + \frac{1}{c}\frac{\partial \Lambda}{\partial t}\right) \\
&= -\frac{1}{c}\frac{\partial \underline{A}}{\partial t} - \underline{\nabla}\,\phi \\
&= \underline{E}
\end{aligned}
\tag{69}
$$

In the above derivation, it has been noted that the order of the derivatives can be interchanged,

$$
\underline{\nabla}\,\frac{\partial \Lambda}{\partial t} = \frac{\partial}{\partial t}\,\underline{\nabla}\,\Lambda
\tag{70}
$$

since Λ is an analytic scalar function.

The gauge invariance allows us the freedom to impose a gauge condition which fixes the gauge. Two gauge conditions which are commonly used are the Lorenz gauge

$$
\underline{\nabla}\cdot\underline{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0
\tag{71}
$$

and the Coulomb or radiation gauge

$$
\underline{\nabla}\cdot\underline{A} = 0
\tag{72}
$$

The Lorenz gauge is manifestly Lorentz invariant, whereas the Coulomb gauge is frequently used in cases where the electrostatic interactions are important.

It is always possible to impose one or the other of these gauge conditions. If the vector and scalar potentials $(\phi, \underline{A})$ do not satisfy the gauge transformation, then one can perform a gauge transformation so that the transformed fields $(\phi', \underline{A}')$ satisfy the gauge condition.

For example, if the fields $(\phi, \underline{A})$ do not satisfy the Lorenz gauge condition, since

$$
\underline{\nabla}\cdot\underline{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = \chi(\underline{r}, t)
\tag{73}
$$

where χ is non-zero, then one can perform the gauge transformation to the new fields $(\phi', \underline{A}')$

$$
\begin{aligned}
\underline{\nabla}\cdot\underline{A}' + \frac{1}{c}\frac{\partial \phi'}{\partial t} &= \underline{\nabla}\cdot\underline{A} - \underline{\nabla}^2\,\Lambda + \frac{1}{c}\frac{\partial \phi}{\partial t} + \frac{1}{c^2}\frac{\partial^2 \Lambda}{\partial t^2} \\
&= \chi - \left(\underline{\nabla}^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\Lambda
\end{aligned}
\tag{74}
$$

The new fields satisfy the Lorentz condition if one chooses Λ to be the solution of the wave equation

$$
\left(\underline{\nabla}^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\Lambda = \chi(\underline{r}, t)
\tag{75}
$$

18

This can always be done, since the driven wave equation always has a solution. Hence, one can always insist that the fields satisfy the gauge condition

$$\underline{\nabla} \cdot \underline{A}' + \frac{1}{c} \frac{\partial \phi'}{\partial t} = 0 \tag{76}$$

Alternatively, if one is to impose the Coulomb gauge condition

$$\underline{\nabla} \cdot \underline{A}' = 0 \tag{77}$$

one can use Poisson's equations to show that one can always find a Λ such that the Coulomb gauge condition is satisfied[4].

In the Lorenz gauge, the equations of motion for the electromagnetic field are given by

$$\left[-\underline{\nabla}^2 + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \underline{A} = \frac{4\pi}{c} \underline{j}$$

$$\left[-\underline{\nabla}^2 + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \phi = 4\pi\rho \tag{78}$$

Hence, $\underline{A}$ and ϕ both satisfy the wave equation, where $\underline{j}$ and ρ are the sources. The solutions are waves which travel with velocity c.

In the Coulomb gauge, the fields satisfy the equations

$$\left[-\underline{\nabla}^2 + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \underline{A} = \frac{4\pi}{c} \underline{j} - \frac{1}{c} \frac{\partial}{\partial t} \underline{\nabla} \phi$$

$$-\underline{\nabla}^2 \phi = 4\pi\rho \tag{79}$$

The second equation is Poisson's equation and has solutions given by

$$\phi(\underline{r}, t) = \int d^3\underline{r}' \, \frac{\rho(\underline{r}', t)}{|\underline{r} - \underline{r}'|} \tag{80}$$

which is an "instantaneous" Coulomb interaction. However, the force from the electric field $\underline{E}$ is not transmitted instantaneously from $\underline{r}'$ to $\underline{r}$, since there is a term in the equation for $\underline{A}$ which compensates for the "instantaneous" interaction described by ϕ.

Exercise:

Consider the case of a uniform magnetic field of magnitude B which is oriented along the z-axis. Using the Coulomb gauge, find a general solution for the vector potential.

_____

[4]Imposing a gauge condition is insufficient to uniquely determine the vector potential $\underline{A}$, since in the case of the Coulomb gauge, the vector potential is only known up to the gradient of any harmonic function Λ.

4 Relativistic Formulation of Electrodynamics

Physical quantities can be classified as either being scalars, vectors or tensors according to how they behave under transformations. Scalars are invariant under Lorentz transformations, and all vectors transform in the same way.

4.1 Lorentz Scalars and Vectors

The space-time four-vector has components given by the time t and the three space coordinates $(x^{(1)}, x^{(2)}, x^{(3)})$ which label an event. The zeroth-component of the four-vector $x^{(0)}$ (the time component) is defined to be ct, where c is the velocity of light, in order that all the components have the dimensions of length. In Minkowski space, the four-vector is defined as having contravariant components $x^\mu = (ct, x^{(1)}, x^{(2)}, x^{(3)})$, while the covariant components are denoted by $x_\mu = (ct, -x^{(1)}, -x^{(2)}, -x^{(3)})$. The invariant length is given by the scalar product

$$x^\mu \, x_\mu \; = \; (\, c\,t\,)^2 \; - \; (\, x^{(1)} \,)^2 \; - \; (\, x^{(2)} \,)^2 \; - \; (\, x^{(3)} \,)^2 \tag{81}$$

where repeated indices are summed over. The invariant length $x^\mu \, x_\mu$ is related to the proper time τ. This definition can be generalized to the scalar product of two arbitrary four-vectors A^μ and B^μ as

$$A^\mu \, B_\mu \; = \; A^{(0)} \, B^{(0)} \; - \; A^{(1)} \, B^{(1)} \; - \; A^{(2)} \, B^{(2)} \; - \; A^{(3)} \, B^{(3)} \tag{82}$$

In special relativity, the four-vector scalar product can be written in terms of the product of the time-index components and the scalar product of the usual three-vectors as

$$A^\mu \, B_\mu \; = \; A^{(0)} \, B^{(0)} \; - \; \underline{A} \cdot \underline{B} \tag{83}$$

The Lorentz invariant four-vector scalar product can also be written as

$$A^\mu \, B_\mu \; = \; g_{\mu,\nu} \, A^\mu \, B^\nu \tag{84}$$

where $g_{\mu,\nu}$ is the Minkowski metric. These equations imply that

$$A_\mu \; = \; g_{\mu,\nu} \, A^\nu \tag{85}$$

That is, the metric tensor transforms contravariant components to covariant components. The Minkowski metric can be expressed as a four by four matrix

$$g_{\mu,\nu} \; = \; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{86}$$

where μ labels the rows and ν labels the columns. If the four-vectors are expressed as column-vectors

$$A^\nu = \begin{pmatrix} A^{(0)} \\ A^{(1)} \\ A^{(2)} \\ A^{(3)} \end{pmatrix} \tag{87}$$

and

$$A_\nu = \begin{pmatrix} A_{(0)} \\ A_{(1)} \\ A_{(2)} \\ A_{(3)} \end{pmatrix} \tag{88}$$

then the transformation from contravariant to covariant components can be expressed as

$$\begin{pmatrix} A_{(0)} \\ A_{(1)} \\ A_{(2)} \\ A_{(3)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} A^{(0)} \\ A^{(1)} \\ A^{(2)} \\ A^{(3)} \end{pmatrix} \tag{89}$$

The inverse transform is expressed as

$$A^\mu = g^{\mu,\nu} A_\nu \tag{90}$$

where, most generally, $g^{\mu,\nu}$ is the inverse metric

$$g^{\mu,\nu} = (g_{\mu,\nu})^{-1} \tag{91}$$

In our particular case of Cartesian coordinates in (flat) Minkowski space, the inverse metric coincides with the metric.

Example: Covariant and Contravariant Components of a Euclidean Vector

Consider a two-dimensional Euclidean space. However, instead of simply considering Cartesian axes, consider two unit vectors $\hat{e}_1$ and $\hat{e}_2$ which are not co-linear

$$\hat{e}_1 \cdot \hat{e}_2 = \cos\theta \tag{92}$$

A general vector $\underline{X}$ can be expressed in terms of its contravariant components $(X^{(1)}, X^{(2)})$ via

$$\underline{X} = X^{(1)} \hat{e}_1 + X^{(2)} \hat{e}_2 \tag{93}$$

The scalar product of two vectors $\underline{X}$ and $\underline{Y}$ can be evaluated by standard means in terms of their contravariant components. The result can be expressed in terms of a matrix product as

$$\underline{X} \cdot \underline{Y} = \begin{pmatrix} X^{(1)} & X^{(2)} \end{pmatrix} \begin{pmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{pmatrix} \begin{pmatrix} Y^{(1)} \\ Y^{(2)} \end{pmatrix} \tag{94}$$

21

involving the metric tensor $g_{\mu,\nu}$

$$\begin{pmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{pmatrix} \tag{95}$$

Likewise, the length $|\underline{X}|$ of the vector can be expressed as

$$|\underline{X}|^2 = \begin{pmatrix} X^{(1)} & X^{(2)} \end{pmatrix} \begin{pmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{pmatrix} \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} \tag{96}$$

The covariant components of the vector $(X_{(1)}, X_{(2)})$ are found from the contravariant components by the action of the metric tensor

$$\begin{pmatrix} X_{(1)} \\ X_{(2)} \end{pmatrix} = \begin{pmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{pmatrix} \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} \tag{97}$$

The covariant components of any vector can be found geometrically by dropping normals from the tip of the vector to the axes. The intersection of a normal with its corresponding axis determines the covariant component. For a Cartesian coordinate system, the covariant components are identical to the contravariant components.

A familiar example of the Lorentz invariant scalar product involves the momentum four-vector with contravariant components $p^\mu \equiv (\frac{E}{c}, p^{(1)}, p^{(2)}, p^{(3)})$ where E is the energy. The covariant components of the momentum four-vector are given by $p_\mu \equiv (\frac{E}{c}, -p^{(1)}, -p^{(2)}, -p^{(3)})$ and the scalar product defines the invariant mass m via

$$p^\mu \, p_\mu = \left(\frac{E}{c}\right)^2 - \underline{p}^2 = m^2 \, c^2 \tag{98}$$

Another scalar product which is frequently encountered is $p^\mu \, x_\mu$ which is given by

$$p^\mu \, x_\mu = E \, t - \underline{p} \cdot \underline{x} \tag{99}$$

This scalar product is frequently seen in the description of planes of constant phase of waves.

4.2 Covariant and Contravariant Derivatives

We shall now generalize the idea of the differential operator $\underline{\nabla}$ to Minkowski space. The generalization we consider, will have to be modified when the metric varies in space, i.e. when $g_{\mu,\nu}$ depends on the coordinates x^μ of the points in space.

Consider a scalar function $\phi(x^\mu)$ defined in terms of the contravariant coordinates x^μ. Under an infinitesimal translation a^μ

$$x^\mu \to x^{\mu\prime} = x^\mu + a^\mu \tag{100}$$

22

the scalar function $\phi(x^{\mu\prime})$ is still a scalar. Therefore, on performing a Taylor expansion, one has

$$\phi(x^{\mu} + a^{\mu}) \;=\; \phi(x^{\mu}) \;+\; a^{\mu} \, \frac{\partial}{\partial x^{\mu}} \, \phi(x^{\mu}) \;+\; \dots \qquad (101)$$

which is also a scalar. Therefore, the quantity

$$a^{\mu} \, \frac{\partial}{\partial x^{\mu}} \, \phi(x^{\mu}) \qquad (102)$$

is a scalar and can be interpreted as a scalar product between the contravariant vector displacement a^{μ} and the covariant gradient

$$\frac{\partial}{\partial x^{\mu}} \, \phi(x^{\mu}) \qquad (103)$$

The covariant gradient can be interpreted in terms of a covariant derivative

$$
\begin{aligned}
\partial_{\mu} \;&=\; \frac{\partial}{\partial x^{\mu}} \\
&=\; \left(\frac{1}{c} \frac{\partial}{\partial t} \, , \, \frac{\partial}{\partial x^{(1)}} , \, \frac{\partial}{\partial x^{(2)}} , \, \frac{\partial}{\partial x^{(3)}} \right) \\
&=\; \left(\frac{1}{c} \frac{\partial}{\partial t} \, , \, \underline{\nabla} \right) \qquad (104)
\end{aligned}
$$

Likewise, one can introduce the contravariant derivative as

$$
\begin{aligned}
\partial^{\mu} \;&=\; \frac{\partial}{\partial x_{\mu}} \\
&=\; \left(\frac{1}{c} \frac{\partial}{\partial t} \, , \, -\frac{\partial}{\partial x^{(1)}} , \, -\frac{\partial}{\partial x^{(2)}} , \, -\frac{\partial}{\partial x^{(3)}} \right) \\
&=\; \left(\frac{1}{c} \frac{\partial}{\partial t} \, , \, -\underline{\nabla} \right) \qquad (105)
\end{aligned}
$$

These covariant and contravariant derivative operators are useful in making relativistic transformational properties explicit. For example, if one defines the four-vector potential A^{μ} via

$$
\begin{aligned}
A^{\mu} \;&=\; \left(\phi \, , \, A^{(1)} \, , \, A^{(2)} \, , \, A^{(3)} \right) \\
&=\; \left(\phi \, , \, \underline{A} \right) \qquad (106)
\end{aligned}
$$

then the Lorenz gauge condition can be expressed as

$$\partial_{\mu} A^{\mu} \;=\; \frac{1}{c} \frac{\partial \phi}{\partial t} \;+\; \underline{\nabla} \cdot \underline{A} \;=\; 0 \qquad (107)$$

23

which is of the form of a Lorentz scalar. Likewise, if one introduces the current density four-vector j^μ with contravariant components

$$
\begin{aligned}
j^\mu &= \left(c\,\rho \; , \; j^{(1)} \; , \; j^{(2)} \; , \; j^{(3)} \right) \\
&= \left(c\,\rho \; , \; \underline{j} \right)
\end{aligned}
\tag{108}
$$

then the condition for conservation of charge can be written as

$$
\begin{aligned}
\frac{\partial \rho}{\partial t} + \underline{\nabla} \cdot \underline{j} &= 0 \\
\partial_\mu j^\mu &= 0
\end{aligned}
\tag{109}
$$

which is a Lorentz scalar. Also, the gauge transformation can also be compactly expressed in terms of a transformation of the contravariant vector potential

$$
A^\mu \;\rightarrow\; A^{\mu\prime} \;=\; A^\mu \;+\; \partial^\mu \Lambda
\tag{110}
$$

where Λ is an arbitrary scalar function. The gauge transformation reduces to

$$
\begin{aligned}
\phi \;\rightarrow\; \phi' &= \phi + \frac{1}{c}\,\frac{\partial \Lambda}{\partial t} \\
\underline{A} \;\rightarrow\; \underline{A}' &= \underline{A} - \underline{\nabla}\,\Lambda
\end{aligned}
\tag{111}
$$

Similarly, one can use the contravariant notation to express the quantization conditions

$$
\begin{aligned}
E &\;\rightarrow\; i\,\hbar\,\frac{\partial}{\partial t} \\
\underline{p} &\;\rightarrow\; -\,i\,\hbar\,\underline{\nabla}
\end{aligned}
\tag{112}
$$

in the form

$$
p^\mu \;\rightarrow\; i\,\hbar\,\partial^\mu
\tag{113}
$$

One can also express the wave equation operator in terms of the scalar product of the contravariant and covariant derivative operators

$$
\partial^\mu\,\partial_\mu \;=\; \frac{1}{c^2}\,\frac{\partial^2}{\partial t^2} \;-\; \underline{\nabla}^2
\tag{114}
$$

Hence, in the Lorenz gauge, the equations of motion for the four-vector potential A^μ can be expressed concisely as

$$
\partial^\nu \partial_\nu A^\mu \;=\; \frac{4\,\pi}{c}\,j^\mu
\tag{115}
$$

However, these equations are not gauge invariant.

4.3 Lorentz Transformations

A Lorentz transform can be defined as any transformation which leaves the scalar product of two four-vectors invariant. Under a Lorentz transformation, an arbitrary four-vector A^μ is transformed to $A^{\mu\prime}$, via

$$A^{\mu\prime} = \Lambda^\mu{}_\nu A^\nu \tag{116}$$

where the repeated index ν is summed over. The inverse transformation is represented by

$$A^\mu = (\Lambda^{-1})^\mu{}_\nu A^{\nu\prime} \tag{117}$$

Since the scalar product is to be relativistically invariant, one requires that

$$A^{\mu\prime} B_\mu{}' = A^\mu B_\mu \tag{118}$$

The left hand-side is evaluated as

$$A^{\mu\prime} B_\mu{}' = \Lambda^\mu{}_\nu A^\nu g_{\mu,\sigma} \Lambda^\sigma{}_\tau B^\tau \tag{119}$$

If the scalar product is to be invariant, the transform must satisfy the condition

$$g_{\nu,\tau} = \Lambda^\mu{}_\nu g_{\mu,\sigma} \Lambda^\sigma{}_\tau \tag{120}$$

If this condition is satisfied, then $\Lambda^\mu{}_\nu$ is a Lorentz transformation.

Like the metric tensor, the Lorentz transformation can be expressed as a four by four matrix

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{pmatrix} \tag{121}$$

where μ labels the rows and ν labels the columns. In terms of the matrices, the condition that Λ is a Lorentz transformation can be written as

$$g = \Lambda^T g \Lambda \tag{122}$$

where Λ^T is the transpose of the matrix Λ, i.e.

$$(\Lambda^T)^\nu{}_\mu = \Lambda^\mu{}_\nu \tag{123}$$

A specific transformation, which is the transformation from a stationary frame to a reference frame moving along the $x^{(1)}$ axis with velocity v, is represented by the matrix

$$\Lambda^\mu{}_\nu = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \begin{pmatrix} 1 & -\frac{v}{c} & 0 & 0 \\ -\frac{v}{c} & 1 & 0 & 0 \\ 0 & 0 & \sqrt{1 - \frac{v^2}{c^2}} & 0 \\ 0 & 0 & 0 & \sqrt{1 - \frac{v^2}{c^2}} \end{pmatrix} \tag{124}$$
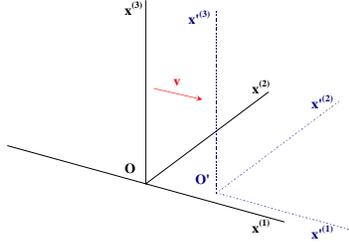
Figure 5: Two inertial frames of reference moving with a constant relative velocity with respect to each other.

which can be seen to satisfy the condition

$$g_{\nu,\tau} \;=\; \Lambda^{\mu}{}_{\nu} \; g_{\mu,\sigma} \; \Lambda^{\sigma}{}_{\tau} \tag{125}$$

which has to be satisfied if $\Lambda^{\mu}{}_{\nu}$ is to represent a Lorentz transformation.
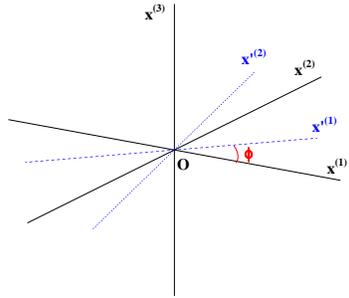
Figure 6: Two inertial frames of reference rotated with respect to each other.

Likewise, the rotation through an angle φ about the $x^{(3)}$-axis represented by

$$\Lambda^{\mu}{}_{\nu} \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi & 0 \\ 0 & -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{126}$$

is a Lorentz transformation, since it also satisfies the condition eqn(120).

Since the boost velocity v and the angles of rotation φ are continuous, one could consider transformations where these quantities are infinitesimal. Such infinitesimal transformations can be expanded as

$$\Lambda^{\mu}{}_{\nu} \;=\; \delta^{\mu}{}_{\nu} \;+\; \epsilon^{\mu}{}_{\nu} \;+\; \ldots \tag{127}$$

26

where $\delta^\mu{}_\nu$ is the Kronecker delta function representing the identity transformation[5] and $\epsilon^\mu{}_\nu$ is a matrix which is first-order in the infinitesimal parameter. The condition on $\epsilon^\mu{}_\nu$ required for $\Lambda^\mu{}_\nu$ to be a Lorentz transform is given by

$$0 \;=\; \epsilon^\mu{}_\nu \; g_{\mu,\tau} \;+\; g_{\nu,\sigma} \; \epsilon^\sigma{}_\tau \tag{129}$$

or, on using the metric tensor to lower the indices, one has

$$\epsilon_{\tau,\nu} \;=\; -\; \epsilon_{\nu,\tau} \tag{130}$$

Hence, an arbitrary infinitesimal Lorentz transformation is represented by an arbitrary antisymmetric 4×4 matrix $\epsilon_{\tau,\nu}$. This matrix occurs in the expression for the transformation matrix

$$\Lambda_{\tau,\nu} \;=\; g_{\tau,\nu} \;+\; \epsilon_{\tau,\nu} \;+\; \ldots \tag{131}$$

which transforms the contravariant components of a vector into covariant components. It follows that, if ν and τ are either both space-indices or are both time-indices, the components of the finite Lorentz transformation matrix $\Lambda^\tau{}_\nu$ are antisymmetric on interchanging τ and ν. Whereas if the pair of indices ν and τ are mixed space and time-indices, the components of the transformation matrix $\Lambda^\tau{}_\nu$ are symmetric.

Exercise:

Show that a Lorentz transformation from the unprimed rest frame to the primed reference frame moving along the $x^{(3)}$-axis with constant velocity v, can be considered as a rotation through an imaginary angle $\theta \;=\; i\,\chi$ in space-time, where $i\,c\,t$ plays the role of a spatial coordinate. Find the equation that determines χ.

4.4   Invariant Form of Maxwell's Equations

In physics, one strives to write the fundamental equations in forms which are independent of arbitrary choices, such as the coordinate system or the choice of gauge condition. However, in particular applications it is expedient to choose the coordinate system and gauge condition in ways that highlight the symmetries and simplify the mathematics.

We shall introduce an antisymmetric field tensor $F^{\mu,\nu}$ which is gauge invariant. That is, the form of $F^{\mu,\nu}$ is independent of the choice of gauge. We

---

[5]The student more adept in index-gymnastics may consider the advantages and disadvantages of replacing the Kronecker delta function $\delta^\mu{}_\nu$ by $g^\mu{}_\nu$, since

$$\delta^\mu{}_\nu \;\equiv\; g^\mu{}_\nu \;=\; g^{\mu,\rho} \, g_{\rho,\nu} \;=\; \delta_{\mu,\nu} \tag{128}$$

shall express the six $(\frac{(16-4)}{2} = 6)$ independent components of the antisymmetric tensor in terms of the four-vector potential A^μ and the contravariant derivative as

$$F^{\mu,\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{132}$$

so the tensor is antisymmetric

$$F^{\mu,\nu} = -F^{\nu,\mu} \tag{133}$$

It is immediately obvious that $F^{\mu,\nu}$ is invariant under gauge transformations

$$A^\mu \rightarrow A^{\mu\prime} = A^\mu + \partial^\mu \Lambda \tag{134}$$

since

$$\partial^\mu \partial^\nu \Lambda - \partial^\nu \partial^\mu \Lambda \equiv 0 \tag{135}$$

Alternatively, explicit evaluation of $F^{\mu,\nu}$ shows that the six independent components can be expressed in terms of the electric and magnetic fields, which are gauge invariant. Components of the field tensor are explicitly evaluated from the definition as

$$
\begin{aligned}
F^{0,1} &= \frac{1}{c}\frac{\partial}{\partial t} A^{(1)} - \frac{\partial}{\partial x_1} \phi \\
&= \frac{1}{c}\frac{\partial}{\partial t} A^{(1)} + \frac{\partial}{\partial x^{(1)}} \phi \\
&= -E^{(1)}
\end{aligned}
\tag{136}
$$

and

$$
\begin{aligned}
F^{1,2} &= \frac{\partial}{\partial x_1} A^{(2)} - \frac{\partial}{\partial x_2} A^{(1)} \\
&= -\frac{\partial}{\partial x^{(1)}} A^{(2)} + \frac{\partial}{\partial x^{(2)}} A^{(1)} \\
&= -B^{(3)}
\end{aligned}
\tag{137}
$$

The non-zero components of the field tensor are related to the spatial components (i, j, k) of the electromagnetic field by

$$F^{i,0} = E^{(i)} \tag{138}$$

and

$$F^{i,j} = -\xi^{i,j,k} B^{(k)} \tag{139}$$

where $\xi^{i,j,k}$ is the Levi-Civita symbol. Therefore, the field tensor can be expressed as the matrix

$$
F^{\mu,\nu} = \begin{pmatrix}
0 & -E^{(1)} & -E^{(2)} & -E^{(3)} \\
E^{(1)} & 0 & -B^{(3)} & B^{(2)} \\
E^{(2)} & B^{(3)} & 0 & -B^{(1)} \\
E^{(3)} & -B^{(2)} & B^{(1)} & 0
\end{pmatrix}
\tag{140}
$$

28

Maxwell's equations can be written in terms of the field tensor as

$$\partial_\nu F^{\nu,\mu} \;=\; \frac{4\,\pi}{c}\,j^\mu \tag{141}$$

For $\mu = i$, the field equations become

$$\frac{1}{c}\,\frac{\partial}{\partial t}\,F^{0,i} \;+\; \frac{\partial}{\partial x^j}\,F^{j,i} \;=\; \frac{4\,\pi}{c}\,j^{(i)}$$

$$-\frac{1}{c}\,\frac{\partial}{\partial t}\,E^{(i)} \;+\; \xi^{i,j,k}\,\frac{\partial}{\partial x^j}\,B^{(k)} \;=\; \frac{4\,\pi}{c}\,j^{(i)}$$

$$-\frac{1}{c}\,\frac{\partial}{\partial t}\,E^{(i)} \;+\; \left(\underline{\nabla}\wedge\underline{B}\right)^{(i)} \;=\; \frac{4\,\pi}{c}\,j^{(i)} \tag{142}$$

while for $\mu = 0$ the equations reduce to

$$\frac{\partial}{\partial x^j}\,F^{j,0} \;=\; \frac{4\,\pi}{c}\,j^{(0)}$$

$$\frac{\partial}{\partial x^j}\,E^{(j)} \;=\; 4\,\pi\,\rho$$

$$\left(\underline{\nabla}\cdot\underline{E}\right) \;=\; 4\,\pi\,\rho \tag{143}$$

since $F^{0,0}$ vanishes. The above field equations are the two Maxwell's equations which involve the sources of the fields. The remaining two sourceless Maxwell equations are expressed in terms of the antisymmetric field tensor as

$$\partial_\mu F_{\nu,\rho} \;+\; \partial_\rho F_{\mu,\nu} \;+\; \partial_\nu F_{\rho,\mu} \;=\; 0 \tag{144}$$

where the indices are permuted cyclically. These internal equations reduce to

$$\underline{\nabla}\cdot\underline{B} \;=\; 0 \tag{145}$$

when μ, ν and ρ are the space indices $(1,2,3)$. When one index taken from the set (μ,ν,ρ) is the time index, and the other two are different space indices, the field equations reduce to

$$\frac{1}{c}\,\frac{\partial \underline{B}}{\partial t} \;+\; \underline{\nabla}\wedge\underline{E} \;=\; 0 \tag{146}$$

If two indices are repeated, the above equations are satisfied identically, due to the antisymmetry of the field tensor.

Alternatively, when expressed in terms of the vector potential, the field equations of motion are equivalent to the wave equations

$$\partial_\nu\partial^\nu A^\mu \;-\; \partial^\mu\left(\partial_\nu A^\nu\right) \;=\; \frac{4\,\pi}{c}\,j^\mu \tag{147}$$

29

Since four-vectors A^μ and j^μ transform as

$$
\begin{aligned}
A^{\mu\prime} &= \Lambda^\mu{}_\nu \, A^\nu \\
j^{\mu\prime} &= \Lambda^\mu{}_\nu \, j^\nu
\end{aligned}
\tag{148}
$$

and likewise for the contravariant derivative

$$
\partial^{\mu\prime} = \Lambda^\mu{}_\nu \, \partial^\nu
\tag{149}
$$

then one can conclude that the field tensor transforms as

$$
F^{\mu,\nu\prime} = \Lambda^\mu{}_\sigma \, \Lambda^\nu{}_\tau \, F^{\sigma,\tau}
\tag{150}
$$

This shows that, under a Lorentz transform, the electric and magnetic fields $(\underline{E}, \underline{B})$ transform into themselves.

Exercise:

Show explicitly, how the components of the electric and magnetic fields change, when the coordinate system is transformed from the unprimed reference frame to a primed reference frame which is moving along the $x^{(3)}$-axis with constant velocity v.

5 The Simplest Classical Field Theory

Consider a string stretched along the x-axis, which can support motion in the y-direction. We shall consider the string to be composed of mass elements $m_i = \rho\, a$, that have fixed x-coordinates denoted by x_i and are separated by a distance a. The mass elements can be displaced along the y-axis. The y-coordinate of the i-th mass element is denoted by y_i. We shall assume that the string satisfies the spatial boundary conditions at each end. We shall assume that the string satisfies periodic boundary conditions, so that $y_0 = y_{N+1}$.

The Lagrangian for the string is a function of the coordinates y_i and the velocities $\frac{dy_i}{dt}$. The Lagrangian is given by

$$
L = \sum_{i=1}^{i=N} \left[\frac{m_i}{2} \left(\frac{dy_i}{dt} \right)^2 - \frac{\kappa_i}{2} \left(y_i - y_{i-1} \right)^2 \right]
\tag{151}
$$

The first term represents the kinetic energy of the mass elements, and the second term represents the increase in the elastic potential energy of the section of the string between the i-th and $(i-1)$-th element as the string is stretched from its equilibrium position. This follows since, Δs_i the length of the section of string between mass element i and $i-1$ in a non-equilibrium position is given by

$$
\begin{aligned}
\Delta s_i^2 &= (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 \\
&= a^2 + (y_i - y_{i-1})^2
\end{aligned}
\tag{152}
$$
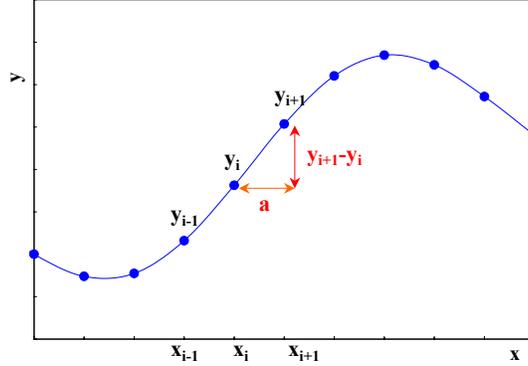
Figure 7: A string composed of a discrete set of particles of masses m_i separated by a distance a along the x-axis. The particles can be moved from their equilibrium positions by displacements y_i transverse to the x-axis.

since the x-coordinates are fixed. Thus, if one assumes that the spring constant for the stretched string segment is κ_i, then the potential energy of the segment is given by

$$V_i \;=\; \frac{\kappa_i}{2} \, (\; y_i \;-\; y_{i-1}\;)^2 \tag{153}$$

We shall consider the case of a uniform string for which $\kappa_i \;=\; \kappa$ for all i.

The equations of motion are obtained by minimizing the action S which is defined as the integral

$$S \;=\; \int_0^T \; dt \; L \tag{154}$$

between an initial configuration at time 0 and a final configuration at time T. The action is a functional of the coordinates y_i and the velocities $\frac{dy_i}{dt}$, which are to be evaluated for arbitrary functions $y_i(t)$. The string follows the trajectory $y_i^{ex}(t)$ which minimizes the action, which travels between the fixed initial value $y_i(0)$ and the final value $y_i(T)$. We shall represent the deviation of an arbitrary trajectory $y_i(t)$ from the extremal trajectory by $\delta y_i(t)$, then

$$\delta y_i(t) \;=\; y_i(t) \;-\; y_i^{ex}(t) \tag{155}$$

The action can be expanded in powers of the deviations δy_i as

$$S \;=\; S_0 \;+\; \delta^1 S \;+\; \delta^2 S \;+\; \ldots \tag{156}$$

where S_0 is the action evaluated for the extremal trajectories. The first-order deviation found by varying δy_i is given by

$$\delta^1 S \;=\; \int_0^T \; dt \; \sum_{i=1}^{i=N} \left[\; m_i \left(\frac{d\delta y_i}{dt} \right) \left(\frac{dy_i^{ex}}{dt} \right) - \kappa \, \delta y_i \left(\; y_i^{ex} - y_{i-1}^{ex} \; \right) + \kappa \, \delta y_i \left(\; y_{i+1}^{ex} - y_i^{ex} \; \right) \right] \tag{157}$$

31

in which $y_i(T)$ and $\frac{dy_i}{dt}$ are to be evaluated for the extremal trajectory. Since the trajectory which the string follows minimizes the action, the term $\delta^1 S$ must vanish for an arbitrary variation δy_i. We can eliminate the time derivative of the deviation by integrating by parts with respect to t. This yields

$$\delta^1 S \;=\; \int_0^T dt \sum_{i=1}^{i=N} \left[\; - \; m_i \, \delta y_i \, \frac{d}{dt}\left(\frac{dy_i^{ex}}{dt}\right) \; - \; \kappa \, \delta y_i \left(\; y_i^{ex} \; - \; y_{i-1}^{ex} \; \right) \; + \; \kappa \, \delta y_i \left(\; y_{i+1}^{ex} \; - \; y_i^{ex} \; \right) \right]$$
$$+ \; \sum_i \; m_i \; \delta y_i(t) \left(\frac{dy_i^{ex}}{dt} \right) \Bigg|_0^T \tag{158}$$

The boundary term vanishes since the initial and final configurations are fixed, so

$$\delta y_i(T) \;=\; \delta y_i(0) \;=\; 0 \tag{159}$$

Hence the first-order variation of the action reduces to

$$\delta^1 S \;=\; \int_0^T dt \sum_{i=1}^{i=N} \; \delta y_i \left[\; - \; m_i \, \frac{d}{dt}\left(\frac{dy_i^{ex}}{dt}\right) \; - \; \kappa \left(\; 2 \, y_i^{ex} \; - \; y_{i-1}^{ex} \; - \; y_{i+1}^{ex} \; \right) \right] \tag{160}$$

The linear variation of the action vanishes for an arbitrary $\delta y_i(t)$, if the term in the square brackets vanishes

$$m_i \, \frac{d}{dt}\left(\frac{dy_i^{ex}}{dt}\right) \; + \; \kappa \left(\; 2 \, y_i^{ex} \; - \; y_{i-1}^{ex} \; - \; y_{i+1}^{ex} \; \right) \;=\; 0 \tag{161}$$

Thus, out of all possible trajectories, the physical trajectory $y_i^{ex}(t)$ is determined by the equation of motion

$$m_i \, \frac{d}{dt}\left(\frac{dy_i}{dt}\right) \;=\; - \; \kappa \left(\; 2 \, y_i \; - \; y_{i-1} \; - \; y_{i+1} \; \right) \tag{162}$$

The momentum p_i which is canonically conjugate to y_i is determined by

$$p_i \;=\; \left(\frac{\partial L}{\partial(\frac{dy_i}{dt})} \right) \tag{163}$$

which yields the momentum as

$$p_i \;=\; m_i \, \frac{dy_i}{dt} \tag{164}$$

The Hamiltonian is defined as the Legendre transform of L, so

$$H \;=\; \sum_i \; p_i \, \frac{dy_i}{dt} \; - \; L \tag{165}$$

The Hamiltonian is only a function of the pairs of canonically conjugate momenta p_i and coordinates y_i. This can be seen, considering infinitesimal changes in y_i, $\frac{dy_i}{dt}$ and p_i. The resulting infinitesimal change in the Hamiltonian dH is expressed as

$$
\begin{aligned}
dH &= \sum_{i=1}^{i=N} \left[dp_i \, \frac{dy_i}{dt} \; + \; p_i \, d(\frac{dy_i}{dt}) \; - \; \left(\frac{\partial L}{\partial (\frac{dy_i}{dt})} \right) d(\frac{dy_i}{dt}) \; - \; \left(\frac{\partial L}{\partial y_i} \right) dy_i \right] \\
&= \sum_{i=1}^{i=N} \left[dp_i \, \frac{dy_i}{dt} \; - \; \left(\frac{\partial L}{\partial y_i} \right) dy_i \right]
\end{aligned}
\tag{166}
$$

since, the terms proportional to the infinitesimal change $d(\frac{dy_i}{dt})$ vanish identically, due to the definition of p_i. From this, one finds

$$
\frac{\partial H}{\partial p_i} \; = \; \frac{dy_i}{dt}
\tag{167}
$$

and

$$
\frac{\partial H}{\partial y_i} \; = \; - \left(\frac{\partial L}{\partial y_i} \right)
\tag{168}
$$

Therefore, the Hamiltonian is only a function of the pairs of canonically conjugate variables p_i and y_i. The Hamiltonian is given by

$$
H \; = \; \sum_{i=1}^{i=N} \left[\frac{p_i^2}{2 \, m_i} \; + \; \frac{\kappa_i}{2} \left(y_i \; - \; y_{i-1} \right)^2 \right]
\tag{169}
$$

When expressed in terms of the Hamiltonian, the equations of motion have the form

$$
\begin{aligned}
\frac{dy_i}{dt} &= \frac{\partial H}{\partial p_i} \\
\frac{dp_i}{dt} &= - \frac{\partial H}{\partial y_i}
\end{aligned}
\tag{170}
$$

The Hamilton equations of motion reduce to

$$
\begin{aligned}
\frac{dy_i}{dt} &= \frac{p_i}{m_i} \\
\frac{dp_i}{dt} &= - \kappa_i \left(y_i \; - \; y_{i-1} \right) + \kappa_{i+1} \left(y_{i+1} \; - \; y_i \right)
\end{aligned}
\tag{171}
$$

for each i value $N \geq i \geq 1$.

One can define the Poisson brackets of two arbitrary quantities A and B in terms of derivatives with respect to the canonically conjugate variables

$$
\left\{ A \, , \, B \right\} \; = \; \sum_{i=1}^{i=N} \left[\frac{\partial A}{\partial y_i} \frac{\partial B}{\partial p_i} \; - \; \frac{\partial B}{\partial y_i} \frac{\partial A}{\partial p_i} \right]
\tag{172}
$$

The Poisson bracket is antisymmetric in A and B

$$\left\{ A \, , \, B \right\} = - \left\{ B \, , \, A \right\} \tag{173}$$

The Poisson brackets of the canonically conjugate variables are given by

$$\left\{ p_i \, , \, y_j \right\} = - \delta_{i,j} \tag{174}$$

and

$$\left\{ p_i \, , \, p_j \right\} = \left\{ y_i \, , \, y_j \right\} = 0 \tag{175}$$

We shall show how energy is conserved by considering a finite segment of the string. For example, we shall consider the segment of the string consisting of the i-th mass element and the string which connects the i-th and $(i-1)$-th mass element. The energy of this segment will be described by $\mathcal{H}_i$, where

$$\mathcal{H}_i = \frac{\rho \, a}{2} \left(\frac{dy_i}{dt} \right)^2 + \frac{\kappa}{2} \left(y_i - y_{i-1} \right)^2 \tag{176}$$

The rate of increase energy in this segment is given by

$$\frac{d\mathcal{H}_i}{dt} = \rho \, a \left(\frac{d^2 y_i}{dt^2} \right) \left(\frac{dy_i}{dt} \right) + \kappa \left(\frac{dy_i}{dt} \right) \left(y_i - y_{i-1} \right) - \kappa \left(\frac{dy_{i-1}}{dt} \right) \left(y_i - y_{i-1} \right) \tag{177}$$

One can use the equation of motion to eliminate the acceleration term, leading to

$$\frac{d\mathcal{H}_i}{dt} = \kappa \left(\frac{dy_i}{dt} \right) \left(y_{i+1} - y_i \right) - \kappa \left(\frac{dy_{i-1}}{dt} \right) \left(y_i - y_{i-1} \right) \tag{178}$$

The increase in energy of this segment, per unit time, is clearly given by the difference of the quantity

$$\mathcal{P}_i = - \kappa \left(\frac{dy_i}{dt} \right) \left(y_{i+1} - y_i \right) \tag{179}$$

at the front end of the segment and $\mathcal{P}_{i-1}$ at the back end of the segment. Since, from continuity of energy, the rate of increase in the energy of the segment must equal the net inflow of energy into the segment, one can identify $\mathcal{P}_i$ as the flux of energy flowing out of the i-th into the $(i+1)$-th segment.

34

5.1 The Continuum Limit

The displacement of each element of the string can be expressed as a function of its position, via

$$y_i = y(x_i) \tag{180}$$

where each segment has length a, so that $x_{i+1} = x_i + a$. The displacement $y(x_{i+1})$ can be Taylor expanded about x_i as

$$y(x_{i+1}) = y(x_i) + a \left. \frac{\partial y}{\partial x} \right|_{x_i} + \frac{a^2}{2!} \left. \frac{\partial^2 y}{\partial x^2} \right|_{x_i} + \ldots \tag{181}$$

We intend to take the limit $a \to 0$, so that only the first few terms of the series need to be retained. The summations over i are to be replaced by integrations

$$\sum_{i=1}^{N} \to \frac{1}{a} \int_0^L dx \tag{182}$$

The tension in the string $\mathcal{T}$ is given by

$$\mathcal{T} = \kappa\, a \tag{183}$$

and this has to be kept constant when the limit $a \to 0$ is taken.

In the continuum limit, the Lagrangian L can be expressed as an integral of the Lagrangian density $\mathcal{L}$ as

$$L = \int_0^L dx\, \mathcal{L} \tag{184}$$

where

$$\mathcal{L} = \frac{1}{2} \left[\rho \left(\frac{dy}{dt} \right)^2 - \kappa\, a \left(\frac{\partial y}{\partial x} \right)^2 \right] \tag{185}$$

The equations of motion are found from the extrema of the action

$$S = \int_0^T dt \int_0^L dx\, \mathcal{L} \tag{186}$$

It should be noted, that in S time and space are treated on the same footing and that $\mathcal{L}$ is a scalar quantity.

In the continuum limit, the Hamiltonian is given by

$$H = \int_0^L dx\, \mathcal{H} \tag{187}$$

where the Hamiltonian density $\mathcal{H}$ is given by

$$\mathcal{H} \; = \; \frac{1}{2} \left[\rho \left(\frac{dy}{dt} \right)^2 \; + \; \kappa \, a \left(\frac{\partial y}{\partial x} \right)^2 \right] \tag{188}$$

and the energy flux $\mathcal{P}$ is given by

$$\mathcal{P} \; = \; - \, \kappa \, a \left(\frac{dy}{dt} \right) \left(\frac{\partial y}{\partial x} \right) \tag{189}$$

The condition of conservation of energy is expressed as the continuity equation

$$\frac{d\mathcal{H}}{dt} \; + \; \frac{\partial \mathcal{P}}{\partial x} \; = \; 0 \tag{190}$$

5.2 Normal Modes

The solutions of the equations of motion are of the form of a (real) superposition of plane waves

$$\psi_k(x) \; = \; \frac{1}{\sqrt{L}} \, \exp \left[\, i \, (\, k \, x \; - \; \omega \, t \,) \, \right] \tag{191}$$

The above expression satisfies the wave equation if the frequency ω satisfies the dispersion relation

$$\omega_k^2 \; = \; v^2 \, k^2 \tag{192}$$

The above dispersion relation yields both positive and negative frequency solutions. If the plane-waves are to satisfy periodic boundary conditions, k must be quantized so that

$$k_n \; = \; \frac{2 \, \pi}{L} \, n \tag{193}$$

for integer n. The positive-frequency solutions shall be written as

$$\psi_k(x) \; = \; \frac{1}{\sqrt{L}} \, \exp \left[\, i \, (\, k_n \, x \; - \; \omega_n \, t \,) \, \right] \tag{194}$$

and the negative frequency solutions as

$$\psi_{-k}^*(x) \; = \; \frac{1}{\sqrt{L}} \, \exp \left[\, i \, (\, k_n \, x \; + \; \omega_n \, t \,) \, \right] \tag{195}$$

These solutions form an orthonormal set since

$$\int_0^L \, dx \, \psi_{k'}^*(x) \, \psi_k(x) \; = \; \delta_{k',k} \tag{196}$$

Hence, a general solution can be written as

$$y(x) \; = \; \sum_k \left(\, c_k(0) \, \psi_k(x) \; + \; c_{-k}^*(0) \, \psi_{-k}^*(x) \, \right) \tag{197}$$

36

where the c_k are arbitrary complex numbers that depend on k. If the time dependence of the $\psi_k(x)$ is absorbed into the complex functions c_k via

$$c_k(t) \; = \; c_k(0) \, \exp\left[\, - \, i \, \omega_k \, t \, \right] \qquad (198)$$

then one has

$$y(x) \; = \; \frac{1}{\sqrt{L}} \, \sum_k \left(\, c_k(t) \, + \, c_{-k}^*(t) \, \right) \, \exp\left[\, i \, k \, x \, \right] \qquad (199)$$

which is purely real. Thus, the field $y(x)$ is determined by the amplitudes of the normal modes, i.e. by $c_k(t)$. The time-dependent amplitude $c_k(t)$ satisfies the equation of motion

$$\frac{d^2 c_k}{dt^2} \; = \; - \, \omega_k^2 \, c_k \qquad (200)$$

and, therefore, behaves like a classical harmonic oscillator. To quantize this classical field theory, one needs to quantize these harmonic oscillators.

The Hamiltonian is expressed as

$$H \; = \; \frac{1}{2 \, a} \, \int_0^L \, dx \, \left[\, \frac{1}{\rho \, a} \, p(x)^2 \, + \, \kappa \, a^2 \, \left(\frac{\partial y(x)}{\partial x} \right)^2 \, \right] \qquad (201)$$

On substituting $y(x)$ in the form

$$y(x) \; = \; \frac{1}{\sqrt{L}} \, \sum_k \left(\, c_k(t) \, + \, c_{-k}^*(t) \, \right) \, \exp\left[\, i \, k \, x \, \right] \qquad (202)$$

and

$$p(x) \; = \; \frac{\rho \, a}{\sqrt{L}} \, \sum_k \left(\, \frac{dc_k}{dt} \, + \, \frac{dc_{-k}^*}{dt} \, \right) \, \exp\left[\, i \, k \, x \, \right] \qquad (203)$$

then after integrating over x, one finds that the energy has the form

$$H \; = \; \frac{\rho}{2} \, \sum_k \left(\frac{dc_{-k}}{dt} \, + \, \frac{dc_k^*}{dt} \right) \left(\frac{dc_k}{dt} \, + \, \frac{dc_{-k}^*}{dt} \right)$$
$$+ \, \frac{\kappa \, a}{2} \, \sum_k \, k^2 \left(\, c_{-k}(t) \, + \, c_k^*(t) \, \right) \left(\, c_k(t) \, + \, c_{-k}^*(t) \, \right) \qquad (204)$$

Furthermore, on using the time-dependence of the Fourier coefficients $c_k(t)$, one has

$$H \; = \; - \, \frac{\rho}{2} \, \sum_k \, \omega_k^2 \left(\, c_{-k}(t) \, - \, c_k^*(t) \, \right) \left(\, c_k(t) \, - \, c_{-k}^*(t) \, \right)$$
$$+ \, \frac{\kappa \, a}{2} \, \sum_k \, k^2 \left(\, c_{-k}(t) \, + \, c_k^*(t) \, \right) \left(\, c_k(t) \, + \, c_{-k}^*(t) \, \right) \qquad (205)$$

37

but the frequency is given by the dispersion relation

$$\omega_k^2 \;=\; v^2 \, k^2 \;=\; \left(\frac{\kappa \, a}{\rho} \right) k^2 \tag{206}$$

Therefore, the expression for the Hamiltonian simplifies to

$$
\begin{aligned}
H \;&=\; \rho \sum_k \omega_k^2 \left(\, c_k^*(t) \, c_k(t) \;+\; c_{-k}(t) \, c_{-k}^*(t) \, \right) \\
&=\; \rho \sum_k \omega_k^2 \left(\, c_k^*(0) \, c_k(0) \;+\; c_{-k}(0) \, c_{-k}^*(0) \, \right)
\end{aligned}
\tag{207}
$$

which is time-independent, since the time-dependent phase factors cancel out. Thus, one can think of the energy as a function of the variables c_k and c_{-k}^*. Since the Hamiltonian is strictly expressed in terms of canonically conjugate coordinates and momenta, one should examine the Poisson brackets of c_k and c_{-k}^*.

The variables $y(x_i)$ and $p(x_j)$ have the Poisson brackets

$$
\left\{ \, p(x_i) \, , \; y(x_j) \, \right\} \;=\; -\, \delta_{i,j}
$$

$$
\left\{ \, p(x_i) \, , \; p(x_j) \, \right\} \;=\; \left\{ \, y(x_i) \, , \; y(x_j) \, \right\} \;=\; 0
\tag{208}
$$

Due to the orthogonality properties of the plane-waves, one has

$$
\left(\, c_k \;+\; c_{-k}^* \, \right) \;\approx\; \frac{a}{\sqrt{L}} \sum_i \, y(x_i) \, \exp \left[\, -\, i \, k \, x_i \, \right]
\tag{209}
$$

and also

$$
-\, i \, \omega_{k'} \, \rho \, a \left(\, c_{k'} \;-\; c_{-k'}^* \, \right) \;\approx\; \frac{a}{\sqrt{L}} \sum_j \, p(x_j) \, \exp \left[\, -\, i \, k' \, x_j \, \right]
\tag{210}
$$

These relations are simply the results of applying the inverse Fourier transform to $y(x)$ and $p(x)$. One can find the Poisson brackets relations between c_k and c_k^* from

$$
\begin{aligned}
&-\, i \, \omega_{k'} \, \rho \left\{ \left(\, c_{k'} \;-\; c_{-k'}^* \, \right) , \left(\, c_k \;+\; c_{-k}^* \, \right) \right\} \\
&=\; -\, \frac{a}{L} \sum_{i,j} \left\{ \, p(x_i) \, , \; y(x_j) \, \right\} \exp \left[\, -\, i \, (\, k \, x_i \;+\; k' \, x_j \,) \, \right] \\
&=\; +\, \frac{a}{L} \sum_{i,j} \delta_{i,j} \, \exp \left[\, -\, i \, (\, k \, x_i \;+\; k' \, x_j \,) \, \right] \\
&=\; +\, \frac{a}{L} \sum_i \exp \left[\, -\, i \, (\, k \;+\; k' \,) \, x_i \, \right] \\
&=\; +\, \delta_{k+k'}
\end{aligned}
\tag{211}
$$

Likewise, one can obtain similar expressions for the other commutation relations. This set of equations can be satisfied by setting

$$\left\{ \, c_{k'}^* \, , \, c_k \, \right\} \; = \; \frac{i}{2 \, \omega_k \, \rho} \, \delta_{k,k'} \tag{212}$$

and

$$\left\{ \, c_{k'}^* \, , \, c_k^* \, \right\} \; = \; \left\{ \, c_{k'} \, , \, c_k \, \right\} \; = \; 0 \tag{213}$$

The above set of Poisson brackets can be recast in a simpler form by defining

$$c_k \; = \; \frac{1}{\sqrt{2 \, \omega_k \, \rho}} \, a_k \tag{214}$$

etc., so that the Poisson brackets reduce to

$$\left\{ \, a_{k'}^* \, , \, a_k \, \right\} \; = \; i \, \delta_{k,k'} \tag{215}$$

and

$$\left\{ \, a_{k'}^* \, , \, a_k^* \, \right\} \; = \; \left\{ \, a_{k'} \, , \, a_k \, \right\} \; = \; 0 \tag{216}$$

where the non-universal factors have cancelled out.

5.3   Rules of Canonical Quantization

The first rule of Canonical Quantization states, "Physical quantities should be represented by operators". Hence a_k and a_k^* should be replaced by the operators $\hat{a}_k$ and $\hat{a}_k^\dagger$. The second rule of Canonical Quantization states, "Poisson Brackets should be replaced by Commutators". Hence,

$$i \, \hbar \left\{ \, A \, , \, B \, \right\} \; \rightarrow \; [\, \hat{A} \, , \, \hat{B} \,] \tag{217}$$

So one has

$$[\, \hat{a}_{k'}^\dagger \, , \, \hat{a}_k \,] \; = \; - \, \hbar \, \delta_{k,k'} \tag{218}$$

and

$$[\, \hat{a}_{k'}^\dagger \, , \, \hat{a}_k^\dagger \,] \; = \; [\, \hat{a}_{k'} \, , \, \hat{a}_k \,] \; = \; 0 \tag{219}$$

To get rid of the annoying $\hbar$ in the commutator, one can set

$$\begin{aligned} \hat{a}_k \; &= \; \sqrt{\hbar} \, \hat{b}_k \\ \hat{a}_k^\dagger \; &= \; \sqrt{\hbar} \, \hat{b}_k^\dagger \end{aligned} \tag{220}$$

Whether it was noted or not, $\hat{b}_k^\dagger$ is the Hermitean conjugate of $\hat{b}_k$. The Hermitean relation can proved by taking the Hermitean conjugate of $\hat{y}(x_i)$, and

noting that the third rule of quantization states, "Measurable quantities are to replaced by Hermitean operators". Therefore, the operator

$$\hat{y}(x) \; = \; \frac{1}{\sqrt{L}} \; \sum_k \; \sqrt{\frac{\hbar}{2 \, \rho \, \omega_k}} \left(\; \hat{b}_k(t) \; + \; \hat{b}^\dagger_{-k}(t) \; \right) \; \exp\left[\; i \; k \; x \; \right] \qquad (221)$$

must be Hermitean. What this means is, the Hermitean conjugate

$$\hat{y}^\dagger(x) \; = \; \frac{1}{\sqrt{L}} \; \sum_k \; \sqrt{\frac{\hbar}{2 \, \rho \, \omega_k}} \left(\; \hat{b}^\dagger_k(t) \; + \; (\, \hat{b}^\dagger_{-k}(t))^\dagger \; \right) \; \exp\left[\; - \; i \; k \; x \; \right] \quad (222)$$

has to be the same as $\hat{y}(x)$. On setting $k = -k'$ in the above equation, one has

$$\hat{y}^\dagger(x) \; = \; \frac{1}{\sqrt{L}} \; \sum_{k'} \; \sqrt{\frac{\hbar}{2 \, \rho \, \omega_{k'}}} \left(\; \hat{b}^\dagger_{-k'}(t) \; + \; (\, \hat{b}^\dagger_{k'}(t))^\dagger \; \right) \; \exp\left[\; + \; i \; k' \; x \; \right] \quad (223)$$

For $\hat{y}^\dagger(x)$ to be equal to $\hat{y}(x)$, it is necessary that the Hermitean conjugate of the operator $\hat{b}^\dagger_{k'}$ is equal to $\hat{b}_{k'}$. This shows that the pair of operators are indeed Hermitean conjugates. The quantum field is represented by the operator

$$\hat{y}(x) \; = \; \frac{1}{\sqrt{L}} \; \sum_{k'} \; \sqrt{\frac{\hbar}{2 \, \rho \, \omega_{k'}}} \left(\; \hat{b}^\dagger_{-k'}(t) \; + \; \hat{b}_{k'}(t) \; \right) \; \exp\left[\; + \; i \; k' \; x \; \right] \qquad (224)$$

where the time-dependent creation and annihilation operators are given by

$$\hat{b}_k(t) \; = \; \hat{b}_k \; \exp\left[\; - \; i \; \omega_k \; t \; \right]$$

$$\hat{b}^\dagger_k(t) \; = \; \hat{b}^\dagger_k \; \exp\left[\; + \; i \; \omega_k \; t \; \right] \qquad (225)$$

The quantized Hamiltonian becomes

$$\begin{aligned} \hat{H} \; &= \; \rho \; \sum_k \; \omega_k^2 \left(\; \hat{c}^\dagger_k \; \hat{c}_k \; + \; \hat{c}_{-k} \; c^\dagger_{-k} \; \right) \\ &= \; \sum_k \; \frac{\hbar \, \omega_k}{2} \left(\; \hat{b}^\dagger_k \; \hat{b}_k \; + \; \hat{b}_{-k} \; \hat{b}^\dagger_{-k} \; \right) \qquad (226) \end{aligned}$$

On transforming $k \rightarrow -k$ in the second term of the summation, one obtains the standard form

$$\hat{H} \; = \; \sum_k \; \frac{\hbar \, \omega_k}{2} \left(\; \hat{b}^\dagger_k \; \hat{b}_k \; + \; \hat{b}_k \; \hat{b}^\dagger_k \; \right) \qquad (227)$$

where the $\hat{b}_k$ and $\hat{b}_k^\dagger$ are to be identified as annihilation and creation operators for the quanta.

The quantum operator $\hat{P}$ corresponding to the classical quantity P

$$P = \int_0^L dx \, \mathcal{P} \tag{228}$$

is evaluated as

$$\hat{P} = -\left(\frac{\kappa \, a}{2 \, \rho} \right) \sum_k \hbar \, k \left(\hat{b}_{-k}^\dagger - \hat{b}_k \right) \left(\hat{b}_k^\dagger + \hat{b}_{-k} \right) \tag{229}$$

where the plane-wave orthogonality properties have been used. This quantity can be expressed as the sum of two terms

$$\hat{P} = -\left(\frac{\kappa \, a}{2 \, \rho} \right) \sum_k \hbar \, k \left(\hat{b}_{-k}^\dagger \, \hat{b}_k^\dagger - \hat{b}_k \, \hat{b}_{-k} \right)$$
$$+ \left(\frac{\kappa \, a}{2 \, \rho} \right) \sum_k \hbar \, k \left(\hat{b}_k \, \hat{b}_k^\dagger - \hat{b}_{-k}^\dagger \, \hat{b}_{-k} \right) \tag{230}$$

The operator $\hat{P}$ can be shown to be equivalent to

$$\hat{P} = v^2 \sum_k \hbar \, k \, \hat{b}_k^\dagger \, \hat{b}_k \tag{231}$$

which obviously is proportional to the sum of the momenta of the quanta. The quantity $\frac{\kappa \, a}{\rho}$ is just the square of the wave velocity v^2. On noting that the quanta travel with velocities given by $v \, \text{sign}(k)$ and have energies given by $\hbar \, \omega_k = \hbar \, v \mid k \mid$, one sees that $\mathcal{P}$ is the expressed as the total energy flux associated with the quanta.

5.4   The Algebra of Boson Operators

The number operator $\hat{n}_k$ can be defined[6] as

$$\hat{n}_k = \hat{b}_k^\dagger \, \hat{b}_k \tag{232}$$

which has eigenstates $\mid n_k >$ with eigenvalues n_k

$$\hat{n}_k \mid n_k > = n_k \mid n_k > \tag{233}$$

The eigenvalues n_k are positive integers, including zero. This can be inferred from the commutation relations

$$[\, \hat{b}_k^\dagger \, , \, \hat{b}_{k'} \,] = - \, \delta_{k,k'} \tag{234}$$

---

[6]P. Jordan and O. Klein, Zeit. für Physik, **45**, 751 (1927).

which has the consequence that

$$
\begin{aligned}
[\ \hat{n}_k\ ,\ \hat{b}^{\dagger}_{k'}\ ] &= +\ \delta_{k,k'}\ \hat{b}^{\dagger}_{k'} \\
[\ \hat{n}_k\ ,\ \hat{b}_{k'}\ ] &= -\ \delta_{k,k'}\ \hat{b}_{k'}
\end{aligned}
\tag{235}
$$

Hence, when $\hat{b}_k$ acts on an eigenstate of $\hat{n}_k$ with eigenvalue n_k it produces another eigenstate of $\hat{n}_k$ but with an eigenvalue of $n_k - 1$, as can be seen since

$$
\begin{aligned}
\hat{n}_k\ \hat{b}_k\ |\ n_k\ > &= \hat{b}_k\ (\ \hat{n}_k\ -\ 1\ )\ |\ n_k\ > \\
&= (\ n_k\ -\ 1\ )\ \hat{b}_k\ |\ n_k\ >
\end{aligned}
\tag{236}
$$

Therefore, since $\hat{b}_k$ lowers the eigenvalue of the number operator by one unit, one can write

$$
\hat{b}_k\ |\ n_k\ > \ = \ C(n_k)\ |\ n_k - 1\ >
\tag{237}
$$

where the complex number $C(n_k)$ has to be determined. The normalization coefficient $C(n_k)$ can be determined by noting that

$$
<\ n_k\ |\ \hat{b}^{\dagger}_k
\tag{238}
$$

is the Hermitean conjugate of the state

$$
\hat{b}_k\ |\ n_k\ >
\tag{239}
$$

On taking the norm of the state and its conjugate, one finds the normalization

$$
\begin{aligned}
<\ n_k\ |\ \hat{b}^{\dagger}_k\ \hat{b}_k\ |\ n_k\ > &= C^*(n_k)\ C(n_k)\ <\ n_k - 1\ |\ n_k - 1\ > \\
&= |\ C(n_k)\ |^2
\end{aligned}
\tag{240}
$$

However, on using the definition of the number operator and the normalization condition, one finds that

$$
|\ C(n_k)\ |^2\ =\ n_k
\tag{241}
$$

so, on choosing the phase factor, one can define

$$
\hat{b}_k\ |\ n_k\ > \ = \ \sqrt{n_k}\ |\ n_k - 1\ >
\tag{242}
$$

as the annihilation operator.

Likewise, one can see that Hermitean conjugate operator $\hat{b}^{\dagger}_k$ when acting on an eigenstate of the number operator increases its eigenvalue by one unit

$$
\begin{aligned}
\hat{n}_k\ \hat{b}^{\dagger}_k\ |\ n_k\ > &= \hat{b}^{\dagger}_k\ (\ \hat{n}_k\ +\ 1\ )\ |\ n_k\ > \\
&= (\ n_k\ +\ 1\ )\ \hat{b}^{\dagger}_k\ |\ n_k\ >
\end{aligned}
\tag{243}
$$

Therefore, one has

$$
\hat{b}^{\dagger}_k\ |\ n_k\ > \ = \ C^{'}(n_k)\ |\ n_k + 1\ >
\tag{244}
$$

The coefficient $C'(n_k)$ is found from the normalization condition

$$
\begin{aligned}
< n_k \mid \hat{b}_k \, \hat{b}_k^\dagger \mid n_k > &= C'^*(n_k) \, C'(n_k) < n_k + 1 \mid n_k + 1 > \\
&= \mid C'(n_k) \mid^2
\end{aligned}
\tag{245}
$$

which with

$$
\hat{b}_k \, \hat{b}_k^\dagger = \hat{n}_k + 1
\tag{246}
$$

yields

$$
\mid C'(n_k) \mid^2 = n_k + 1
\tag{247}
$$

Since, the phase factor has already been determined by the Hermitean conjugate equation, one has

$$
\hat{b}_k^\dagger \mid n_k > = \sqrt{n_k + 1} \mid n_k + 1 >
\tag{248}
$$

which raises the eigenvalue of the number operator.

Hence, one sees that the eigenvalues of the number operator are separated by integers. Furthermore, the smallest eigenvalue corresponds to $n_k = 0$, since for $n_k = 0$ the equation

$$
\hat{b}_k \mid n_k > = \sqrt{n_k} \mid n_k - 1 >
\tag{249}
$$

reduces to

$$
\hat{b}_k \mid 0 > = 0
\tag{250}
$$

Hence, the hierarchy of states produced by the annihilation operator acting on a number operator eigenstate terminates at $n_k = 0$. Thus, the eigenvalues of the number operator n_k can have integer values 0 , 1 , 2 , 3 ,... , ∞.

Therefore, any arbitrary number operator eigenstate $\mid \{n_k\} >$, in which the number of excitations (n_k) in the each normal mode has been specified, can be written in terms of the vacuum state $\mid 0 >$ and the creation operators as

$$
\mid \{n_k\} > = \prod_k \left(\frac{(\hat{b}_k^\dagger)^{n_k}}{\sqrt{n_k!}} \right) \mid 0 >
\tag{251}
$$

The repeated operation of the creation operator $\hat{b}_k^\dagger$ creates a state with n_k bosonic excitations present in mode k and the denominator provides the correct normalization for this state.

Any arbitrary state $\mid \Psi >$ can be expressed as a linear superposition of number operator eigenstates

$$
\mid \Psi > = \sum_{\{n_k\}} C(\{n_k\}) \mid \{n_k\} >
\tag{252}
$$

where the sum runs over all possible number eigenstates, and the complex coefficients $C(\{n_k\})$ are arbitrary except that they must satisfy the normalization condition

$$
\sum_{\{n_k\}} \mid C(\{n_k\}) \mid^2 = 1
\tag{253}
$$

5.5 The Classical Limit

The classical limit of the quantum field theory can be characterized by the limit in which the field operator can be replaced by a function. This requires that the "classical" states are not only described as states with large numbers of quanta in the excited normal modes, but also that the state is a linear superposition of states with different number of quanta, with a reasonable well defined phase of the complex coefficients. For a quantum state to ideally represent a given classical state, one needs the quantum state to be composed of a coherent superposition of states with different numbers of quanta.

That states which are eigenstates of the number operators ($| \{n_k\} >$) can not represent classical states, can be seen by noting that the expectation value of the field operator is zero

$$< \{n_k\} \mid \hat{y}(x) \mid \{n_k\} > = 0 \tag{254}$$

follows from the expectation value of the creation and annihilation operators

$$< \{n_k\} \mid a_k \mid \{n_k\} > = 0 \tag{255}$$

Despite the fact that the average value of the field is zero, the fluctuation in the field amplitude is infinite since

$$< \{n_k\} \mid \hat{y}(x)^2 \mid \{n_k\} > = \frac{1}{L} \sum_{k'} \frac{\hbar}{2 \rho \, \omega_{k'}} < \{n_k\} \mid \left(\hat{b}_{k'}^\dagger + \hat{b}_{-k'} \right) \left(\hat{b}_{k'} + \hat{b}_{-k'}^\dagger \right) \mid \{n_k\} >$$

$$= \frac{1}{L} \sum_{k'} \frac{\hbar}{2 \rho \, \omega_{k'}} \left(1 + 2 \, n_{k'} \right) \tag{256}$$

and the zero-point contribution diverges logarithmically at the upper and lower limits of integration.

Hence, the eigenstates of the number operator. or equivalently $\hat{H}$, do not describe the classical states of the string. Classical states must be expressed as a linear superposition of energy eigenstates.

6 Classical Field Theory

The dynamics of a multi-component classical field ϕ^α is governed by a Lagrange density $\mathcal{L}$, which is a scalar quantity that is a function of the fields ϕ^α and their derivatives $\partial_\mu \phi^\alpha$. The equations of motion for the classical field are determined by the principle of extremal action. That is, the classical fields are those for which the action S

$$S = \int dt' \int d^3\underline{x} \, \mathcal{L} \left(\phi^\alpha, \partial_\mu \phi^\alpha \right) \tag{257}$$

is extremal. An arbitrary field ϕ^α can be expressed in terms of the extremal value ϕ^α_{ex} and the deviation $\delta\phi^\alpha$ as

$$\phi^\alpha = \phi^\alpha_{ex} + \delta\phi^\alpha \tag{258}$$

The space and time derivatives of the arbitrary field can also be expressed as the derivatives of the sum of the extremal field and the deviation

$$\partial_\nu \phi^\alpha = \partial_\nu \phi^\alpha_{ex} + \partial_\nu \delta\phi^\alpha \tag{259}$$

The first-order change in the action δS is given by

$$\delta S = \int_0^t dt' \int d^3\underline{x} \left[\delta\phi^\alpha \, \frac{\partial}{\partial \phi^\alpha} \, \mathcal{L}\left(\phi^\alpha_{ex} , \partial_\mu \phi^\alpha_{ex} \right) + (\partial_\nu \delta\phi^\alpha) \, \frac{\partial}{\partial(\partial_\nu \phi^\alpha)} \, \mathcal{L}\left(\phi^\alpha_{ex} , \partial_\mu \phi^\alpha_{ex} \right) \right] \tag{260}$$

On integrating by parts with respect to x^ν in the last term, and on assuming appropriate boundary conditions, one finds

$$\delta S = \int_0^t dt' \int d^3\underline{x} \, \delta\phi^\alpha \left[\frac{\partial}{\partial \phi^\alpha} \, \mathcal{L}\left(\phi^\alpha_{ex} , \partial_\mu \phi^\alpha_{ex} \right) - \partial_\nu \left[\frac{\partial}{\partial(\partial_\nu \phi^\alpha)} \, \mathcal{L}\left(\phi^\alpha_{ex} , \partial_\mu \phi^\alpha_{ex} \right) \right] \right] \tag{261}$$

which has to vanish for an arbitrary choice of $\delta\phi_\alpha$. Hence, one obtains the Euler-Lagrange equations

$$\frac{\partial}{\partial \phi^\alpha} \, \mathcal{L}\left(\phi^\alpha_{ex} , \partial_\mu \phi^\alpha_{ex} \right) = \partial_\nu \left[\frac{\partial}{\partial(\partial_\nu \phi^\alpha)} \, \mathcal{L}\left(\phi^\alpha_{ex} , \partial_\mu \phi^\alpha_{ex} \right) \right] \tag{262}$$

This set of equations determine the time dependence of the classical fields $\phi^\alpha_{ex}(x)$. That is, out of all possible fields with components ϕ_α, the equations of motion determine the physical field which has the components ϕ^α_{ex}. It is convenient to define the field momentum density $\pi^0_\alpha(x)$ conjugate to ϕ^α as

$$\pi^0_\alpha(x^\nu) = \frac{1}{c} \, \frac{\partial}{\partial(\partial_0 \phi^\alpha)} \, \mathcal{L}\left(\phi^\beta , \partial_\mu \phi^\beta \right) \tag{263}$$

The Hamiltonian density $\mathcal{H}$ is then defined as the Legendre transform

$$\mathcal{H} = c \sum_\alpha \pi^0_\alpha \, (\partial_0 \phi^\alpha) - \mathcal{L} \tag{264}$$

which eliminates the time-derivative of the fields in terms of the momentum density of the fields.

Exercise:

Starting from the Lorentz scalar Lagrangian

$$\mathcal{L} = \frac{1}{2} \left[(\partial_\mu \phi) (\partial^\mu \phi) - \left(\frac{m \, c}{\hbar} \right)^2 \phi^2 \right] \tag{265}$$

45

for a real scalar field ϕ, determine the Euler-Lagrange equation and the Hamiltonian density $\mathcal{H}$.

Exercise:

Consider the Lagrangian density

$$\mathcal{L} = \frac{1}{2} \left[(\partial_\mu \psi^*)(\partial^\mu \psi) - \left(\frac{m\,c}{\hbar}\right)^2 |\psi|^2 \right] \tag{266}$$

for a complex scalar field ψ. Treat ψ and ψ^* as independent fields.
(i) Determine the Euler-Lagrange equation and the Hamiltonian density $\mathcal{H}$.
(ii) By Fourier transforming with respect to space and time, determine the form of the general solution for ψ.

Exercise:

The Lagrangian density for the complex field ψ representing a charged particle is given by

$$\mathcal{L} = -\frac{\hbar^2}{2\,m} \left(\underline{\nabla}\,\psi^*\right) \cdot \left(\underline{\nabla}\,\psi\right) - \frac{\hbar}{2\,i} \left[\psi^* \left(\frac{\partial\psi}{\partial t}\right) - \left(\frac{\partial\psi^*}{\partial t}\right)\psi \right] - \psi^*\,V(x)\,\psi \tag{267}$$

(i) Determine the equation of motion, and the Hamiltonian density $\mathcal{H}$.
(ii) Consider the case $V(x) \equiv 0$, then by Fourier transforming with respect to space and time, determine the form of the general solution for ψ.

6.1 The Hamiltonian Formulation

The Hamiltonian formulation reserves a special role for time, and so is not Lorentz covariant. However, the Hamiltonian formulation is the most convenient formulation for quantizing fields. The Hamilton equations of motion are determined from the Hamiltonian

$$H = \int d^3\underline{x}\, \mathcal{H} \tag{268}$$

by noting that $\mathcal{H}$ is only a functional of π^0_α and ϕ^α. This can be seen, since as

$$H = \int d^3\underline{x} \left(c \sum_\alpha \pi^0_\alpha \, (\partial_0 \phi^\alpha) - \mathcal{L} \right) \tag{269}$$

then, the first-order variation of the Hamiltonian δH is given by

$$\delta H = \int d^3\underline{x} \left[c \sum_\alpha \left(\delta\pi^0_\alpha \, (\partial_0\phi^\alpha) + \pi^0_\alpha \, (\partial_0\delta\phi^\alpha) \right) - \delta\mathcal{L} \right] \tag{270}$$

but, from the Lagrangian formulation of field theory, one has

$$\frac{1}{c}\,\delta\mathcal{L} \;=\; \delta\phi^{\alpha}\,(\partial_0\pi^0_{\alpha}\,) \;+\; (\partial_0\delta\phi^{\alpha})\,\pi^0_{\alpha} \tag{271}$$

where the Euler-Lagrange equations were substituted into the first term. Therefore, the variation in the Hamiltonian is given by

$$\delta H \;=\; \int d^3\underline{x}\,c\,\sum_{\alpha}\,\left(\,\delta\pi^0_{\alpha}\,(\partial_0\phi^{\alpha}) \;-\; \delta\phi^{\alpha}\,(\partial_0\pi^0_{\alpha})\,\right) \tag{272}$$

which does not involve the time derivative of the fields. This implies that the Hamiltonian is a function of the fields π^0_{α}, ϕ^{α} and their derivatives. On calculating the variation of $\mathcal{H}$ using the independent variables π^0_{α} and ϕ^{α}, and integrating by parts, one finds that the Hamiltonian equations of motion are given by

$$c\,\partial_0\phi^{\alpha} \;=\; \left(\frac{\partial\mathcal{H}}{\partial\pi^0_{\alpha}}\right) \;-\; \underline{\nabla}\left(\frac{\partial\mathcal{H}}{\partial(\underline{\nabla}\pi^0_{\alpha})}\right)$$

$$-\,c\,\partial_0\pi^0_{\alpha} \;=\; \left(\frac{\partial\mathcal{H}}{\partial\phi^{\alpha}}\right) \;-\; \underline{\nabla}\left(\frac{\partial\mathcal{H}}{\partial(\underline{\nabla}\phi^{\alpha})}\right) \tag{273}$$

The structure of these equations are similar to those of the classical mechanics of point particles. Similar to classical mechanics of point particles, one can define Poisson Brackets with fields. When quantizing the fields, the Poisson Bracket relations between the fields can be replaced by commutation relations.

6.2 Symmetry and Conservation Laws

Emmy Noether produced a theorem linking continuous symmetries of a Lagrangian to conservation laws[7].

6.2.1 Conservation Laws

Consider a Lagrangian density $\mathcal{L}$ which is a function of a set of fields $\phi^{\alpha}(x)$ and their derivatives defined in a Minkowski space x. Consider how the Lagrangian density changes for a particular choice of a combination of infinitesimal transformations of the field components

$$\phi^{\alpha}(x) \;\rightarrow\; \phi^{\alpha\prime}(x) \;=\; \phi^{\alpha}(x) \;+\; \delta\phi^{\alpha}(x) \tag{274}$$

and, as a consequence, the derivatives of the field components also transform as

$$\partial_{\mu}\phi^{\alpha}(x) \;\rightarrow\; \partial_{\mu}\phi^{\alpha\prime}(x) \;=\; \partial_{\mu}\phi^{\alpha}(x) \;+\; \partial_{\mu}\delta\phi^{\alpha}(x) \tag{275}$$

---

[7]E. Noether, Nachr. d. Kgl. Gessch. d. Wiss. Gottingen, K1. Math. Phys. (1918) 235.

Under this combined transformation, the Lagrangian density changes by an infinitesimal amount $\delta\mathcal{L}$, given by

$$\delta\mathcal{L} = \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}\right)\partial_\mu\delta\phi^\alpha + \frac{\partial\mathcal{L}}{\partial\phi^\alpha}\delta\phi^\alpha \qquad (276)$$

where the field index α is to be summed over. However, the generalized momentum density $\pi_\alpha^\mu(x)$ is defined by

$$\pi_\alpha^\mu(x) = \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}\right) \qquad (277)$$

so

$$\delta\mathcal{L} = \pi_\alpha^\mu\,(\partial_\mu\delta\phi^\alpha) + \frac{\partial\mathcal{L}}{\partial\phi^\alpha}\delta\phi^\alpha \qquad (278)$$

The Euler-Lagrange equation for each field ϕ^α is given by

$$\partial_\mu\pi_\alpha^\mu - \frac{\partial\mathcal{L}}{\partial\phi^\alpha} = 0 \qquad (279)$$

where ϕ^α satisfies the appropriate boundary conditions. Thus, on adding and subtracting a term

$$(\partial_\mu\,\pi_\alpha^\mu)\,\delta\phi^\alpha \qquad (280)$$

to $\delta\mathcal{L}$, one finds

$$
\begin{aligned}
\delta\mathcal{L} &= \left(\pi_\alpha^\mu\,\partial_\mu\delta\phi^\alpha + (\partial_\mu\,\pi_\alpha^\mu)\,\delta\phi^\alpha\right) + \left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha} - \partial_\mu\pi_\alpha^\mu\right)\delta\phi^\alpha \\
&= \partial_\mu\left(\pi_\alpha^\mu\,\delta\phi^\alpha\right) + \left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha} - \partial_\mu\pi_\alpha^\mu\right)\delta\phi^\alpha \\
&= \partial_\mu\left(\pi_\alpha^\mu\,\delta\phi^\alpha\right) \qquad (281)
\end{aligned}
$$

since the last term in the second line vanishes if the fields ϕ^α satisfy the Euler-Lagrange equations. If the Lagrangian is invariant under the transformation, then $\delta\mathcal{L} = 0$, so

$$\partial_\mu\left(\pi_\alpha^\mu\,\delta\phi^\alpha\right) = 0 \qquad (282)$$

where the field index α is to be summed over. The above equation can be re-written as a continuity equation

$$\partial_\mu j^\mu = 0 \qquad (283)$$

where the conserved current $j^\mu(x)$ is given by

$$
\begin{aligned}
j^\mu(x) &\propto \pi_\alpha^\mu(x)\,\delta\phi^\alpha(x) \\
&\propto \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}\right)\delta\phi^\alpha(x) \qquad (284)
\end{aligned}
$$

48

up to a constant of proportionality. The normalization of the conserved current is arbitrary and can be chosen at will. Since it is recognized that $\delta\phi^\alpha$ is infinitesimal, the normalization is chosen by introducing an infinitesimal constant ϵ via

$$
\begin{aligned}
\epsilon \, j^\mu(x) &= \pi^\mu_\alpha(x) \, \delta\phi^\alpha(x) \\
&= \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\alpha)} \right) \delta\phi^\alpha(x)
\end{aligned}
\tag{285}
$$

The conserved charge Q is defined as the integral over all space of the time component of the current density $j^{(0)}$. That is, the conserved charge is given by

$$
Q = \int d^3\underline{x} \, j^{(0)}(x)
\tag{286}
$$

or, more specifically

$$
\begin{aligned}
\epsilon \, Q &= \int d^3\underline{x} \, \pi^{(0)}_\alpha(x) \, \delta\phi^\alpha(x) \\
&= \int d^3\underline{x} \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \phi^\alpha)} \right) \delta\phi^\alpha(x)
\end{aligned}
\tag{287}
$$

Since ϵ is a constant, the total charge Q is constant. Therefore, the total time derivative of Q vanishes

$$
\frac{dQ}{dt} = 0
\tag{288}
$$

The spatial components of j^μ form the current density vector.

6.2.2 Noether Charges

Consider the infinitesimal variation of a complex field $\phi_\alpha(x)$ defined by

$$
\phi^\alpha(x) \rightarrow \phi'^\alpha(x) = \phi^\alpha(x) + i \, \epsilon \sum_\beta \lambda^\alpha{}_\beta \, \phi^\beta(x)
\tag{289}
$$

If this infinitesimal variation leads to $\mathcal{L}$ being invariant, one has a conserved current

$$
j^\mu = i \sum_{\alpha,\beta} \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\alpha)} \right) \lambda^\alpha{}_\beta \, \phi^\beta(x)
\tag{290}
$$

An important example is given by the infinitesimal transformation

$$
\begin{aligned}
\psi' &= \psi + i \, \epsilon \, \psi \\
\psi^{*\prime} &= \psi^* - i \, \epsilon \, \psi^*
\end{aligned}
\tag{291}
$$

49

where ψ and its complex conjugate ψ^* are regarded as independent fields. The transformation represents a an infinitesimal constant shift of the phase of the field[8]. The conserved current is

$$j^\mu \; = \; - \, i \, \left[\, \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} \right) \psi(x) \; - \; \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi^*)} \right) \psi^*(x) \, \right] \tag{292}$$

which is the electromagnetic current density four-vector.

Exercise:

The Lagrangian density for the complex Schrödinger field representing a charged particle is given by

$$\mathcal{L} \; = \; - \frac{\hbar^2}{2 \, m} \left(\underline{\nabla} \, \psi^* \right) \cdot \left(\underline{\nabla} \, \psi \right) - \frac{\hbar}{2 \, i} \left[\, \psi^* \left(\frac{\partial \psi}{\partial t} \right) - \left(\frac{\partial \psi^*}{\partial t} \right) \psi \, \right] - \psi^* \, V(x) \, \psi \tag{293}$$

(i) Determine the conserved Noether charges.

Exercise:

Determine the Noether charges for a complex Klein-Gordon field theory, governed by the Lagrangian density

$$\mathcal{L} \; = \; \frac{1}{2} \left[\, (\, \partial_\mu \psi^* \,) \, (\, \partial^\mu \psi \,) \; - \; \left(\frac{m \, c}{\hbar} \right)^2 \, | \, \psi \, |^2 \, \right] \tag{294}$$

6.2.3 Noether's Theorem

The basic theorem can be generalized to the case where the Lagrangian density is not invariant under the infinitesimal transformation, but instead changes by a combination of total derivatives. That is,

$$\delta \mathcal{L} \; = \; \epsilon \, \partial_\mu \Lambda^\mu \tag{295}$$

---

[8]This particular transformation is a specific example of a gauge transformations of the first kind, in which

$$\psi'(x) \; = \; \exp \left[\, - \, i \, \frac{q}{\hbar \, c} \, \Lambda(x) \, \right] \psi(x)$$

A gauge transformation of the second kind is one in which the field changes according to

$$A^{\mu \prime} \; = \; A^\mu \; + \; (\partial^\mu \, \Lambda)$$

Since $\hat{p}^\mu \; = \; i \, \hbar \, \partial^\mu$, the combination of these transformations keep the quantity $(\hat{p}^\mu - \frac{q}{c} \, A^\mu)\psi$ invariant

for some analytic vector function with components Λ^μ. This type of transformation does not change the total action. If the Lagrangian changes by the above amount for the combined transformation $\delta\phi^\alpha$

$$\phi^\alpha(x) \;\rightarrow\; \phi^{\alpha\prime}(x) \;=\; \phi^\alpha(x) \;+\; \delta\phi^\alpha(x) \tag{296}$$

then as has been previously shown

$$
\begin{aligned}
\delta\mathcal{L} \;&=\; \left(\pi_\alpha^\mu \,(\partial_\mu \delta\phi^\alpha) \;+\; (\partial_\mu \pi_\alpha^\mu)\, \delta\phi^\alpha \right) \;+\; \left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha} \;-\; \partial_\mu \pi_\alpha^\mu \right) \delta\phi^\alpha \\
&=\; \partial_\mu \left(\pi_\alpha^\mu \,\delta\phi^\alpha \right) \;+\; \left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha} \;-\; \partial_\mu \pi_\alpha^\mu \right) \delta\phi^\alpha \\
&=\; \partial_\mu \left(\pi_\alpha^\mu \,\delta\phi^\alpha \right)
\end{aligned}
\tag{297}
$$

one has

$$\epsilon\, \partial_\mu \Lambda^\mu \;=\; \partial_\mu \left(\pi_\alpha^\mu \,\delta\phi^\alpha \right) \tag{298}$$

If the conserved currents are identified as

$$\epsilon\, j^\mu \;=\; \left(\pi_\alpha^\mu \,\delta\phi^\alpha \right) \;-\; \epsilon\, \Lambda^\mu \tag{299}$$

then the continuity condition

$$\partial_\mu j^\mu \;=\; 0 \tag{300}$$

holds.

6.3   The Energy-Momentum Tensor

An example of Noether's theorem is given by the transformation

$$\phi^\alpha(x) \;\rightarrow\; \phi^\alpha(x+\epsilon) \;=\; \phi^\alpha(x) \;+\; \epsilon^\mu \,(\partial_\mu \phi^\alpha) \tag{301}$$

which represents an infinitesimal space-time translation. This is a symmetry appropriate to a Lagrangian density $\mathcal{L}$ which has no explicit x dependence. We shall assume that the Lagrangian density only depends on the field ϕ^α and its derivatives $\partial_\nu \phi^\alpha$

$$\mathcal{L} \;=\; \mathcal{L}\!\left(\phi^\alpha, (\partial_\nu \phi^\alpha) \right) \tag{302}$$

In this case, the change in the Lagrangian density is given by the total derivative

$$
\begin{aligned}
\delta\mathcal{L} \;&=\; \left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu \phi^\alpha)} \right) (\partial_\nu \delta\phi^\alpha) \;+\; \left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha} \right) \delta\phi^\alpha \\
&=\; \epsilon^\mu \left[\left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu \phi^\alpha)} \right) \left(\partial_\nu \partial_\mu \phi^\alpha \right) \;+\; \left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha} \right) \partial_\mu \phi^\alpha \right] \\
&=\; \epsilon^\mu \,\partial_\mu \mathcal{L}
\end{aligned}
\tag{303}
$$

where the last line follows since the Lagrangian only depends implicitly on x^μ through the fields. Hence, the change in the Lagrangian is a total derivative

$$\delta\mathcal{L} \;=\; \epsilon^\mu \, \partial_\mu \Lambda \tag{304}$$

where $\Lambda \;=\; \mathcal{L}$. Therefore, for transformations of the type

$$\phi^\alpha \;\to\; \phi^\alpha \;+\; \epsilon^\mu \, (\partial_\mu \phi^\alpha) \tag{305}$$

Noether's theorem takes the form

$$
\begin{aligned}
\epsilon^\mu \left(\partial_\mu \mathcal{L} \right) \;&=\; \epsilon^\mu \left[\left(\frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi^\alpha)} \right) \partial_\nu \left(\partial_\mu \phi^\alpha \right) \;+\; \left(\frac{\partial \mathcal{L}}{\partial \phi^\alpha} \right) \partial_\mu \phi^\alpha \right] \\
&=\; \epsilon^\mu \left[\left(\frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi^\alpha)} \right) \partial_\nu \left(\partial_\mu \phi^\alpha \right) \;+\; \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi^\alpha)} \right) \partial_\mu \phi^\alpha \right] \\
&=\; \epsilon^\mu \, \partial_\nu \left[\left(\frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi^\alpha)} \right) \left(\partial_\mu \phi^\alpha \right) \right]
\end{aligned}
\tag{306}
$$

where the Euler-Lagrange equation has been used in the second line. Thus, the fields satisfy the continuity conditions

$$0 \;=\; \epsilon^\mu \, \partial_\nu \left[\left(\frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi^\alpha)} \right) \left(\partial_\mu \phi^\alpha \right) \;-\; \delta^\nu{}_\mu \, \mathcal{L} \right] \tag{307}$$

where

$$
\begin{aligned}
\delta^\mu{}_\nu \;&=\; 1 \quad \text{if } \mu \;=\; \nu \\
\delta^\mu{}_\nu \;&=\; 0 \quad \text{otherwise}
\end{aligned}
\tag{308}
$$

The conserved current density is identified as

$$T^\nu{}_\mu \;=\; \left[\left(\frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi^\alpha)} \right) \left(\partial_\mu \phi^\alpha \right) \;-\; \delta^\nu{}_\mu \, \mathcal{L} \right] \tag{309}$$

which is the energy-momentum density T^ν_μ. The energy momentum tensor satisfies the conservation law

$$\partial_\nu \, T^\nu{}_\mu \;=\; 0 \tag{310}$$

The second-rank tensor can be written in contravariant form as

$$T^{\nu,\mu} \;=\; \left[\left(\frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi^\alpha)} \right) \left(\partial^\mu \phi^\alpha \right) \;-\; g^{\nu,\mu} \, \mathcal{L} \right] \tag{311}$$

where the metric tensor has been used to raise the index μ. The component with $\mu \;=\; \nu \;=\; 0$ is the Hamiltonian density $\mathcal{H}$ for the fields

$$\mathcal{H} \;=\; T^{0,0} \;=\; \left[\left(\frac{\partial \mathcal{L}}{\partial (\partial_0 \phi^\alpha)} \right) \left(\partial_0 \phi^\alpha \right) \;-\; \mathcal{L} \right] \tag{312}$$

so the total energy of the field is given by

$$E = \int d^3\underline{x}\, \mathcal{H} = \int d^3\underline{x}\, T^{0,0} \tag{313}$$

The energy is conserved since

$$\frac{1}{c}\frac{\partial \mathcal{H}}{\partial t} + \sum_j \frac{\partial}{\partial x^{(j)}} T^{j,0} = 0 \tag{314}$$

where the components $c\, T^{j,0}$ represents the components of the energy-density flux. Likewise, the components

$$T^{0,j} = c\, \pi_\alpha^{(0)} \left(\partial^{(j)} \phi^\alpha \right) \tag{315}$$

are related to the momentum density since the total momentum of the field is given by

$$P^{(j)} = \int d^3\underline{x}\, \frac{1}{c}\, T^{0,j} \tag{316}$$

Since $T^{0,j}$ is the momentum density, one expects that the components of the orbital angular momentum density are proportional to

$$M^{0,j,k} = T^{0,j}\, x^{(k)} - T^{0,k}\, x^{(j)} \tag{317}$$

One can define a third-rank tensor via

$$M^{\mu,\nu,\rho} = T^{\mu,\nu}\, x^\rho - T^{\mu,\rho}\, x^\nu \tag{318}$$

The divergence of the third-rank tensor is evaluated as

$$\begin{aligned}
\partial_\mu M^{\mu,\nu,\rho} &= \left(\partial_\mu T^{\mu,\nu} \right) x^\rho + T^{\mu,\nu}\, \delta^\rho{}_\mu - \left(\partial_\mu T^{\mu,\rho} \right) x^\nu - T^{\mu,\rho}\, \delta^\nu{}_\mu \\
&= T^{\rho,\nu} - T^{\nu,\rho} \tag{319}
\end{aligned}$$

where the conservation law for $T^{\mu,\nu}$ and the condition

$$\partial_\mu x^\rho = \delta_\mu{}^\rho \tag{320}$$

expressing the independence of the variables x^ρ and x^μ have been used. The divergence of the third-rank tensor vanishes if $T^{\mu,\nu}$ is symmetric. Thus, the angular momentum tensor $M^{\mu,\nu,\rho}$ is conserved if the energy-momentum tensor is symmetric.

It should be noted that the tensor $T^{\mu,\nu}$ is only symmetric for scalar fields. This is related to the fact that a vector or tensor field carries a non-zero intrinsic angular momentum. It is possible to incorporate an additional term in the

53

momentum-energy tensor of a vector field to make it symmetric.

Exercise:

(i) Determine the momentum-energy tensor for a complex scalar field ψ governed by the Lagrangian density

$$\mathcal{L} \; = \; \frac{1}{2} \left[\, (\, \partial_\mu \psi^* \,) \, (\, \partial^\mu \psi \,) \; - \; \left(\frac{m \, c}{\hbar} \right)^2 \, | \, \psi \, |^2 \, \right] \qquad (321)$$

(ii) Find the forms of the energy and momentum density of the field.
(iii) Using the form of the general solution, find expressions for the total energy and momentum of the field in terms of the Fourier components of the field.

Exercise:

(i) Determine the energy-momentum tensor for the Lagrangian density for the complex Schrödinger field representing a charged particle given by

$$\mathcal{L} \; = \; - \, \frac{\hbar^2}{2 \, m} \left(\, \underline{\nabla} \, \psi^* \, \right) \cdot \left(\, \underline{\nabla} \, \psi \, \right) - \frac{\hbar}{2 \, i} \left[\, \psi^* \left(\frac{\partial \psi}{\partial t} \right) - \left(\frac{\partial \psi^*}{\partial t} \right) \psi \, \right] - \psi^* \, V(x) \, \psi$$

$$(322)$$

(ii) Find the forms of the energy and momentum density of the field.
(iii) Find the forms of the generalized orbital angular momentum density of the field.
(iv) Consider the case where $V(x) \equiv 0$. Using the form of the general solution, find expressions for the total energy and momentum of the field in terms of the Fourier components of the field.

7   The Electromagnetic Lagrangian

The Lagrangian for a source-free electromagnetic field must be gauge invariant and must be a Lorentz scalar. An appropriate scalar Lagrange density can be constructed as

$$\mathcal{L} \; = \; - \, \frac{1}{16 \, \pi} \, F^{\mu,\nu} \, F_{\mu,\nu} \qquad (323)$$

where A^μ are the fields. The constant of proportionality is merely a matter of convention. The Euler-Lagrange equations are found by expressing the Lagrangian density in the symmetrical form

$$\mathcal{L} \; = \; - \, \frac{1}{16 \, \pi} \, F^{\mu,\nu} \, g_{\mu,\sigma} \, g_{\nu,\tau} \, \, F^{\sigma,\tau} \qquad (324)$$

From the above expression, it is seen that the two factors of the antisymmetrical second-rank field tensors produce identical variations of the action. Furthermore, in this form the Lagrangian only depends on the contravariant derivatives

of the contravariant components of the field. This form allows variations to be made directly without using the properties of the metric tensor. The first-order variation of the action can be expressed as

$$
\begin{aligned}
\delta S &= -\frac{2}{16\,\pi\,c} \int d^4x \left[\left(\partial_\mu \delta A_\nu \right) F^{\mu,\nu} - \left(\partial_\nu \delta A_\mu \right) F^{\mu,\nu} \right] \\
&= \frac{2}{16\,\pi\,c} \int d^4x \left[\delta A_\nu \, \partial_\mu \left(F^{\mu,\nu} - F^{\nu,\mu} \right) \right] \\
&= \frac{1}{4\,\pi\,c} \int d^4x \, \delta A_\nu \left(\partial_\mu F^{\mu,\nu} \right)
\end{aligned}
\tag{325}
$$

where the second line has been obtained by integrating by parts and the last line was obtained by using the antisymmetric nature of the field tensor. The vanishing of the first-order variation of the action δS, for arbitrary δA_ν, yields the Euler-Lagrange equation

$$
\partial_\mu F^{\mu,\nu} = 0
\tag{326}
$$

which is the same as Maxwell's equations in the absence of any sources.

In the absence of the source, the Lagrangian density is gauge invariant. This can be seen by noting that the contravariant field tensor $F^{\mu,\nu}$ is gauge invariant, and the covariant tensor is obtained from the contravariant tensor by lowering both indices with the metric tensor. The contravariant field tensor can be expressed as the matrix

$$
F^{\mu,\nu} \equiv
\begin{pmatrix}
0 & -E^{(1)} & -E^{(2)} & -E^{(3)} \\
E^{(1)} & 0 & -B^{(3)} & B^{(2)} \\
E^{(2)} & B^{(3)} & 0 & -B^{(1)} \\
E^{(3)} & -B^{(2)} & B^{(1)} & 0
\end{pmatrix}
\tag{327}
$$

and the co-variant field tensor can be expressed as the matrix

$$
F_{\mu,\nu} \equiv
\begin{pmatrix}
0 & E^{(1)} & E^{(2)} & E^{(3)} \\
-E^{(1)} & 0 & -B^{(3)} & B^{(2)} \\
-E^{(2)} & B^{(3)} & 0 & -B^{(1)} \\
-E^{(3)} & -B^{(2)} & B^{(1)} & 0
\end{pmatrix}
\tag{328}
$$

in which the sign of the terms with mixed time and space indices have changed. Therefore, the Lagrangian density can be expressed in terms of the electromagnetic fields as

$$
\mathcal{L} = \frac{1}{8\,\pi} \left(\underline{E}^2 - \underline{B}^2 \right)
\tag{329}
$$

Since the Lagrangian density is completely expressed in terms of the electromagnetic field, it is gauge invariant.

In the presence of source densities, the Lagrangian density is extended to include the interaction to become

$$\mathcal{L} = -\frac{1}{16\,\pi}\,F^{\mu,\nu}\,F_{\mu,\nu} - \frac{1}{c}\,A_\mu\,j^\mu \tag{330}$$

This interaction term is the only Lorentz scalar that one can form with the four-vector current and the field. It should be noted that the last term is not gauge invariant. This action yields the equation of motion

$$\partial_\mu F^{\mu,\nu} = \frac{4\,\pi}{c}\,j^\nu \tag{331}$$

as expected.

The lack of gauge invariance in the interaction Lagrangian

$$\mathcal{L}_{int} = -\frac{1}{c}\,A^\mu\,j_\mu \tag{332}$$

does not affect the equations of motion. On performing the gauge transformation

$$A^\mu \rightarrow A^{\mu\prime} = A^\mu + \partial^\mu\Lambda \tag{333}$$

one finds that the interaction part of the Lagrangian density is transformed to

$$\mathcal{L}_{int}{}' = -\frac{1}{c}\left(A^\mu + \partial^\mu\Lambda \right)j_\mu \tag{334}$$

Since charge is conserved, the current density must satisfy the continuity equation

$$\partial^\mu j_\mu = 0 \tag{335}$$

The continuity condition can be used to express the interaction as the untransformed Lagrangian density and a perfect derivative

$$\mathcal{L}_{int}{}' = -\frac{1}{c}\,A^\mu\,j_\mu - \frac{1}{c}\,\partial^\mu(\,\Lambda\,j_\mu\,) \tag{336}$$

The perfect derivative term only adds a constant term to the action which does not affect the equations of motion[9]. Hence, although the Lagrangian density is not gauge invariant in the presence of sources, the Lagrangian equations of motion are gauge invariant.

The momentum density conjugate to A_μ is calculated as

$$\pi^{0,\mu} = -\frac{c}{4\,\pi}\,F^{0,\mu} \tag{337}$$

---

[9]The change in the form of the interaction Lagrangian density produced by a gauge transformation should be taken as a warning against considering quantities in a field theory as being localized.

which vanishes for $\mu = 0$, indicating that the scalar potential A_0 is not a dynamic variable. This suggests that it may be appropriate to completely fix the scalar potential by a choice of gauge, such as the Coulomb gauge which leads to the scalar potential ϕ being fixed by Poisson's equation. In the presence of sources, the Hamiltonian density is expressed as

$$
\begin{aligned}
\mathcal{H} &= -\frac{1}{4\pi} \left(\partial_0 A_\nu \right) F^{0,\nu} - \mathcal{L} \\
&= -\frac{1}{4\pi} \left(F_{0,\nu} + \partial_\nu A_0 \right) F^{0,\nu} - \mathcal{L} \\
&= -\frac{1}{4\pi} \left(F_{0,\nu} + \partial_\nu A_0 \right) F^{0,\nu} + \frac{1}{8\pi} \left(\underline{B}^2 - \underline{E}^2 \right) + \frac{1}{c} j^\mu A_\mu \\
&= +\frac{1}{8\pi} \left(\underline{E}^2 + \underline{B}^2 \right) - \frac{1}{4\pi} \left(\underline{\nabla} \cdot \underline{E} \right) A_0 + \frac{1}{c} j^\mu A_\mu \\
&\quad + \frac{1}{4\pi} \underline{\nabla} \cdot \left(A^{(0)} \underline{E} \right)
\end{aligned}
\tag{338}
$$

The fourth line has been derived by noting that the non-zero components of $F^{0,\mu}$ are only non-zero for space-like μ and are given by

$$
F^{0,i} = -E^{(i)} \tag{339}
$$

Thus, the first term in the third line is given by

$$
-\frac{1}{4\pi} F_{0,\nu} \, F^{0,\nu} = +\frac{1}{4\pi} \underline{E}^2 \tag{340}
$$

which can be combined with the term

$$
-\frac{1}{8\pi} \left(\underline{E}^2 - \underline{B}^2 \right) \tag{341}
$$

originating from the Lagrangian density. This combination results in the term

$$
\frac{1}{8\pi} \left(\underline{E}^2 + \underline{B}^2 \right) \tag{342}
$$

which is recognized as the usual expression for the energy density of a free electromagnetic field. On substituting eqn(339) into the second term in the third line, one finds

$$
+\frac{1}{4\pi} \left(\underline{\nabla} A_0 \right) \cdot \underline{E} \tag{343}
$$

which can be expressed as

$$
\frac{1}{4\pi} \left(\underline{\nabla} A_0 \right) \cdot \underline{E} = \frac{1}{4\pi} \underline{\nabla} \cdot \left(A_0 \underline{E} \right) - \frac{1}{4\pi} A_0 \left(\underline{\nabla} \cdot \underline{E} \right) \tag{344}
$$

This relation has been used in arriving at the fourth line of eqn(338). Since the divergence of the electric field satisfies Gauss's law

$$
\underline{\nabla} \cdot \underline{E} = 4\pi \rho \tag{345}
$$

the expression given in eqn(344) simplifies to

$$
\frac{1}{4\pi} \left(\underline{\nabla} A_0 \right) . \underline{E} = \frac{1}{4\pi} \underline{\nabla} . \left(A_0 \, \underline{E} \right) - A_0 \, \rho \tag{346}
$$

Therefore, the Hamiltonian density can be expressed as

$$
\mathcal{H} = + \frac{1}{8\pi} \left(\underline{E}^2 + \underline{B}^2 \right) - \rho \, A_0 + \frac{1}{c} j^\mu \, A_\mu
$$
$$
+ \frac{1}{4\pi} \underline{\nabla} . \left(A^{(0)} \, \underline{E} \right) \tag{347}
$$

On combining the term $\rho \, A_0$ with the last term

$$
\frac{1}{c} j^\mu \, A_\mu = \rho \, A_0 - \frac{1}{c} \underline{j} . \underline{A} \tag{348}
$$

which originates from the Lagrangian interaction $(-\mathcal{L}_{int})$, one finds that the terms proportional to $A_0 \, \rho$ in the Hamiltonian density cancel. On neglecting the total derivative term $[+ \frac{1}{4\pi} \underline{\nabla} . (\phi \, \underline{E})]$, one finds that the Hamiltonian density reduces to

$$
\mathcal{H} = \frac{1}{8\pi} \left(\underline{E}^2 + \underline{B}^2 \right) - \frac{1}{c} \underline{j} . \underline{A} \tag{349}
$$

The first term is the energy density of the free electromagnetic field and the second term represents the energy of the interaction between the electromagnetic field and "charged particles". It should be noted that the interaction Hamiltonian is expressed entirely in terms of an interaction between the current density and the vector potential, which demonstrates that the Hamiltonian is not invariant under a Lorentz transformation

$$
\mathcal{H}_{int} = - \frac{1}{c} \underline{j} . \underline{A} \tag{350}
$$

but is invariant under rotations in space. This situation is to be contrasted with the interaction term in the Lagrangian which was Lorentz invariant as it explicitly included an interaction between the scalar potential and the charge density.

7.1 Conservation Laws for Electromagnetic Fields

The Lagrangian density $\mathcal{L}$ of an electromagnetic field is given by the Lorentz scalar

$$
\mathcal{L} = - \frac{1}{16\pi} F^{\mu,\nu} \, F_{\mu,\nu} - \frac{1}{c} j^\mu \, A_\mu \tag{351}
$$

or

$$
\mathcal{L} = - \frac{1}{16\pi} \left[\left(\partial^\mu A^\nu \right) - \left(\partial^\nu A^\mu \right) \right] \left[\left(\partial_\mu A_\nu \right) - \left(\partial_\nu A_\mu \right) \right] - \frac{1}{c} j_\mu \, A^\mu \tag{352}
$$

The Noetherian energy-momentum tensor $T^{\nu,\mu}$ is found from

$$
\begin{aligned}
T^\nu{}_\mu &= \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A^\rho)} \right) \left(\partial_\mu A^\rho \right) - \delta^\nu{}_\mu \, \mathcal{L} \\
&= \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\rho)} \right) \left(\partial_\mu A_\rho \right) - \delta^\nu{}_\mu \, \mathcal{L}
\end{aligned}
\tag{353}
$$

The derivative of the Lagrangian density is evaluated as

$$
\begin{aligned}
\left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\rho)} \right) &= -\frac{1}{8\,\pi} \left(F^{\nu,\rho} - F^{\rho,\nu} \right) \\
&= -\frac{1}{4\,\pi} F^{\nu,\rho}
\end{aligned}
\tag{354}
$$

Therefore, the energy-momentum density is found as

$$
T^\nu{}_\mu = -\frac{1}{4\,\pi} F^{\nu,\rho} \left(\partial_\mu A_\rho \right) - \delta^\nu{}_\mu \, \mathcal{L}
\tag{355}
$$

On raising the index μ with the metric tensor, one has the contravariant second-rank tensor

$$
T^{\nu,\mu} = -\frac{1}{4\,\pi} F^{\nu,\rho} \left(\partial^\mu A_\rho \right) - g^{\nu,\mu} \, \mathcal{L}
\tag{356}
$$

The energy-momentum tensor is not gauge invariant, as it explicitly involves the fields A^μ. On using the expression for the source-free Lagrangian density

$$
\mathcal{L} = \frac{1}{8\,\pi} \left(\underline{E}^2 - \underline{B}^2 \right)
\tag{357}
$$

one finds that the time components of $T^{\mu,\nu}$ are given by

$$
T^{0,0} = \frac{1}{8\,\pi} \left(\underline{E}^2 + \underline{B}^2 \right) + \frac{1}{4\,\pi} \underline{\nabla} \cdot \left(\phi \, \underline{E} \right)
\tag{358}
$$

The expression $T^{0,0}$ is the Hamiltonian density $\mathcal{H}$, in the absence of sources, which represents the energy density of the free field. The momentum density is given by the mixed time and space components, and is given by

$$
T^{0,j} = -\frac{1}{4\,\pi} F^{0,\rho} \left(\partial^{(j)} A_\rho \right)
\tag{359}
$$

but since $F^{\mu,\nu}$ is antisymmetric, only the terms where ρ is a spatial index are non-zero. Hence, one has

$$
\begin{aligned}
T^{0,j} &= -\frac{1}{4\,\pi} \sum_i F^{0,i} \left(\partial^{(j)} A_i \right) \\
&= +\frac{1}{4\,\pi} \sum_i F^{0,i} \left(\partial^{(j)} A^{(i)} - \partial^{(i)} A^{(j)} \right) + \frac{1}{4\,\pi} \sum_i F^{0,i} \left(\partial^{(i)} A^{(j)} \right)
\end{aligned}
\tag{360}
$$

where the relation between the space-like components of the covariant and contravariant four-vector $A_i = - A^{(i)}$ has been used. Since the time component of the field tensor is given by

$$F^{0,i} = - E^{(i)} \tag{361}$$

and[10]

$$\left(\partial^{(i)} A^{(j)} - \partial^{(j)} A^{(i)} \right) = - \sum_k \xi^{i,j,k} B^{(k)} \tag{362}$$

one finds that the momentum density is given by

$$
\begin{aligned}
T^{0,j} &= - \frac{1}{4\pi} \sum_{i,k} \xi^{i,j,k} E^{(i)} B^{(k)} - \frac{1}{4\pi} \sum_i E^{(i)} \left(\partial^{(i)} A^{(j)} \right) \\
&= \frac{1}{4\pi} \left(\underline{E} \wedge \underline{B} \right)^{(j)} + \frac{1}{4\pi} \underline{E} \cdot \left(\underline{\nabla} A^{(j)} \right)
\end{aligned}
\tag{363}
$$

On noting that in the absence of sources, one has

$$\underline{\nabla} \cdot \underline{E} = 0 \tag{364}$$

and by adding a term proportional to $A^{(j)} \left(\underline{\nabla} \cdot \underline{E} \right)$ to the expression for $T^{0,j}$ in eqn(363), one arrives at the result

$$T^{0,j} = \frac{1}{4\pi} \left(\underline{E} \wedge \underline{B} \right)^{(j)} + \frac{1}{4\pi} \underline{\nabla} \cdot \left(A^{(j)} \underline{E} \right) \tag{365}$$

The components $T^{0,\nu}$, apart from the terms involving total derivatives which integrate out to zero, are related to the total energy and the components of the total momentum of the electromagnetic field. The components of $T^{\mu,\nu}$ satisfy the continuity equations

$$\partial_\mu T^{\mu,\nu} = 0 \tag{366}$$

which represent the conservation of energy and momentum. The other mixed time and spatial components of the energy-momentum tensor are evaluated as

$$T^{j,0} = \frac{1}{4\pi} \left(\underline{E} \wedge \underline{B} \right)^{(j)} + \frac{1}{4\pi} \left[\left(\underline{\nabla} \wedge \left(\phi \underline{B} \right) \right)^{(j)} - \frac{1}{c} \frac{\partial}{\partial t} \left(\phi E^{(j)} \right) \right] \tag{367}$$

The components $T^{j,0}$ represent the components of the energy flux.

It should be noted that the energy-momentum tensor $T^{\mu,\nu}$ is not symmetric. This has the consequence that the covariant generalization of the angular momentum to the third-rank tensor

$$M^{\mu,\nu,\rho} = T^{\mu,\nu} x^\rho - T^{\mu,\rho} x^\nu \tag{368}$$

---

[10]Since the vector relationship $\underline{B} = \underline{\nabla} \wedge \underline{A}$ involves the covariant derivative, there is a negative sign in the analogous expression involving the contravariant derivative.

is not conserved as the energy-momentum tensor is not symmetric. Additional terms can be added to the energy-momentum tensor[11], to create a symmetric tensor $\Theta^{\mu,\nu}$. These extra terms account for the intrinsic angular momentum of the photon.

The symmetric energy-momentum tensor $\Theta^{\mu,\nu}$ can be found by substituting

$$(\partial^\nu A^\lambda) \ = \ - \ F^{\lambda,\nu} \ + \ (\partial^\lambda A^\nu) \tag{369}$$

into the expression for $T^{\mu,\nu}$, to yield

$$T^{\mu,\nu} \ = \ \frac{1}{4\ \pi} \ \left[\ g^{\mu,\rho} \ F_{\rho,\lambda} \ F^{\lambda,\nu} \ + \ \frac{1}{4} \ g^{\mu,\nu} \ F_{\rho,\lambda} \ F^{\rho,\lambda} \ \right] \ - \ \frac{1}{4\ \pi} \ g^{\mu,\rho} \ F_{\rho,\lambda} \ (\partial^\lambda A^\nu) \tag{370}$$

The first two terms are symmetric w.r.t. μ and ν and are gauge invariant. These two terms will form the basis for $\Theta^{\mu,\nu}$, which will be expressed as

$$\Theta^{\mu,\nu} \ = \ \frac{1}{4\ \pi} \ \left[\ g^{\mu,\rho} \ F_{\rho,\lambda} \ F^{\lambda,\nu} \ + \ \frac{1}{4} \ g^{\mu,\nu} \ F_{\rho,\lambda} \ F^{\rho,\lambda} \ \right] \tag{371}$$

The expression $\Theta^{\mu,\nu}$ is symmetric under the interchange of μ and ν, as can be seen by writing

$$
\begin{aligned}
\Theta^{\mu,\nu} \ &= \ \frac{1}{4\ \pi} \ \left[\ F^\mu{}_\lambda \ F^{\lambda,\nu} \ + \ \frac{1}{4} \ g^{\mu,\nu} \ F_{\rho,\lambda} \ F^{\rho,\lambda} \ \right] \\
&= \ \frac{1}{4\ \pi} \ \left[\ F^{\mu,\lambda} \ F_\lambda{}^\nu \ + \ \frac{1}{4} \ g^{\mu,\nu} \ F_{\rho,\lambda} \ F^{\rho,\lambda} \ \right]
\end{aligned}
\tag{372}
$$

If $\Theta^{\mu,\nu}$ and $T^{\mu,\nu}$ are to represent the same set of conserved quantities, the last term in eqn(370) must be expressible as a total derivative. That this is true can be seen by examining the asymmetric term

$$- \ \frac{1}{4\ \pi} \ g^{\mu,\rho} \ F_{\rho,\lambda} \ (\partial^\lambda A^\nu) \ = \ - \ \frac{1}{4\ \pi} \ F^{\mu,\lambda} \ (\partial_\lambda A^\nu) \tag{373}$$

where the index ρ was raised by using the metric tensor. On combining the above expression with the source free Maxwell equation

$$(\partial_\lambda F^{\mu,\lambda}) \ = \ 0 \tag{374}$$

---

[11] J. Belinfante, Physica **6**, 887 (1939) has shown that the modified tensor $\Theta^{\mu,\nu}$ defined by

$$\Theta^{\mu,\nu} \ = \ T^{\mu,\nu} \ + \ \left(\partial_\rho \ \Lambda^{\rho;\mu,\nu} \right)$$

where $\Lambda^{\rho;\mu,\nu}$ is an arbitrary tensor that is antisymmetric under the interchange of the first pair of indices

$$\Lambda^{\rho;\mu,\nu} \ = \ - \ \Lambda^{\mu;\rho,\nu}$$

will automatically satisfy the same continuity conditions as $T^{\mu,\nu}$ and leave the total energy and momentum unaltered.

one obtains

$$-\frac{1}{4\,\pi}\,g^{\mu,\rho}\,F_{\rho,\lambda}\,(\partial^\lambda A^\nu) \;=\; -\frac{1}{4\,\pi}\left[\,F^{\mu,\lambda}\,(\partial_\lambda A^\nu)\;+\;A^\nu\,(\partial_\lambda F^{\mu,\lambda})\,\right]$$
$$=\;-\frac{1}{4\,\pi}\,\partial_\lambda\left(\,F^{\mu,\lambda}\,A^\nu\,\right) \tag{375}$$

which is a total covariant derivative of a third-rank tensor which is antisymmetric under the interchange of μ and λ. Furthermore, adding this term does to the energy-momentum tensor does not change the energy-momentum conservation law. This can be seen by observing that the difference between the two forms of energy-momentum conservation law involves the double derivative

$$\partial_\mu\,(\,\Theta^{\mu,\nu}\;-\;T^{\mu,\nu}\,)\;=\;-\frac{1}{4\,\pi}\,\partial_\mu\,\partial_\lambda\left(\,F^{\lambda,\mu}\,A^\nu\,\right) \tag{376}$$

and $F^{\lambda,\mu}$ is antisymmetric. On interchanging the order of the derivatives in the right hand side, switching the summation labels, and using the antisymmetric property of $F^{\lambda,\mu}$, one has

$$\partial_\mu\,(\,\Theta^{\mu,\nu}\;-\;T^{\mu,\nu}\,)\;=\;-\frac{1}{4\,\pi}\,\partial_\mu\,\partial_\lambda\left(\,F^{\lambda,\mu}\,A^\nu\,\right)$$
$$=\;-\frac{1}{4\,\pi}\,\partial_\lambda\,\partial_\mu\left(\,F^{\lambda,\mu}\,A^\nu\,\right)$$
$$=\;-\frac{1}{4\,\pi}\,\partial_\mu\,\partial_\lambda\left(\,F^{\mu,\lambda}\,A^\nu\,\right)$$
$$=\;+\frac{1}{4\,\pi}\,\partial_\mu\,\partial_\lambda\left(\,F^{\lambda,\mu}\,A^\nu\,\right) \tag{377}$$

On comparing the right hand sides of the first and last line, one finds that they have opposite signs and, therefore, they are zero. Thus, the difference between continuity relations vanish

$$\partial_\mu\,(\,\Theta^{\mu,\nu}\;-\;T^{\mu,\nu}\,)\;=\;0 \tag{378}$$

Hence, since $T^{\mu,\nu}$ is conserved, then the symmetrized energy-momentum tensor $\Theta^{\mu,\nu}$ is also conserved.

Thus, the symmetric energy-momentum tensor $\Theta^{\mu,\nu}$ expressed by

$$\Theta^{\mu,\nu}\;=\;\frac{1}{4\,\pi}\left[\,g^{\mu,\rho}\,F_{\rho,\lambda}\,F^{\lambda,\nu}\;+\;\frac{1}{4}\,g^{\mu,\nu}\,F_{\rho,\lambda}\,F^{\rho,\lambda}\,\right] \tag{379}$$

is a conserved quantity. The purely temporal component is given by

$$\Theta^{0,0}\;=\;\frac{1}{8\,\pi}\left(\,\underline{E}^2\;+\;\underline{B}^2\,\right) \tag{380}$$

and the mixed temporal and spatial components are given by

$$\Theta^{0,j} \;=\; \frac{1}{4\,\pi} \left(\, \underline{E} \;\wedge\; \underline{B} \,\right)^{(j)} \tag{381}$$

The temporal and spatial components of $\Theta^{0,\mu}$ are, respectively, recognized as being the energy-density and the momentum-density vector of the field. The components $\Theta^{j,0}$ are recognized as forming the Poynting vector which represents the energy flux of the electromagnetic field. The spatial components are given by

$$\Theta^{i,j} \;=\; -\,\frac{1}{4\,\pi} \left[\, E^{(i)}\,E^{(j)} \;+\; B^{(i)}\,B^{(j)} \;-\; \frac{1}{2}\,\delta^{i,j}\,(\,\underline{E}^2 \;+\; \underline{B}^2\,)\,\right] \tag{382}$$

Noether's theorem is purely classical, but there are generalizations for quantum fields. Quantum generalizations includes the Ward-Takahashi and Taylor-Slavnov identities.

Exercise:

Evaluate the components $T^{j,0}$ and $T^{i,j}$ of the (asymmetric) energy-momentum tensor for a source-free electromagnetic field.

Exercise:

Show that in the presence of sources, the symmetric energy-momentum tensor has components with the form

$$\begin{aligned}
\Theta^{0,0} &\;=\; \frac{1}{8\,\pi}\left(\,\underline{E}^2 \;+\; \underline{B}^2\,\right) \;-\; \frac{1}{c}\,\underline{j}\cdot\underline{A} \\[4pt]
\Theta^{0,j} &\;=\; \frac{1}{4\,\pi}\left(\,\underline{E}\;\wedge\;\underline{B}\,\right)^{(j)} \;-\; \rho\,\underline{A}^{(j)}
\end{aligned} \tag{383}$$

Verify the form of the conservation laws for energy and momentum.

Exercise:

Show that the extra term added to the tensor $T^{i,j}$ in order that $\Theta^{i,j}$ will be symmetric produces a contribution to the angular momentum density of the form

$$\mathcal{S}^{0,j} \;=\; \frac{1}{4\,\pi}\left(\,\underline{E}\;\wedge\;\underline{A}\,\right)^{(j)} \tag{384}$$

which is the intrinsic spin density of the electromagnetic field.

7.2 Massive Spin-One Particles

The electromagnetic theory has been unified with the theory of weak interactions. This generalization requires the existence of two new types of spin-one particles in addition to the photon, which together mediate the electro-weak interaction. These new particles have non-zero mass. The massive spin-one particle has to satisfy the equation[12]

$$p^\mu \, p_\mu \;=\; m^2 \, c^2 \tag{385}$$

and with the quantization condition,

$$p^\mu \;\rightarrow\; \hat{p}^\mu \;=\; i \, \hbar \, \frac{\partial}{\partial x_\mu} \tag{386}$$

the four-vector field A^μ must satisfy the Klein-Gordon equation

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \;-\; \underline{\nabla}^2 \;+\; \left(\frac{m \, c}{\hbar} \right)^2 \right] A^\mu \;=\; \frac{4 \, \pi}{c} \, j^\mu \tag{387}$$

where $\hbar$ no longer drops out. This equation can be derived from the Lagrangian

$$\mathcal{L} \;=\; - \, \frac{1}{16 \, \pi} \, F^{\mu,\nu} \, F_{\mu,\nu} \;+\; \frac{1}{8 \, \pi} \left(\frac{m \, c}{\hbar} \right)^2 A^\mu \, A_\mu \;-\; \frac{1}{c} \, j^\mu \, A_\mu \tag{388}$$

For example, on varying A^μ, one obtains the equation of motion

$$\partial_\nu F^{\nu,\mu} \;+\; \left(\frac{m \, c}{\hbar} \right)^2 A^\mu \;=\; \frac{4 \, \pi}{c} \, j^\mu \tag{389}$$

Neither the Lagrangian, nor the equation of motion are gauge invariant. The appropriate gauge condition can be enforced by imposing conservation of charge[13]

$$\partial_\mu j^\mu \;=\; 0 \tag{390}$$

On taking the four-divergence of the equation of motion, one finds

$$\partial_\mu \partial_\nu F^{\nu,\mu} \;+\; \left(\frac{m \, c}{\hbar} \right)^2 \partial_\mu A^\mu \;=\; \frac{4 \, \pi}{c} \, \partial_\mu j^\mu \tag{391}$$

The first term on the left-hand side vanishes due to the definition of $F^{\mu,\nu}$, since

$$F^{\nu,\mu} \;=\; \partial^\nu A^\mu \;-\; \partial^\mu A^\nu \tag{392}$$

one finds

$$\partial_\nu F^{\nu,\mu} \;=\; \partial_\nu \partial^\nu A^\mu \;-\; \partial^\mu \partial_\nu A^\nu \tag{393}$$

---

[12]A. Proca, J. Phys. et Radium **7**, 147 (1936).

[13]Note that, unlike the massless photon, charge conservation has to be imposed as an additional assumption.

therefore

$$
\begin{aligned}
\partial_\mu \partial_\nu F^{\nu,\mu} &= \partial_\nu \partial^\nu \partial_\mu A^\mu - \partial_\mu \partial^\mu \partial_\nu A^\nu \\
&= 0
\end{aligned}
\tag{394}
$$

The term on the right-hand side of eqn(391) also vanishes, because it was chosen to impose charge conservation. Hence, one finds that A^μ for a massive spin-one particle must satisfy the Lorenz gauge condition

$$
\partial_\mu A^\mu = 0
\tag{395}
$$

Exercise:

Starting from the expression eqn(379), determine the symmetrized energy-momentum tensor for the massive vector field. Hence, find the energy and momentum densities.

7.3 Polarizations of Massive Spin-One Particles

The four-vector potential with contravariant components $A^\mu(x)$ can be Fourier transformed as

$$
A^\mu(x) = \frac{1}{\sqrt{2 V}} \sum_{\underline{k}} \left(A^\mu(\underline{k}) \, \exp[\, i \, k^\nu \, x_\nu \,] \; + \; \text{c.c.} \right)
\tag{396}
$$

which results in four components associated with four polarization vectors which are denoted by $\hat{e}^\mu(\underline{k})$. For massive photons, if one assumes charge conservation, the gauge fields must satisfy the Lorentz gauge condition. The Lorentz gauge condition

$$
\partial_\mu \, A^\mu(x) = 0
\tag{397}
$$

results in the Fourier components satisfying the condition

$$
k_\mu \, A^\mu(\underline{k}) = 0
\tag{398}
$$

which involving the photon's four-momentum. Apparently, the gauge condition reduces the number of independent components of the four-vector potential from four to three. However, we are used to thinking that massless photons only have two independent polarizations. In the following, we shall see how this comes about.

For a photon with mass m, one can apply the Lorentz gauge condition in the laboratory frame. We shall choose the z-axis as the direction of the photon's three-momentum $\underline{k}$. With this choice, the photon's four-momentum has

components $(k^{(0)}, 0, 0, k^{(3)})$ where $k^{(0)} = \sqrt{(\frac{mc}{\hbar})^2 + \underline{k}^2}$. The Lorentz condition becomes

$$k^{(0)} \, A^{(0)}(\underline{k}) \; - \; k^{(3)} \, A^{(3)}(\underline{k}) \; = \; 0 \tag{399}$$

Hence, $A^{(0)}(\underline{k}) = \frac{k^{(3)}}{k^{(0)}} A^{(3)}(\underline{k})$. This implies that in the photon's rest frame, the four-vector potential only has three spatial components and the scalar potential is zero. In any case, there are only three-independent components of the four-vector potential. The four-vector potential can be expressed in terms of the three independent components as

$$
\begin{aligned}
A(x) \;\; = \;\; & \frac{1}{\sqrt{2\,V}} \sum_{\underline{k}} \Bigg[\left(A^{(3)}(\underline{k}) \, (\frac{k^{(3)}}{k^{(0)}} \, \hat{e}^{(0)}(\underline{k}) \; + \; \hat{e}^{(3)}(\underline{k}) \,) \right. \\
& \left. + \; A^{(1)}(\underline{k}) \, \hat{e}^{(1)}(\underline{k}) \; + \; A^{(2)}(\underline{k}) \, \hat{e}^{(2)}(\underline{k}) \right) \; \exp[\; i \; k^{\nu} \; x_{\nu} \;] \; + \; \text{c.c.} \Bigg]
\end{aligned}
\tag{400}
$$

Thus, the four-vector associated with the (longitudinal) contravariant component $A^{(3)}(\underline{k})$ has a polarization which is proportional to $\frac{k^{(3)}}{k^{(0)}} \, \hat{e}^{(0)}(\underline{k}) \; + \; \hat{e}^{(3)}(\underline{k})$. The vector with components $A^{(3)}(\underline{k})$ is physical since it lead to an electromagnetic field. The electric and magnetic fields are defined by

$$
\begin{aligned}
\underline{E} \;\; &= \;\; - \, \frac{1}{c} \, \frac{\partial}{\partial t} \, \underline{A} \; - \; \underline{\nabla} \, \phi \\
\underline{B} \;\; &= \;\; \underline{\nabla} \, \wedge \, \underline{A}
\end{aligned}
\tag{401}
$$

The longitudinal component of the electric field has a Fourier amplitude given by

$$
\begin{aligned}
E^{(3)}(\underline{k}) \;\; &= \;\; - \, i \left(k^{(0)} \, A^{(3)}(\underline{k}) \; - \; k^{(3)} \, A^{(0)}(\underline{k}) \right) \\
&= \;\; - \, i \left(k^{(0)} \; - \; \frac{k^{(3)2}}{k^{(0)}} \right) \, A^{(3)}(\underline{k}) \\
&= \;\; - \, i \, \frac{\left(\frac{m\,c}{\hbar} \right)^2}{k^{(0)}} \, A^{(3)}(\underline{k})
\end{aligned}
\tag{402}
$$

and the longitudinal component of the magnetic field $B^{(3)}(\underline{k})$ is zero since

$$0 \;\; = \;\; - \, i \left(k^{(3)} \, \hat{e}^{(3)} \, \wedge \, \underline{A}(\underline{k}) \right) . \, \hat{e}^{(3)}(\underline{k}) \tag{403}$$

Thus, the component $A^{(3)}(\underline{k})$ contributes to the $\underline{E}$ field but not the $\underline{B}$ field. For massive photons, the unsymmmetrized energy-momentum tensor $T^{\mu,\nu}$ is expressed in terms of contributions from, not only the $\underline{E}$ and $\underline{B}$ fields, but also

from the gauge field A^μ. The tensor is evaluated as

$$T^{\mu,\nu} = \frac{1}{16\,\pi}\left(g^{\mu,\nu}\,F_{\rho,\sigma}\,F^{\rho,\sigma} - 4\,F^{\mu,\rho}\,\partial^\nu\,A_\rho\right) - \frac{g^{\mu,\nu}}{8\,\pi}\left(\frac{m\,c}{\hbar}\right)^2 A^\rho\,A_\rho + g^{\mu,\nu}\,\frac{j^\rho\,A_\rho}{c} \tag{404}$$

The energy-momentum tensor can be "symmetrized" by adding and subtracting a term

$$+\,\frac{4}{16\,\pi}\,F^{\mu,\rho}\,\partial_\rho\,A^\nu \tag{405}$$

to express the second term in a gauge-invariant form. We then note that that remaining term (with the negative sign) could be put into the form of a divergence of an antisymmetric third rank-tensor, if we could combine it with a term of the form

$$-\frac{1}{4\,\pi}\,(\,\partial_\rho\,F^{\mu,\rho}\,)\,A^\nu \;=\; +\frac{1}{4\,\pi}\,(\,\partial_\rho\,F^{\rho,\mu}\,)\,A^\nu \tag{406}$$

This can be accomplished by adding zero in the form of A^ν times the Euler-Lagrange equation

$$+\,\partial_\rho\,F^{\rho,\mu}\;+\;\left(\frac{m\,c}{\hbar}\right)^2 A^\mu\;-\;\frac{4\,\pi}{c}\,j^\mu\;=\;0 \tag{407}$$

therefore, completing the divergence at the expense of adding extra mass and source terms. The divergence of third-rank antisymmetric tensor can be dropped, leading to the "symmetrized" energy-momentum tensor given by

$$\begin{aligned}
\Theta^{\mu,\nu} \;=\;\; & \frac{1}{16\,\pi}\left(g^{\mu,\nu}\,F_{\rho,\sigma}\,F^{\rho,\sigma}\;+\;4\,F^{\mu,\rho}\,F_\rho{}^\nu\right) \\
& -\frac{g^{\mu,\nu}}{8\,\pi}\left(\frac{m\,c}{\hbar}\right)^2 A^\rho\,A_\rho\;+\;\frac{1}{4\,\pi}\left(\frac{m\,c}{\hbar}\right)^2 A^\mu\,A^\nu \\
& +\,g^{\mu,\nu}\,\frac{j^\rho\,A_\rho}{c}\;-\;\frac{j^\mu\,A^\nu}{c}
\end{aligned} \tag{408}$$

It is seen that the coupling of the electromagnetic fields to the current densities spoils the symmetry of the energy-momentum tensor. Hence, the charge currents act as sources of angular momentum and results in the electromagnetic field's angular momentum not being conserved. The energy-density $\mathcal{H}$ is given by $\Theta^{0,0}$ so

$$\mathcal{H} \;=\; \frac{1}{8\,\pi}\left[\,\underline{E}^2\;+\;\underline{B}^2\,\right]\;+\;\frac{1}{8\,\pi}\left(\frac{m\,c}{\hbar}\right)^2\left(A^{(0)\,2}\;+\;\underline{A}^2\right)\;-\;\frac{\underline{j}\cdot\underline{A}}{c} \tag{409}$$

and the momentum density P is given by

$$P \;=\; \frac{1}{4\,\pi\,c}\left(\underline{E}\;\wedge\;\underline{B}\right)\;+\;\frac{1}{4\,\pi\,c}\left(\frac{m\,c}{\hbar}\right)^2 A^{(0)}\,\underline{A}\;-\;\frac{j^{(0)}\,\underline{A}}{c^2} \tag{410}$$

Thus, the longitudinal photon with components $A^{(3)}(\underline{k})$ and $A^{(0)}(\underline{k})$ does have physical effects, as do the two transverse photons $A^{(1)}(\underline{k})$ and $A^{(2)}(\underline{k})$. Therefore, the four-vector potential of a massive electromagnetic field has three physical components.

Exercise:

Determine the contribution to the energy and momentum of the electromagnetic field which originates from the coupled time and space-like (longitudinal) components of the four-vector potential.

For a massless photon $m = 0$, so one sees that the longitudinal component of the electric field $E^{(3)}(\underline{k})$ is zero. Also since for $m = 0$ one has $k^{(0)} = k^{(3)}$, so the Lorentz gauge condition requires that the component $A^{(3)}(\underline{k})$ is equal to $A^{(0)}(\underline{k})$ and, therefore, the longitudinal component is associated with a single polarization vector ($\hat{e}^{(0)}(\underline{k}) + \hat{e}^{(3)}(\underline{k})$). Since $A^{(0)}(\underline{k}) = A^{(3)}(\underline{k})$, the norm of the vector $(A^{(0)}(\underline{k}), 0, 0, A^{(3)}(\underline{k}))$ is identical zero. For $m = 0$, the longitudinal component $A^{(3)}(\underline{k})$ neither contributes to the $\underline{E}$ field nor to the $\underline{B}$ field. Since the energy and momentum densities are expressed in terms of the $\underline{E}$ and $\underline{B}$ fields via

$$\mathcal{H} = \frac{1}{8\,\pi} \left[\underline{E}^2 + \underline{B}^2 \right] \tag{411}$$

and

$$P = \frac{1}{4\,\pi\,c} \left(\underline{E} \wedge \underline{B} \right) \tag{412}$$

the component $A^{(3)}(\underline{k})$ or equivalently $A^{(0)}(\underline{k})$ has no physical effect. Hence, it can be deemed unphysical. Therefore, the four-vector potential of a massless electromagnetic field only has two physical components $A^{(1)}(\underline{k})$ and $A^{(2)}(\underline{k})$ which are the components transverse to the direction of propagation.

7.4   The Propagator for Massive Photons.

The four-vector potential can be expressed entirely in terms of the currents, via the introduction of a photon propagator. The equation of motion for the massive spin-one particle is

$$\partial_\mu \, F^{\mu,\nu} + \left(\frac{m\,c}{\hbar} \right)^2 A^\nu = \frac{4\,\pi}{c} \, j^\nu \tag{413}$$

On Fourier transforming the linear equation with respect to space and time, one obtains

$$\left[- k_\mu \, k^\mu + \left(\frac{m\,c}{\hbar} \right)^2 \right] A^\nu(k) + k_\mu \, k^\nu \, A^\mu(k) = \frac{4\,\pi}{c} \, j^\nu(k) \tag{414}$$

which can be re-written as

$$\left[\left(- k_\rho \, k^\rho + \left(\frac{m \, c}{\hbar} \right)^2 \right) \delta_\mu^\nu + k_\mu \, k^\nu \right] A^\mu(k) = \frac{4 \, \pi}{c} \, j^\nu(k) \qquad (415)$$

The photon propagator is defined via

$$A^\mu(k) = D^{\mu,\lambda}(k) \, j_\lambda(k) \qquad (416)$$

which leads to the phonon propagator being determined as the solution of a matrix equation

$$\left[\left(- k_\rho \, k^\rho + \left(\frac{m \, c}{\hbar} \right)^2 \right) g^{\nu,\mu} + k^\nu \, k^\mu \right] D_{\mu,\lambda}(k) = \frac{4 \, \pi}{c} \, \delta_\lambda^\nu \qquad (417)$$

The above matrix equation can be solved by first writing the propagator as

$$D^{\mu,\nu}(k) = g^{\mu,\nu} \, A(k^2) + k^\mu \, k^\nu \, B(k^2) \qquad (418)$$

and then solving for the unknown quantities $B(k^2)$ and $A(k^2)$. This leads to the following expression for the photon propagator

$$D^{\mu,\nu}(k) = \left(\frac{4\pi}{c} \right) \frac{\left(g^{\mu,\nu} - \frac{k^\mu \, k^\nu}{(\frac{m \, c}{\hbar})^2} \right)}{\left[\left(\frac{m \, c}{\hbar} \right)^2 - k^\rho \, k_\rho \right]} \qquad (419)$$

Thus, the Fourier component of the four-vector potential is given by

$$A^\mu(k) = \frac{\left(g^{\mu,\nu} - \frac{k^\mu \, k^\nu}{(\frac{m \, c}{\hbar})^2} \right)}{\left[\left(\frac{m \, c}{\hbar} \right)^2 - k^\rho \, k_\rho \right]} \left(\frac{4\pi}{c} \right) j_\nu(k) \qquad (420)$$

Hence, $A^\mu(x)$ can be found from the inverse Fourier transform. It should be noted that, if one imposes current conservation

$$k^\nu \, j_\nu(k) = 0 \qquad (421)$$

the second term in the numerator of the photon propagator has no physical effect. We also note that

$$k_\mu \, A^\mu(k) = \left(\frac{4 \, \pi}{c} \right) \left(\frac{\hbar}{m \, c} \right)^2 k^\nu \, j_\nu(k) \qquad (422)$$

So, once again we see that assumption of charge conservation enforces the Lorentz gauge condition, even if the massive photon is involved in a virtual process.

8 Symmetry Breaking and Mass Generation

We shall first look at an example of Goldstone's theorem which states that, if a system described by a Lagrangian which has a continuous symmetry (and only short-ranged interactions) has a broken symmetry state then the system supports a branch of small amplitude excitations with a dispersion relation ω_k that vanishes at $k = 0$. We shall then examine the situation in which the system is coupled by long-ranged interactions, as modelled by an electromagnetic field. As was first pointed out by Anderson, the long-ranged interactions alter the excitation spectrum of the symmetry broken state by removing the Goldstone modes and generating a branch of massive excitations.

8.1 Symmetry Breaking and Goldstone Bosons

Consider a Lagrangian density for a complex scalar field of the form

$$
\mathcal{L} \; = \; (\; \partial_\mu \psi^* \;) \; (\; \partial^\mu \psi \;) \; - \; \left(\frac{m \; c}{2 \; \hbar \; \phi_0} \right)^2 \left(\; \psi^* \; \psi \; - \; \phi_0^2 \; \right)^2 \tag{423}
$$

The Lagrangian density is invariant under the continuous global symmetry

$$
\psi \; \rightarrow \psi' \; = \; \exp \left[\; - \; i \; \alpha \; \right] \psi \tag{424}
$$

for any real constant α. The static or minimum energy solution corresponds to

$$
| \; \psi \; | \; = \; \phi_0 \tag{425}
$$

which leaves the phase of ψ undetermined. Since the phase of ψ is continuous, the ground state is infinitely degenerate. If one writes

$$
\begin{aligned}
\psi \; &= \; \phi_1 \; + \; i \; \phi_2 \\
\psi^* \; &= \; \phi_1 \; - \; i \; \phi_2
\end{aligned} \tag{426}
$$

then the Lagrangian can be written as a Lagrangian density involving the two real scalar fields ϕ_1 and ϕ_2. The Lagrangian density has a $U(1)$ symmetry which corresponds to the rotation of ψ around a circle about the origin in the (ϕ_1, ϕ_2) plane.

We shall assume the field ψ representing the physical ground state corresponds to only one of the infinite number of possible candidates. The physical state must have a phase, which shall be defined as zero. That is, one starts with a ground state $\psi = \phi_0$, and then consider the small amplitude excitations. A low-energy excited state corresponds to the complex field

$$
\psi \; = \; \phi_0 \; + \; \delta\psi \tag{427}
$$
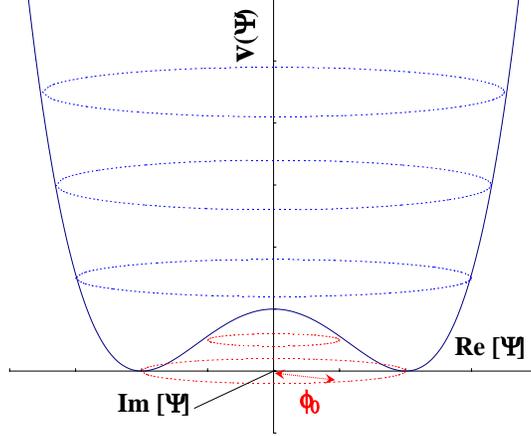
Figure 8: The potential $V[\psi]$ described by the Lagrangian is invariant under global rotations of the phase of ψ. The minima occurs at values of ψ which have magnitudes ϕ_0, therefore, the uniform static field is infinitely degenerate.

where $\delta\psi$ is static and uniform and can be considered to be very small. The small amplitude complex field $\delta\psi$ can be expressed in terms of its real and imaginary parts

$$\delta\psi \; = \; \chi_1 \; + \; i \; \chi_2 \tag{428}$$

The Lagrangian density takes the form

$$\mathcal{L} \; = \; (\, \partial_\mu \chi_1 \,)\,(\, \partial^\mu \chi_1 \,) + (\, \partial_\mu \chi_2 \,)\,(\, \partial^\mu \chi_2 \,) - \left(\frac{m\,c}{2\,\hbar\,\phi_0}\right)^2 \left(\, 2\,\phi_0\,\chi_1 + \chi_1^2 + \chi_2^2 \,\right)^2 \tag{429}$$

If one only considers infinitesimally small amplitude oscillations, one only needs consider terms quadratic in the fields. The quadratic Lagrangian density $\mathcal{L}_{Free}$ describes non-interacting fields. The quadratic Lagrangian density is given by

$$\mathcal{L}_{Free} \; = \; (\, \partial_\mu \chi_1 \,)\,(\, \partial^\mu \chi_1 \,) \; - \; \left(\frac{m\,c}{\hbar}\right)^2 \chi_1^2 \; + \; (\, \partial_\mu \chi_2 \,)\,(\, \partial^\mu \chi_2 \,) \tag{430}$$

The symmetry breaking has resulted in the complex field breaking up into two fields: The first field χ_1 describes massive excitations m and the second field χ_2 describes massless excitations. The first field χ_1 has plane-wave solutions if the energy and momentum are related via the dispersion relation

$$\omega^2 \; = \; c^2\,k^2 \; + \; \left(\frac{m\,c^2}{\hbar}\right)^2 \tag{431}$$

and represents excitations which corresponds to a "stretching" of ϕ_0. It is massive since this excitation moves the field away from the minimum of the potential. The second excitation χ_2 represents $\delta\psi$ which is transverse to ϕ_0 in

71

the (ϕ_1, ϕ_2) plane. This last excitation is known as a Goldstone boson[14]. The Goldstone boson has a dispersion relation

$$\omega^2 \; = \; c^2 \; k^2 \tag{432}$$

which vanishes at $k = 0$. The Goldstone boson dynamically restores the spontaneously broken $U(1)$ symmetry since, at $k = 0$, it just corresponds to a change of the value of the (static and uniform) broken symmetry field from $(\phi_0, 0)$ to the new direction (ϕ_0, χ_2). Therefore, if infinitely many zero-energy Goldstone bosons are excited in the system, the resulting state should correspond to a new ground state with a different value of the phase. As noted by Anderson[15] prior to Goldstone's work, the Goldstone theorem breaks down when long-ranged interactions are present. Anderson was responsible for the concept of mass generation through symmetry breaking due to the coupling with gauge fields[16]. This concept was subsequently developed by Peter Higgs[17] and Tom Kibble and coworkers[18].

8.2 The Kibble-Higgs Mechanism

We shall now consider the coupling of a scalar field ψ with charge q to a gauge field A^μ. The Lagrangian density is related to the sum of the Lagrangian density for the electromagnetic field and the Lagrangian density for the charged scalar particle. The coupling between the fields is found from the minimum coupling assumption

$$\hat{p}^\mu \; \rightarrow \; \hat{p}^{\mu\prime} \; = \; \hat{p}^\mu \; - \; \frac{q}{c} \; A^\mu \tag{433}$$

which becomes

$$i \; \hbar \; \partial^\mu \; \rightarrow \; i \; \hbar \; \partial^\mu \; - \; \frac{q}{c} \; A^\mu \tag{434}$$

Therefore, the Lagrangian density for the coupled fields has the form

$$\begin{aligned}
\mathcal{L} \; = \; & \left(\partial_\mu \; - \; i \; \frac{q}{\hbar \; c} \; A_\mu \right) \psi^* \left(\partial^\mu \; + \; i \; \frac{q}{\hbar \; c} \; A^\mu \right) \psi \; - \; \frac{1}{16 \; \pi} \; F^{\mu,\nu} \; F_{\mu,\nu} \\
& - \; \left(\frac{m \; c}{2 \; \hbar \; \phi_0} \right)^2 \left(\psi^* \; \psi \; - \; \phi_0^2 \right)^2
\end{aligned} \tag{435}$$

[14] J. Goldstone, Il Nuovo Cimento, **19**, 154 (1961).

[15] P. W. Anderson, Phys. Rev., **112** 1900 (1958).

[16] P.W. Anderson, "Plasmons, Gauge Invariance, and Mass". Physical Review 130, 439 (1963).

[17] P.W. Higgs, "Broken Symmetries, Massless Particles and Gauge Fields", Physics Letters, 12, 132 (1964): P.W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons". Physical Review Letters 13, 508 (1964):P.W. Higgs, "Spontaneous Symmetry Breaking without massless Bosons", Physical Review 145, 1156 (1966)

[18] G.S. Guralnik, C.R. Hagen, and T.W.B. Kibble, "Global Conservation Laws and Massless Particles". Physical Review Letters 13, 585-587 (1964): T.W.B. Kibble, "Symmetry Breaking in non-Abelian Gauge Theories", Physical Review, 155, 1554 (1967).

The Lagrangian density is invariant under the local gauge transformation[19]

$$
\begin{aligned}
\psi \quad &\rightarrow \quad \psi^{'} \;=\; \exp\left[\,-\, i\, \frac{q}{\hbar\, c}\, \Lambda\, \right]\, \psi \\
A^{\mu} \quad &\rightarrow \quad A^{\mu\prime} \;=\; A^{\mu} \,+\, \partial^{\mu}\Lambda
\end{aligned}
\tag{436}
$$

where $\Lambda = \Lambda(x^{\mu})$ is a function of space-time. The system has minimum energy when ψ has a constant value with a magnitude given by

$$
\mid \psi \mid \;=\; \phi_0
\tag{437}
$$

and the A_{μ} vanish. Any local gauge transformation leads to a state with the same energy, therefore, the ground state is infinitely degenerate.

We shall assume that a physical system spontaneously breaks the symmetry in that it corresponds to a specific constant value of Λ. We shall choose the local gauge $\Lambda(x)$ such that the field ψ representing the excited states is purely real. However, once the gauge has been fixed, no further gauge transformations can be made.

The small amplitude excitations can be expressed as

$$
\psi \;=\; \phi_0 \,+\, \delta\psi
\tag{438}
$$

The fluctuations can be expressed as

$$
\delta\psi \;=\; \chi_1
\tag{439}
$$

and on substituting in the Lagrangian and collecting the quadratic terms, one obtains

$$
\begin{aligned}
\mathcal{L}_{Free} \;=\; &(\, \partial_{\mu}\chi_1\,)\,(\,\partial^{\mu}\chi_1\,) \,-\, \left(\frac{m\,c}{\hbar}\right)^2 \chi_1^2 \\
&-\, \frac{1}{16\,\pi}\, F^{\mu,\nu}\, F_{\mu,\nu} \,+\, \left(\frac{q\,\phi_0}{\hbar\,c}\right)^2 A_{\mu}\, A^{\mu}
\end{aligned}
\tag{440}
$$

Therefore, one finds that the charged boson field has a mass m and the gauge field has acquired a mass m_A given by

$$
m_A^2 \;=\; 8\,\pi\, \left(\frac{q\,\phi_0}{c^2}\right)^2
\tag{441}
$$

Hence, by coupling an electromagnetic field with two components to a scalar charged boson field, one has found a massive vector boson gauge-field with three independent components. The massless spin-less component of the charged boson field which described the Goldstone mode has become the longitudinal mode

---

[19]V. Fock, Z. Phys., **39**, 226, (1927).

of the gauge field. More specifically, the field χ_2 which initially corresponded to the Goldstone mode became unphysical when the massless vector field was introduced as it could be gauged away. This is seen by writing

$$\psi \;=\; \phi_0 \;+\; \chi_1 \;+\; i\,\chi_2 \;\approx\; \left(\,\phi_0 \;+\; \chi_1\,\right)\, \exp\left[\, i\,\frac{\chi_2}{\phi_0}\,\right] \qquad (442)$$

so χ_2 could be removed by a gauge transformation involving the particular choice of Λ

$$\Lambda \;=\; \frac{\chi_2}{\left(\dfrac{q\,\phi_0}{\hbar\,c}\right)} \qquad (443)$$

9 Gravitational Interactions

We introduce the field theory of the Gravitational Interaction, in the weak field limit, by analogy with electromagnetism. More specifically, we shall develop the classical field theory of the massive graviton in parallel with the Proca's theory of a massive spin-one particle that we discussed previously. This approach has the disadvantage that it does not have the same beautiful geometric basis as Einstein's field equations.

Electromagnetic fields are mediated by spin-one bosons and gravitational fields are mediated by spin-two bosons. Electromagnetic forces between two-particles with like charges is repulsive but Gravitational forces are always attractive. These facts are quantified by Coulomb's law

$$V(\underline{r}_1 - \underline{r}_2) \;=\; \frac{q_1\,q_2}{\mid \underline{r}_1 \;-\; \underline{r}_2 \mid} \qquad (444)$$

and Newton's Force law

$$V(\underline{r}_1 - \underline{r}_2) \;=\; -\;\frac{G\,m_1\,m_2}{\mid \underline{r}_1 \;-\; \underline{r}_2 \mid} \qquad (445)$$

where G is the gravitational constant. The value of G was first determined as $G \;=\; 6.754 \;\times\; 10^{-11} \; m^3 \; sec^{-2} \; kgm^{-1}$ by Henry Cavendish in 1798. However, G remains the most poorly known physical constant. Recent high precision measurements of G yield values with error bars that are mutually exclusive. The high precision values differ in the third significant digit. The above two potentials show that the electric and gravitational fields have the same type of classical Green's functions associated with massless bosons. Therefore, we expect that the difference in the signs of the interaction is caused by the spin values of the bosons involved in the exchanges. We shall examine the difference of the signs of the electromagnetic and gravitational forces by examining the

quantum propagators.

Consider (massive) electromagnetic radiation in the presence of a source. The Lagrangian density formulated by Proca[20] is given by

$$\mathcal{L} = -\frac{1}{16\pi} F^{\mu,\nu} F_{\mu,\nu} + \frac{1}{8\pi} \left(\frac{m\,c}{\hbar}\right)^2 A^\mu A_\mu - \frac{1}{c} A^\mu j_\mu \tag{446}$$

where

$$F^{\mu,\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{447}$$

The mass term spoils the gauge invariance of the source-free Lagrangian. The action S is given by

$$S = \int d^4x\, \mathcal{L} \tag{448}$$

which, on integrating by parts, can be expressed as

$$S = \int d^4x \left[\frac{1}{8\pi} A_\mu \left(\partial_\rho \partial^\rho + \left(\frac{m\,c}{\hbar}\right)^2 \right) g^{\mu,\nu} - \partial^\mu \partial^\nu - \frac{1}{c} j^\nu \right] A_\nu \tag{449}$$

The equation of motion for the four-vector potential is

$$\left[\left(\partial_\rho \partial^\rho + \left(\frac{m\,c}{\hbar}\right)^2 \right) g^{\mu,\nu} - \partial^\mu \partial^\nu \right] A_\nu = \frac{4\pi}{c} j^\mu \tag{450}$$

If charge is conserved

$$\partial_\mu j^\mu = 0 \tag{451}$$

On operating on the equation of motion with ∂_μ, one finds that because of the finite mass the field A_ν must satisfy the condition

$$\partial^\nu A_\nu = 0 \tag{452}$$

Thus, conservation of charge removes the choice of gauge. In the rest frame of a photon, $k = (\frac{m\,c}{\hbar})(1,0,0,0)$, so one has $A_0(k) = 0$ and so one finds that the massive photon has three components. This is expected for a particle with spin $S = 1$. The electromagnetic propagator can be defined as the kernel of the integral relation

$$A_\nu(x) = \frac{4\pi}{c} \int d^4x'\, D_{\nu,\lambda}(x,x')\, j^\lambda(x') \tag{453}$$

Thus, the propagator satisfies the partial differential equation

$$\left[\left(\partial_\rho \partial^\rho + \left(\frac{m\,c}{\hbar}\right)^2 \right) g^{\mu,\nu} - \partial^\mu \partial^\nu \right] D_{\nu,\lambda}(x,x') = \delta^\mu_\lambda\, \delta^4(x - x') \tag{454}$$

[20]A. Proca, C.R. Acad. Sci., Paris **203**, 709-711 (1936), J. de Physique et Radium, **8**, 23-28 (1937).

which can be solved by Fourier transforming. On performing the Fourier transformation, one finds the algebraic equation

$$\left[\left(- k_\rho \ k^\rho + \left(\frac{m \ c}{\hbar} \right)^2 \right) g^{\mu,\nu} + k^\mu \ k^\nu \right] D_{\nu,\lambda}(k) = \delta_\lambda^\mu \qquad (455)$$

which can be solved and has the solution

$$D_{\nu,\lambda}(k) = \left(\frac{- g_{\nu,\lambda} + \frac{k_\nu \ k_\lambda}{(\frac{m \ c}{\hbar})^2}}{k^\rho \ k_\rho - (\frac{m \ c}{\hbar})^2} \right) \qquad (456)$$

We shall determine the interaction mediated by the photon, by considering the contribution to the energy density from two charges. For a static charge density, the interaction energy can be expressed as minus half the interaction Lagrangian. For static charges, the kinetic energy terms of the free-field Lagrangian do not contribute to the energy. Furthermore, since the static field free Lagrangian is quadratic in the field, the equation of motion can be used to show that it cancels half the interaction term which is linear in the field. The interaction Lagrangian can be written as

$$L_{int} = - \frac{1}{c} \int d^3\underline{r} \ j_\mu \ A^\mu \qquad (457)$$

which, for $t = 0$, becomes

$$L_{int} = - \frac{1}{c} \int \frac{dk_0'}{(2 \ \pi)} \int \frac{d^4k}{(2 \ \pi)^4} \ j_\mu^*(k_0', \underline{k}) \ A^\mu(k_0, \underline{k}) \qquad (458)$$

On eliminating the four-vector potential, the interaction portion of the Lagrangian can be expressed as

$$L_{int} = - \frac{1}{c} \int \frac{dk_0'}{(2 \ \pi)} \int \frac{d^4k}{(2 \ \pi)^4} \ j^\nu(k_0', \underline{k})^* \ \frac{4 \ \pi}{c} \left(\frac{- g_{\nu,\lambda} + \frac{k_\nu \ k_\lambda}{(\frac{m \ c}{\hbar})^2}}{k^\rho \ k_\rho - (\frac{m \ c}{\hbar})^2} \right) j^\lambda(k_0, \underline{k}) \qquad (459)$$

Using the condition of continuity of charge

$$k_\nu \ j^\nu(k) = 0 \qquad (460)$$

the interaction can be written as

$$\begin{aligned} L_{int} &= - \frac{1}{c} \int \frac{dk_0'}{(2 \ \pi)} \int \frac{d^4k}{(2 \ \pi)^4} \ j^\nu(k_0', \underline{k})^* \ \frac{4 \ \pi}{c} \left(\frac{- g_{\nu,\lambda}}{k^\rho \ k_\rho - (\frac{m \ c}{\hbar})^2} \right) j^\lambda(k_0, \underline{k}) \\ &= + \frac{4 \ \pi}{c^2} \int \frac{dk_0'}{(2 \ \pi)} \int \frac{d^4k}{(2 \ \pi)^4} \left(\frac{j^\nu(k_0', \underline{k})^* \ j_\nu(k_0, \underline{k})}{k^\rho \ k_\rho - (\frac{m \ c}{\hbar})^2} \right) \end{aligned} \qquad (461)$$

For two static charges, for which

$$\rho(k) = 2 \ \pi \ \delta(k_0) \ \rho(\underline{k}) \qquad (462)$$

one has

$$L_{int} = \frac{4\pi}{c^2} \int \frac{d^3\underline{k}}{(2\pi)^3} \left(\frac{\rho(\underline{k})^* \, \rho(\underline{k})}{-\underline{k}^2 - (\frac{m\,c}{\hbar})^2} \right) \tag{463}$$

Thus, the interaction Lagrangian between two like charges is negative and this corresponds to a positive (replusive) interaction potential V.

The massive photon corresponds to fluctuations of the four-vector potential A_ν, which has four components. If we assume the continuity of charge, then the four-vector potential must satisfy the gauge condition

$$k^\nu \, A_\nu(k) = 0 \tag{464}$$

Therefore, the massive photon has three independent components corresponding to the three different S^z eigenvalues of an $S = 1$ particle. Consider the form of the photon propagator

$$D_{\nu,\lambda}(k) = \left(\frac{-\,g_{\nu,\lambda} + \frac{k_\nu \, k_\lambda}{(\frac{m\,c}{\hbar})^2}}{k^\rho \, k_\rho - (\frac{m\,c}{\hbar})^2} \right) \tag{465}$$

The numerator originates from the product of polarization vectors $\epsilon^\alpha(k)$ for the massive spin-one particle. The polarization vectors can can be chosen as

$$\begin{aligned} \epsilon^{(1)} &= (0,1,0,0) \\ \epsilon^{(2)} &= (0,0,1,0) \\ \epsilon^{(3)} &= (0,0,0,1) \end{aligned} \tag{466}$$

and for a photon at rest, one has

$$k^\nu = ((\frac{m\,c}{\hbar}),0,0,0) \tag{467}$$

The polarization four-vectors and the energy-momentum four-vector satisfy the gauge condition

$$k^\nu \, \epsilon_\nu^\alpha(k) = 0 \tag{468}$$

which is Lorentz invariant. The propagator involves a polarization vector $\epsilon_\nu^\alpha(k)$ at the source where the photon is emitted. The photon then propagates at the sink, where it is absorbed. The absorption involves the factor $\epsilon_\lambda^\alpha(k)$. Since the polarization is not measured, it is summed over. Thus, the numerator of the propagator is given by factor

$$\sum_\alpha \epsilon_\nu^\alpha(k) \, \epsilon_\lambda^\alpha(k) \tag{469}$$

Since this quantity transforms as a Lorentz tensor, it can only be expressed in terms of the metric $g_{\nu,\lambda}$ and the product $k_\nu \, k_\lambda$. The combination can be fixed, up to a normalization, by noting that since one must use the gauge condition

$$k^\nu \, \epsilon_\nu^\alpha(k) = 0 \tag{470}$$

77

the propagator must satisfy the condition

$$k^\nu \, D_{\nu,\lambda}(k) \;=\; 0 \tag{471}$$

Thus, if the propagator has the form

$$D_{\nu,\lambda}(k) \;=\; \left(\frac{A \, g_{\nu,\lambda} \;+\; B \, k_\nu \, k_\lambda}{k^\rho \, k_\rho \;-\; (\frac{m \, c}{\hbar})^2} \right) \tag{472}$$

the gauge condition requires that

$$A \, k_\lambda \;+\; B \, k^\nu \, k_\nu \, k_\lambda \;=\; 0 \tag{473}$$

or

$$A \;+\; B \left(\frac{m \, c}{\hbar} \right)^2 \;=\; 0 \tag{474}$$

Therefore, the numerator of the propagator can be expressed as

$$A \left(\, g_{\nu,\lambda} \;-\; \frac{k_\nu \, k_\lambda}{(\frac{m \, c}{\hbar})^2} \, \right) \tag{475}$$

The constant of proportionality A can be fixed by considering the massive photon in its rest frame, and choosing one spatial index for ν and λ in the sum over polarization components (eg. $\nu = \lambda = 1$). This fixes $A = -1$. Thus, the numerator of the photon propagator is given by

$$G_{\nu,\lambda}(k) \;=\; \left(\, - \, g_{\nu,\lambda} \;+\; \frac{k_\nu \, k_\lambda}{(\frac{m \, c}{\hbar})^2} \, \right) \tag{476}$$

We have shown that the photon propagator is given by

$$D_{\nu,\lambda}(k) \;=\; \left(\frac{G_{\nu,\lambda}(k)}{k^\rho \, k_\rho \;-\; (\frac{m \, c}{\hbar})^2} \right) \tag{477}$$

9.1 Mathematical Structure of General Relativity

The fully non-linear Einstein field equation is given by

$$G_{\mu,\nu} \;+\; \Lambda \, g_{\mu,\nu} \;=\; \frac{8 \, \pi \, G}{c^3} \, T_{\mu,\nu} \tag{478}$$

where Λ is the cosmological constant, $G_{\mu,\nu}$ is the Einstein tensor and $T_{\mu,\nu}$ is the energy-momentum tensor. The Einstein tensor is given in terms of the Ricci tensor $R_{\mu,\nu}$ via

$$G_{\mu,\nu} \;=\; R_{\mu,\nu} \;-\; \frac{1}{2} \, g_{\mu,\nu} \, R \tag{479}$$

78

where the Ricci scalar R is defined as

$$R = g^{\mu,\nu} \, R_{\mu,\nu} \tag{480}$$

The Ricci tensor is obtained from the Riemann tensor via the contraction

$$R_{\mu,\nu} = R^{\lambda}_{\mu,\lambda,\nu} \tag{481}$$

and the Riemann tensor is given in terms of the Christoffel symbols via

$$R^{\rho}_{\sigma,\mu,\nu} = \partial_\mu \, \Gamma^{\rho}_{\nu,\sigma} - \partial_\nu \, \Gamma^{\rho}_{\mu,\sigma} + \Gamma^{\rho}_{\mu,\lambda} \, \Gamma^{\lambda}_{\nu,\sigma} - \Gamma^{\rho}_{\nu,\lambda} \, \Gamma^{\lambda}_{\mu,\sigma} \tag{482}$$

where, in turn, the Christoffel symbols $\Gamma^{\sigma}_{\rho,\mu}$ are given in terms of the metric tensors by

$$\Gamma^{\sigma}_{\rho,\mu} = \frac{1}{2} \, g^{\sigma,\rho} \left(\partial_\mu \, g_{\nu,\rho} + \partial_\nu \, g_{\rho,\mu} - \partial_\rho \, g_{\mu,\nu} \right) \tag{483}$$

In the absence of a source, the Lagrangian density of the source-free gravitational field is given by

$$\mathcal{L} = \sqrt{-g} \, R \tag{484}$$

where $g = \det g_{\mu,\nu}$ is negative.

9.2   Linearized Gravity

The graviton is supposed to correspond to a particle with spin $S = 2$. The field of the graviton corresponds to the fluctuating part of the metric tensor $h_{\mu,\nu}$. Therefore, $h_{\mu,\nu}$ is the linearized part of a symmetric second-rank tensor, where the lowest-order piece corresponds to the metric tensor $\eta_{\mu\nu}$ for flat Minkowski space. Thus

$$g_{\mu,\nu} = \eta_{\mu,\nu} + h_{\mu,\nu} \tag{485}$$

where

$$h_{\mu,\nu} \ll 1 \tag{486}$$

Since the metric tensor is symmetric, it has 10 independent components. However, under a general coordinate transformation

$$x_\mu \rightarrow x'_\mu = x_\mu - \Lambda_\mu \tag{487}$$

the full metric changes according to

$$g_{\mu,\nu} \rightarrow g'_{\mu,\nu} = g_{\sigma,\tau} \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\tau}{\partial x'^\nu} \tag{488}$$

so the linearized metric transforms as

$$h_{\mu,\nu} \rightarrow h'_{\mu,\nu} = h_{\mu,\nu} + \partial_\mu \, \Lambda_\nu + \partial_\nu \, \Lambda_\mu \tag{489}$$

Thus, both metrics describe the same physical gravitational field and so gravitational theory has a gauge invariance. The gravitational gauge transformation is very similar to the electromagnetic gauge transformation

$$A_\mu \; \to \; A'_\mu \; = \; A_\mu \; + \; \partial_\mu \, \Lambda \qquad (490)$$

We can use this similarity to motivate the expectation that a graviton has $S = 2$. On introducing a mass for the photon, we found that charge is not automatically conserved. Furthermore, we found that if one enforces charge conservation, then the electromagnetic field must satisfy the gauge condition

$$\partial^\mu \, A_\mu \; = \; 0 \qquad (491)$$

Therefore, the four-vector potential has only three independent components as expected for a massive spin-one particle. Likewise, on giving the graviton a mass, one expects that the massive graviton's field should satisfy four different Lorentz gauge conditions

$$\partial^\mu \, h_{\mu,\nu} \; = \; 0 \qquad (492)$$

If $h_{\mu,\nu}$ is to represent a particle with $S = 2$, we anticipate that it must also be traceless

$$\eta^{\mu,\nu} \, h_{\mu,\nu} \; = \; 0 \qquad (493)$$

That is, a massive $S = 2$ graviton's field must have 5 independent components which correspond to the five different eigenvalues of S^z. Therefore, since the field $h_{\mu,\nu}$ has 10 components, it must satisfy five conditions.

In the Lorentz gauge, the linearized Einstein equations reduce to the form

$$\partial_\rho \, \partial^\rho \, h_{\mu,\nu} \; - \; \frac{1}{2} \, \eta_{\mu,\nu} \, \partial_\rho \, \partial^\rho \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \; = \; - \left(\frac{16 \, \pi \, G}{c^3} \right) \, T_{\mu,\nu} \qquad (494)$$

where the energy-momentum tensor $T_{\mu,\nu}$ is assumed to be symmetric. On multiplying by $\eta^{\mu,\nu}$ and summing over the indices μ and ν, one finds

$$(\, 1 \; - \; \frac{d}{2} \,) \, \partial_\rho \, \partial^\rho \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \; = \; - \left(\frac{16 \, \pi \, G}{c^3} \right) \, \eta^{\mu,\nu} \, T_{\mu,\nu} \qquad (495)$$

where $d = 4$. This equation can be used to eliminate the second term in the linearized equation of motion, leading to

$$\partial_\rho \, \partial^\rho \, h_{\mu,\nu} \; = \; - \left(\frac{16 \, \pi \, G}{c^3} \right) \left(\, T_{\mu,\nu} \; - \; \frac{1}{2} \, \eta_{\mu,\nu} \, \eta^{\sigma,\tau} \, T_{\sigma,\tau} \, \right) \qquad (496)$$

The second term in the source projects out unwanted components of the energy-momentum tensor. For the vacuum, where $T_{\mu,\nu} = 0$, the linearized equations take the form

$$\partial_\rho \, \partial^\rho \, h_{\mu,\nu} \; = \; 0 \qquad (497)$$

which is just the relativistic wave equation for a massless spin-two particle. For a graviton of mass m, the wave equation could be expected to have the form

$$\left(\partial^\rho \, \partial_\rho \; + \; \left(\frac{m \, c}{\hbar} \right)^2 \right) h_{\mu,\nu} \; = \; 0 \tag{498}$$

9.3   The Massive $S = 2$ Graviton

The source-free equations of motion

$$\left(\partial^\rho \, \partial_\rho \; + \; \left(\frac{m \, c}{\hbar} \right)^2 \right) h_{\mu,\nu} \; = \; 0 \tag{499}$$

describing a massive $S = 2$ particle can be found by considering Fierz-Pauli Lagrangian density for a symmetric second-rank tensor field[21]. The Lagrangian density is given by

$$\begin{aligned}
\mathcal{L} \; = \; &+ \frac{1}{2} \, \partial_\lambda \, h_{\mu,\nu} \, \partial^\lambda \, h^{\mu,\nu} \; - \; \frac{1}{2} \, \partial_\lambda \, \eta_{\mu,\nu} \, h^{\mu,\nu} \, \partial^\lambda \, \eta_{\sigma,\tau} \, h^{\sigma,\tau} \\
&- \, \partial_\mu \, h_{\nu,\lambda} \, \partial^\nu \, h^{\mu,\lambda} \; + \; \partial_\mu \, h^{\mu,\nu} \, \partial_\nu \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \\
&- \frac{1}{2} \left(\frac{m \, c}{\hbar} \right)^2 \left[h_{\mu,\nu} \, h^{\mu,\nu} \; - \; \eta^{\mu,\nu} \, h_{\mu,\nu} \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \right]
\end{aligned} \tag{500}$$

up to a constant factor. The form of the mass term is not enforced by any known symmetry. In fact, like the Proca Lagrangian for the massive photon, the mass term in the Fierz-Pauli Lagrangian density violates the gauge symmetry. The principle of extremal action leads to the equations of motion

$$\begin{aligned}
0 \; = \; &- \, \partial_\lambda \, \partial^\lambda \, h^{\mu,\nu} \; + \; \eta^{\mu,\nu} \, \partial_\lambda \, \partial^\lambda \, \eta_{\sigma,\tau} \, h^{\sigma,\tau} \\
&+ \, \partial_\lambda \, \partial^\mu \, h^{\lambda,\nu} \; + \; \partial_\lambda \, \partial^\nu \, h^{\lambda,\mu} \\
&- \, \partial^\mu \, \partial^\nu \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \; - \; \eta^{\mu,\nu} \, \partial_\sigma \, \partial_\tau \, h^{\sigma,\tau} \\
&- \left(\frac{m \, c}{\hbar} \right)^2 \left[h^{\mu,\nu} \; - \; \eta^{\mu,\nu} \, \eta_{\sigma,\tau} \, h^{\sigma,\tau} \right]
\end{aligned} \tag{501}$$

On operating on the equation with ∂_μ and on noting that only the mass term remains, one finds the equation

$$\partial^\mu \, h_{\mu,\nu} \; - \; \partial_\nu \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \; = \; 0 \tag{502}$$

This is a mandatory "Lorentz gauge" condition for linearized (massive) general relativity. Here this is not a choice since the condition is enforced by the

---

[21]M. Fierz and W. Pauli, Proc. Roy. Soc. London, **187**, 211-232 (1937).

existence of the mass. Substituting this mandatory condition back into the equations of motion yields

$$\partial^\rho \, \partial_\rho \, h_{\mu,\nu} \, - \, \partial_\mu \, \partial_\nu \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \, + \, \left(\frac{m \, c}{\hbar} \right)^2 \left(h_{\mu,\nu} \, - \, \eta_{\mu,\nu} \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \right) \; = \; 0 \quad (503)$$

On taking the trace of the above equation, if $m \neq 0$, one recovers the condition that $h_{\mu,\nu}$ must be traceless

$$\eta^{\mu,\nu} \, h_{\mu,\nu} \; = \; 0 \tag{504}$$

as anticipated. Furthermore, this also implies the four mandatory "gauge conditions"

$$\partial^\mu \, h_{\mu,\nu} \; = \; 0 \tag{505}$$

must be satisfied. On applying the above five constraints to the equation of motion, one finds

$$\left(\partial^\rho \, \partial_\rho \, + \, \left(\frac{m \, c}{\hbar} \right)^2 \right) h_{\mu,\nu} \; = \; 0 \tag{506}$$

Thus, the Fierz-Pauli Lagrangian for $h_{\mu,\nu}$ does indeed describe a massive particle with $S = 2$, and we have confirmed the form of the five conditions on the field. In the presence of a source, the equation could be expected to have the form

$$\left(\partial^\rho \, \partial_\rho \, + \, \left(\frac{m \, c}{\hbar} \right)^2 \right) h_{\mu,\nu} \; = \; - \left(\frac{16 \, \pi \, G}{c^3} \right) \tilde{T}_{\mu,\nu} \tag{507}$$

where $\tilde{T}_{\mu,\nu}$ is an appropriate expression linear in the energy-momentum tensor which radiates a tensor field with five independent components. The unwanted components of $\tilde{T}_{\mu,\nu}$ must be projected out, similar to what we have done for the linearized theory of general relativity. The propagator takes care of this projection. The equation of motion is a linear equation with a source, so it can be solved by introducing a propagator which satisfies the equation

$$\left(\partial^\rho \, \partial_\rho \, + \, \left(\frac{m \, c}{\hbar} \right)^2 \right) D_{\mu,\nu;\sigma,\tau}(x, x') \; = \; \frac{1}{2} \left(\delta_\mu^\sigma \, \delta_\nu^\tau \, + \, \delta_\mu^\tau \, \delta_\nu^\sigma \right) \delta^4(x - x') \tag{508}$$

so that the solution of the inhomogeneous equation can be written as

$$h_{\mu,\nu}(x) \; = \; \left(\frac{16 \, \pi \, G}{c^3} \right) \int d^4x' \, D_{\mu,\nu;\sigma,\tau}(x, x') \, T_{\sigma,\tau}(x') \tag{509}$$

which involves a convolution of the propagator with the source.

The above analysis implies that the Fourier transformed propagator for a graviton with mass m must have the form

$$D_{\mu,\nu;\sigma,\tau}(k) \; = \; \left(\frac{B_{\mu,\nu;\sigma,\tau}(k)}{k^\rho \, k_\rho \, - \, (\frac{m \, c}{\hbar})^2} \right) \tag{510}$$

The graviton's field is normalized according to

$$\sum_\alpha \epsilon^\alpha_{\mu,\nu}(\underline{k}) \, \epsilon^\alpha_{\mu,\nu}(\underline{k}) \;=\; \hat{I} \tag{511}$$

The numerator of the graviton propagator must involve a combination which transforms as a tensor under a Lorentz transformation. The propagator's form must be such that the symmetric tensor $h_{\mu,\nu}$ must satisfy the four gauge conditions and be traceless. Using the five conditions repeatedly, one can show that the numerator is proportional to

$$\sum_\alpha \epsilon^\alpha_{\mu,\nu}(\underline{k}) \, \epsilon^\alpha_{\sigma,\tau}(\underline{k}) \;\propto\; \left(G_{\mu,\sigma}(k) \, G_{\nu,\tau}(k) + G_{\mu,\tau}(k) \, G_{\nu,\sigma}(k) \right) - \frac{2}{3} \, G_{\mu,\nu}(k) \, G_{\sigma,\tau}(k) \tag{512}$$

where

$$G_{\mu,\nu}(k) \;=\; \left(\; - \; \eta_{\mu,\nu} \; + \; \frac{k_\mu \, k_\nu}{(\frac{m\,c}{\hbar})^2} \; \right) \tag{513}$$

The sign of the propagator is fixed by considering a massive graviton in its rest frame and setting $\mu = \sigma = 1$ and $\nu = \tau = 2$. This leads to

$$D_{\mu,\nu;\sigma,\tau}(k) \;=\; \frac{1}{2} \left[\frac{ \left(G_{\mu,\sigma}(k) \, G_{\nu,\tau}(k) \; + \; G_{\mu,\tau}(k) \, G_{\nu,\sigma}(k) \right) \; - \; \frac{2}{3} \, G_{\mu,\nu}(k) \, G_{\sigma,\tau}(k) }{ k^\rho \, k_\rho \; - \; (\frac{m\,c}{\hbar})^2 } \right] \tag{514}$$

Einstein's equation links the Riemann curvature, or equivalently $h^{\mu,\nu}$, to the metric tensor $T^{\mu,\nu}$. Since energy and momentum are conserved, in flat space the energy-momentum tensor satisfies

$$\partial_\mu \, T^{\mu,\nu} \;=\; 0 \tag{515}$$

or, on Fourier transforming

$$k_\mu \, T^{\mu,\nu}(k) \;=\; 0 \tag{516}$$

The above condition eliminates the the k-dependent terms in the propagator when coupled to the energy-momentum tensor, so one can replace the $G_{\mu,\nu}(k)$ by $- \eta_{\mu,\nu}$. It should be noted that the energy density is positive

$$T^{0,0}(k) \;>\; 0 \tag{517}$$

and for time-like indices $\eta_{0,0} \;=\; 1$. The interaction Lagrangian corresponding to two static masses is given by

$$L_{int} \;=\; - \left(\frac{8 \, \pi \, G}{c^2} \right) \int \frac{d^3 \underline{k}}{(\,2\,\pi\,)^3} \, T^{0,0\,*}(0, \underline{k}) \, \frac{1 \, + \, 1 \, - \, \frac{2}{3}}{- \, \underline{k}^2 \, - \, (\frac{m\,c}{\hbar})^2 \, + \, i \, \eta} \, T^{0,0}(0, \underline{k}) \tag{518}$$

so for static mass densities, the interaction Lagrangian is positive. Thus, the gravitational interaction between two masses is attractive. However, after we

have introduced the factor of $-\frac{1}{2}$ to obtain the static interaction potential, we that the result differs from the result expected from Newton's law of gravity by a factor of $\frac{4}{3}$. That is, one expects that the factor of $\frac{2}{3}$ in the last term of the propagator should be replaced by unity. The propagator corresponding to the linearized theory of gravity (for which $m = 0$) does have the $\frac{2}{3}$ replaced by 1. Due to this discrepancy, we shall examine the connection between the propagator and the source in more detail below.

9.4   The Sourced Equations

The Fierz-Pauli Lagrangian in the presence of a source can be written as

$$
\begin{aligned}
\mathcal{L} \;=\; & + \frac{c^4}{32 \, \pi \, G} \left[\, \partial_\lambda \, h_{\mu,\nu} \, \partial^\lambda \, h^{\mu,\nu} \; - \; \partial_\lambda \, \eta_{\mu,\nu} \, h^{\mu,\nu} \, \partial^\lambda \, \eta_{\sigma,\tau} \, h^{\sigma,\tau} \right. \\
& - \, 2 \, \partial_\mu \, h_{\nu,\lambda} \, \partial^\nu \, h^{\mu,\lambda} \; + \; 2 \, \partial_\mu \, h^{\mu,\nu} \, \partial_\nu \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \\
& - \left. \left(\frac{m \, c}{\hbar} \right)^2 \left(\, h_{\mu,\nu} \, h^{\mu,\nu} \; - \; \eta^{\mu,\nu} \, h_{\mu,\nu} \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \right) \right] \\
& - \, c \, h_{\mu,\nu} \, T^{\mu,\nu} \hspace{5cm} (519)
\end{aligned}
$$

In the presence of a source, the Fierz-Pauli equation of motion becomes

$$
\begin{aligned}
- \left(\frac{16 \, \pi \, G}{c^3} \right) T^{\mu,\nu} \;=\; & \partial_\lambda \, \partial^\lambda \, h^{\mu,\nu} \; - \; \eta^{\mu,\nu} \, \partial_\lambda \, \partial^\lambda \, \eta_{\sigma,\tau} \, h^{\sigma,\tau} \\
& - \, \partial_\lambda \, \partial^\mu \, h^{\lambda,\nu} \; - \; \partial_\lambda \, \partial^\nu \, h^{\lambda,\mu} \\
& + \, \partial^\mu \, \partial^\nu \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \; + \; \eta^{\mu,\nu} \, \partial_\sigma \, \partial_\tau \, h^{\sigma,\tau} \\
& + \left(\frac{m \, c}{\hbar} \right)^2 \left[\, h^{\mu,\nu} \; - \; \eta^{\mu,\nu} \, \eta_{\sigma,\tau} \, h^{\sigma,\tau} \right] \quad (520)
\end{aligned}
$$

On operating on the equation with ∂_μ and on noting that only the mass term remains, one finds the equation

$$
\left(\frac{m \, c}{\hbar} \right)^2 \left[\, \partial^\mu \, h_{\mu,\nu} \; - \; \partial_\nu \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \right] \;=\; - \left(\frac{16 \, \pi \, G}{c^3} \right) \partial_\mu \, T^{\mu,\nu} \qquad (521)
$$

We see that in the absence of a graviton mass, the Lagrangian theory of the graviton requires that the energy-momentum tensor must be conserved. For the massive graviton, the conservation of energy and momentum is not automatically ensured, but is an independent assumption. We shall assume that the energy-momentum tensor remains a conserved quantity.

On assuming conservation of energy and momentum, one can substitute the condition

$$
\partial^\mu \, h_{\mu,\nu} \; - \; \partial_\nu \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \;=\; 0 \qquad (522)
$$

into the equation of motion, yielding

$$\partial^\rho \, \partial_\rho \, h_{\mu,\nu} - \partial_\mu \, \partial_\nu \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} + \left(\frac{m \, c}{\hbar} \right)^2 \left(h_{\mu,\nu} - \eta_{\mu,\nu} \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \right) = - \left(\frac{16 \, \pi \, G}{c^3} \right) T_{\mu,\nu}$$

(523)

Taking the trace, we find that

$$(\, d \, - \, 1 \,) \, \eta^{\sigma,\tau} \, h_{\sigma,\tau} \left(\frac{m \, c}{\hbar} \right)^2 = \left(\frac{16 \, \pi \, G}{c^3} \right) \eta^{\sigma,\tau} \, T_{\sigma,\tau}$$

(524)

so that $h_{\mu,\nu}$ is not traceless in the presence of a source. On substituting the trace in the condition, one finds that the four mandatory gauge conditions are modified to become

$$\partial^\mu \, h_{\mu,\nu} = \frac{1}{d \, - \, 1} \left(\frac{16 \, \pi \, \hbar^2 \, G}{m^2 \, c^5} \right) \partial_\nu \, \eta^{\sigma,\tau} \, T_{\sigma,\tau}$$

(525)

Thus, automatic conservation of energy and momentum and gauge invariance appear to be linked. That is, when energy and momentum conservation has to be inferred from other considerations, a specific gauge condition must be enforced. Substituting the expression for the trace of $h_{\mu,\nu}$ into the equations of motion leads to

$$\partial^\rho \, \partial_\rho \, h_{\mu,\nu} + \left(\frac{m \, c}{\hbar} \right)^2 h_{\mu,\nu} = - \left(\frac{16 \, \pi \, G}{c^3} \right) \left[T_{\mu,\nu} - \frac{1}{d \, - \, 1} \left(\eta_{\mu,\nu} + \frac{\partial_\mu \, \partial_\nu}{(\frac{m \, c}{\hbar})^2} \right) \eta^{\sigma,\tau} \, T_{\sigma,\tau} \right]$$

(526)

On Fourier transforming, one obtains

$$h_{\mu,\nu}(k) = \frac{\left(\frac{16 \, \pi \, G}{c^3} \right)}{k^\rho \, k_\rho \, - \, (\frac{m \, c}{\hbar})^2} \left[T_{\mu,\nu}(k) - \frac{1}{d \, - \, 1} \left(\eta_{\mu,\nu} - \frac{k_\mu \, k_\nu}{(\frac{m \, c}{\hbar})^2} \right) \eta^{\sigma,\tau} \, T_{\sigma,\tau}(k) \right]$$

(527)

or

$$h_{\mu,\nu}(k) = \frac{\left(\frac{16 \, \pi \, G}{c^3} \right)}{k^\rho \, k_\rho \, - \, (\frac{m \, c}{\hbar})^2} \left[\eta_{\mu,\sigma} \, \eta_{\nu,\tau} - \frac{1}{d \, - \, 1} \left(\eta_{\mu,\nu} - \frac{k_\mu \, k_\nu}{(\frac{m \, c}{\hbar})^2} \right) \eta_{\sigma,\tau} \right] T^{\sigma,\tau}(k)$$

(528)

As we are assuming that energy and momentum are conserved

$$k^\mu \, T_{\mu,\nu}(k) = 0$$

(529)

the above equation can be re-written as

$$h_{\mu,\nu}(k) = \frac{\left(\frac{16 \, \pi \, G}{c^3} \right)}{k^\rho \, k_\rho \, - \, (\frac{m \, c}{\hbar})^2} \left[G_{\mu,\sigma}(k) \, G_{\nu,\tau}(k) - \frac{1}{d \, - \, 1} \, G_{\mu,\nu}(k) \, G_{\sigma,\tau}(k) \right] T^{\sigma,\tau}(k)$$

(530)

85

Since $h_{\mu,\nu}$ is symmetric in the indices, one should also symmetrize the propagator. The symmetrized propagator is given by

$$h_{\mu,\nu}(k) = \frac{\left(\frac{8\,\pi\,G}{c^3}\right)}{k^\rho\,k_\rho - \left(\frac{m\,c}{\hbar}\right)^2} \left[G_{\mu,\sigma}(k)\,G_{\nu,\tau}(k) + G_{\mu,\tau}(k)\,G_{\nu,\sigma}(k) - \frac{2}{d-1}\,G_{\mu,\nu}(k)\,G_{\sigma,\tau}(k) \right] T^{\sigma,\tau}(k)$$

$$(531)$$

in agreement with the result of our previous calculation. The factor of $\frac{2}{d-1}$ is a real discrepancy with the $m = 0$ result which contains $\frac{2}{d-2}$. This discrepancy is due to the fact that in the limit $m \to 0$, the 5 excitations of the Fierz-Pauli theory decouple into helicity ± 2 excitations, helicity ± 1 vector excitations and a scalar particle. The vector particles do not mediate an interaction, but the scalar particle does. The scalar particle is the so-called longitudinal graviton. However, the theory with m exactly equal to zero only has two helicity ± 2 excitations and does not include the longitudinal graviton. The discontinuity between the $m = 0$ theory and the limit $m \to 0$ was investigated by Vainshtein[22].

9.5   The Modes and Energies of Massive Gravitons

The Fierz-Pauli Lagrangian density

$$\begin{aligned}
\mathcal{L} \;=\; & + \frac{c^4}{32\,\pi\,G}\left[\, \partial_\lambda\, h_{\mu,\nu}\,\partial^\lambda\,h^{\mu,\nu} \;-\; \partial_\lambda\,\eta_{\mu,\nu}\,h^{\mu,\nu}\,\partial^\lambda\,\eta_{\sigma,\tau}\,h^{\sigma,\tau}\right. \\
& - 2\,\partial_\mu\,h_{\nu,\lambda}\,\partial^\nu\,h^{\mu,\lambda} \;+\; 2\,\partial_\mu\,h^{\mu,\nu}\,\partial_\nu\,\eta^{\sigma,\tau}\,h_{\sigma,\tau} \\
& \left. - \left(\frac{m\,c}{\hbar}\right)^2\left(\, h_{\mu,\nu}\,h^{\mu,\nu} \;-\; \eta^{\mu,\nu}\,h_{\mu,\nu}\,\eta^{\sigma,\tau}\,h_{\sigma,\tau}\right)\right]
\end{aligned}$$

$$(532)$$

can be re-written by integrating the third term by parts twice, leading to

$$\begin{aligned}
\mathcal{L} \;=\; & + \frac{c^4}{32\,\pi\,G}\left[\, \partial_\lambda\, h_{\mu,\nu}\,\partial^\lambda\,h^{\mu,\nu} \;-\; \partial_\lambda\,\eta_{\mu,\nu}\,h^{\mu,\nu}\,\partial^\lambda\,\eta_{\sigma,\tau}\,h^{\sigma,\tau}\right. \\
& - 2\,\partial^\nu\,h_{\nu,\lambda}\,\partial_\mu\,h^{\mu,\lambda} \;+\; 2\,\partial_\mu\,h^{\mu,\nu}\,\partial_\nu\,\eta^{\sigma,\tau}\,h_{\sigma,\tau} \\
& \left. - \left(\frac{m\,c}{\hbar}\right)^2\left(\, h_{\mu,\nu}\,h^{\mu,\nu} \;-\; \eta^{\mu,\nu}\,h_{\mu,\nu}\,\eta^{\sigma,\tau}\,h_{\sigma,\tau}\right)\right]
\end{aligned}$$

$$(533)$$

The vanishing of the trace and the mandatory gauge conditions obtained from the equation of motion can be used to reduce the Lagrangian density to an effective Lagrangian density

$$\mathcal{L}_{eff} \;=\; + \frac{c^4}{32\,\pi\,G}\left(\, \partial_\lambda\, h_{\mu,\nu}\,\partial^\lambda\,h^{\mu,\nu} \;-\; \left(\frac{m\,c}{\hbar}\right)^2 h_{\mu,\nu}\,h^{\mu,\nu}\right)$$

$$(534)$$

[22]A.I. Vainshtein, Phys. Lett. B, **39**, 393-394 (1972).

Hence, the energy of the field can be expressed as

$$P_0 = \frac{c^3}{32 \pi G} \int d^3\underline{r} \left(\partial_0 h_{\mu,\nu} \partial_0 h^{\mu,\nu} + \underline{\nabla} h_{\mu,\nu} \cdot \underline{\nabla} h^{\mu,\nu} + \left(\frac{m c}{\hbar} \right)^2 h_{\mu,\nu} h^{\mu,\nu} \right) \tag{535}$$

or, on Fourier transforming,

$$P_0 = \frac{c^3 \pi^2}{4 G} \int d^3\underline{k} \left(k_0^2 + \underline{k}^2 + \left(\frac{m c}{\hbar} \right)^2 \right) h_{\mu,\nu}(k) h^{\mu,\nu}(k) \tag{536}$$

so the energy is positive.

The fields are restricted by the four constraints

$$k^\mu h_{\mu,\nu}(k) = 0 \tag{537}$$

and the trace condition

$$\eta_{\mu,\nu} h^{\mu,\nu}(k) = 0 \tag{538}$$

The four constraints lead to the elimination of the time-like components of the tensor as independent variables since

$$h_{0,i}(k) = - \sum_j \frac{k^j}{k^0} h_{i,j}(k) \tag{539}$$

and

$$h_{0,0}(k) = \sum_{i,j} \frac{k^i}{k^0} \frac{k^j}{k^0} h_{i,j}(k) \tag{540}$$

The trace condition is

$$h_{0,0}(k) - h_{1,1}(k) - h_{2,2}(k) - h_{3,3}(k) = 0 \tag{541}$$

Therefore, we have three independent off-diagonal amplitudes $h_{1,2}$, $h_{1,3}$, and $h_{2,3}$. The independent diagonal amplitudes can be chosen as $h_{3,3}$ and the difference $h_{1,1} - h_{2,2}$, since the sum $h_{1,1} + h_{2,2}$ is determined in terms of $h_{0,0} - h_{3,3}$ by the trace condition. The energy-density was found to be proportional to

$$h_{\mu,\nu}(k) h^{\mu,\nu}(k) \tag{542}$$

which can be re-written in terms of the independent amplitudes as

$$h_{\mu,\nu}(k) h^{\mu,\nu}(k) = \left| \sum_{i,j} \frac{k^i}{k^0} \frac{k^j}{k^0} h_{i,j}(k) \right|^2 - 2 \sum_j \left| \sum_i \frac{k^i}{k^0} h_{i,j}(k) \right|^2 + \sum_{i,j} |h_{i,j}(k)|^2 \tag{543}$$

87

For $\underline{k}$ directed along the z-direction, one has

$$
\begin{aligned}
h_{\mu,\nu}(k)\, h^{\mu,\nu}(k) \;=\;& \left(\frac{k^2}{k_0^2}-1\right)^2 |h_{3,3}(k)|^2 \;-\; 2\left(\frac{k^2}{k_0^2}-1\right)\left(|h_{3,1}(k)|^2 \;+\; |h_{3,2}(k)|^2\right) \\
& + 2\,|h_{1,2}(k)|^2 \;+\; |h_{1,1}(k)|^2 \;+\; |h_{2,2}(k)|^2 \\
\;=\;& \frac{3}{2}\left(\frac{k^2}{k_0^2}-1\right)^2 |h_{3,3}(k)|^2 \;-\; 2\left(\frac{k^2}{k_0^2}-1\right)\left(|h_{3,1}(k)|^2 \;+\; |h_{3,2}(k)|^2\right) \\
& + 2\,|h_{1,2}(k)|^2 \;+\; \frac{1}{2}\,|\,h_{1,1}(k)\,-\,h_{2,2}(k)\,|^2 \tag{544}
\end{aligned}
$$

Thus, the massive graviton field has five independent components.

If one assumes that

$$
k_0^2 \;=\; \underline{k}^2 \;+\; \left(\frac{m\,c}{\hbar}\right)^2 \tag{545}
$$

then one sees that the two off-diagonal modes $h_{3,1}(k)$ and $h_{3,2}(k)$ have an energy proportional to

$$
4\left(\frac{m\,c}{\hbar}\right)^2\left(|h_{3,1}(k)|^2 \;+\; |h_{3,2}(k)|^2\right) \tag{546}
$$

It can be shown that the amplitudes do not diverge in the limit $m \to 0$, so these modes do not carry energy in this limit. The two modes have helicities of $S^z = \pm 1$. This can be seen by noting that the linear combinations

$$
h_{3,1}(k) \;\pm\; i\, h_{3,2}(k) \tag{547}
$$

transform as

$$
h_{3,1}(k)' \;\pm\; i\, h_{3,2}(k)' \;=\; \exp[\,\mp\, i\, \varphi\,]\left(h_{3,1}(k) \;\pm\; i\, h_{3,2}(k)\right) \tag{548}
$$

under a rotation of φ around the z-axis. Thus, these combinations have angular momenta of ± 1 around the z-axis. The physical effects of the modes $h_{3,1}(k)$ and $h_{3,2}(k)$ vanish when $m \to 0$.

The two modes $h_{1,1}(k) - h_{2,2}(k)$ and $h_{1,2}(k)$ carry energy in the limit of zero mass, have the helicity $S^z = \pm 2$ and correspond to the graviton. That is, the linear combination

$$
\left(\frac{h_{1,1}(k)-h_{22}(k)}{2}\right) \;\pm\; i\, h_{1,2}(k) \tag{549}
$$

transform as

$$
\exp[\,\mp\, i\, 2\, \varphi\,]\left[\left(\frac{h_{1,1}(k)-h_{22}(k)}{2}\right) \;\pm\; i\, h_{1,2}(k)\right] \tag{550}
$$

under a rotation of φ around the z-axis and, thus, correspond to a helicity $S^z = \pm 2$.

The remaining mode $h_{3,3}(k)$ is the longitudinal graviton which corresponds to $S^z = 0$. The amplitude of the longitudinal graviton is expected to diverge in the limit $m \to 0$. This can be seen by examining the equation for the source of the gravitational radiation

$$\left(k^\rho \, k_\rho - (\frac{m \, c}{\hbar})^2 \right) h_{\mu,\nu}(k) \; = \; \left(\frac{16 \, \pi \, G}{c^3} \right) \left[\eta_{\mu,\sigma} \, \eta_{\nu,\tau} - \frac{1}{d \, - \, 1} \left(\eta_{\mu,\nu} - \frac{k_\mu \, k_\nu}{(\frac{m \, c}{\hbar})^2} \right) \eta_{\sigma,\tau} \right] T^{\sigma,\tau}(k)$$
$$(551)$$

and noting that, for $\underline{k}$ along the z-direction, the term in the source which diverges when $m \to 0$ only couples to $h_{3,3}(k)$. Thus, we have

$$h_{3,3}(k) \; \sim \; \frac{1}{(\frac{m \, c}{\hbar})^2} \tag{552}$$

and, therefore, this mode yields a finite contribution to the energy. It is this longitudinal mode which gives rise to the discontinuity at $m = 0$.

10   Quantization of the Electromagnetic Field

Following the work of Dirac[23], the energy, momentum and angular momentum of the electromagnetic field shall be reduced into contributions from a set of normal modes. A particular normal mode will correspond to a particular wave vector and a particular polarization of the field. The normal modes can be described in terms of a set of harmonic oscillators and, when quantized, the normal modes will be described by quantum mechanical harmonic oscillators.

In the absence of sources, the (classical) wave equation for the vector potential has the form

$$\left[\; - \; \underline{\nabla}^2 \; + \; \frac{1}{c^2} \, \frac{\partial^2}{\partial t^2} \; \right] \underline{A} \; = \; 0 \tag{553}$$

when the Coulomb gauge condition is imposed

$$\underline{\nabla} \, . \, \underline{A} \; = \; 0 \tag{554}$$

The Fourier transformation, with respect to space is defined as

$$\underline{A}(\underline{k}, t) \; = \; \frac{1}{\sqrt{V}} \; \int d^3\underline{r} \; \exp \left[\; i \, \underline{k} \, . \, \underline{r} \; \right] \underline{A}(\underline{r}, t) \tag{555}$$

---

[23]P. A. M. Dirac, Proc. Roy. Soc. **A 114**, 243 (1927).
In this paper Dirac uses two different approaches to quantizing electromagnetism. In one approach he treated a single photon as satisfying a single-particle Schrödinger equation, that has a similar form to Maxwell's equations. The other approach treated the fields as dynamical variables and then quantized them. Dirac then showed that these two methods produce equivalent results. By doing this, Dirac created second quantization.

where V is the volume of the system. The inverse Fourier Transform is given by

$$\underline{A}(\underline{r},t) \;=\; \frac{1}{\sqrt{V}} \sum_{\underline{k}} \exp\Big[-\, i\, \underline{k}\,.\,\underline{r} \Big] \, \underline{A}(\underline{k},t) \tag{556}$$

On Fourier transforming the wave equation with respect to space and time, one finds the equation of motion

$$\left[\underline{k}^2 \;+\; \frac{1}{c^2} \left(\frac{\partial^2}{\partial t^2} \right) \right] A(\underline{k},t) \;=\; 0 \tag{557}$$

and the Coulomb gauge condition becomes

$$\underline{k}\,.\,\underline{A}(\underline{k},t) \;=\; 0 \tag{558}$$

We shall look for solutions for $\underline{A}(\underline{k},t)$ that have a time dependence given by linear superpositions of the terms proportional to

$$\exp\Big[\mp\, i\, \omega_k\, t \Big] \tag{559}$$

By substituting the above terms into the wave equation, it is found that linear superpositions of plane-waves are solutions of Maxwell's equation but only if the frequency ω_k and wave vector k are related via the dispersion relation

$$\omega_k^2 \;=\; c^2\, k^2 \tag{560}$$

The gauge condition also requires that the vector potential is oriented perpendicular to the direction of propagation. Therefore, an arbitrary plane-wave solution can be represented as a linear superposition of two polarized waves with polarizations described by two mutually orthogonal unit vectors denoted by $\hat{\epsilon}_\alpha(\underline{k})$. The polarization vectors satisfy

$$\begin{aligned} \underline{k}\,.\,\hat{\epsilon}_\alpha(\underline{k}) \;&=\; 0 \\ \hat{\epsilon}_\alpha(\underline{k})\,.\,\hat{\epsilon}_\beta(\underline{k}) \;&=\; \delta_{\alpha,\beta} \end{aligned} \tag{561}$$

We shall assume that three vectors $\Big(\underline{k}, \hat{\epsilon}_1(\underline{k}), \hat{\epsilon}_2(\underline{k}) \Big)$ form a mutually orthogonal coordinate system. We shall define

$$\begin{aligned} \hat{\epsilon}_1(-\underline{k}) \;&=\; \hat{\epsilon}_1(\underline{k}) \\ \hat{\epsilon}_2(-\underline{k}) \;&=\; \hat{\epsilon}_2(\underline{k}) \end{aligned} \tag{562}$$

The algebraic equations for $\underline{A}(\underline{k})$ can be solved trivially. One can express the vector potential as a linear superposition

$$\underline{A}(\underline{r},t) \;=\; \frac{1}{\sqrt{V}} \sum_{\underline{k},\alpha} \hat{\epsilon}_\alpha(\underline{k})\, \exp\Big[-\, i\, \underline{k}\,.\,\underline{r} \Big] \, \Phi_\alpha(\underline{k},t) \tag{563}$$

90
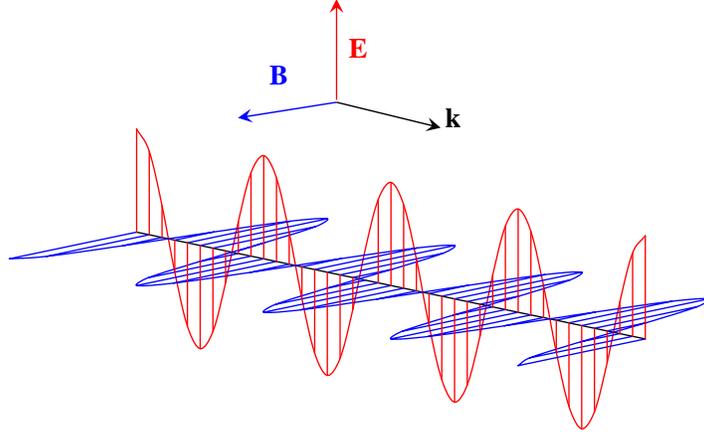
Figure 9: The normal modes of the classical electromagnetic field are plane-polarized waves, in which $\underline{E}$ and $\underline{B}$ are transverse to the direction of propagation $\underline{k}$, and oscillate in phase.

However, since the vector potential is real

$$\underline{A}(\underline{r}, t) = \underline{A}^*(\underline{r}, t) \tag{564}$$

one must have

$$\Phi_\alpha(\underline{k}, t) = \Phi_\alpha^*(-\underline{k}, t) \tag{565}$$

Therefore, if $\Phi_\alpha(\underline{k})$ and $\Phi_\alpha^*(\underline{k})$ are to be considered as being independent fields, then one must restrict $\underline{k}$ to have values in a volume of $\underline{k}$-space that does not contain both $\underline{k}$ and $-\underline{k}$ for any fixed value of $\underline{k}$. This curiosity is associated with the fact that, for purely real fields, particles are identical to their anti-particles.

10.1   The Lagrangian and Hamiltonian Density

The Lagrangian density $\mathcal{L}$ for the electromagnetic field can be expressed as

$$\mathcal{L} = \frac{1}{8\,\pi} \left[\underline{E}^2 - \underline{B}^2 \right] \tag{566}$$

in the Coulomb gauge, the electromagnetic field is given by

$$\begin{aligned} \underline{E} &= -\frac{1}{c}\,\frac{\partial \underline{A}}{\partial t} \\ \underline{B} &= \underline{\nabla} \wedge \underline{A} \end{aligned} \tag{567}$$

Hence, the Lagrangian density is expressed as

$$\mathcal{L} = \frac{1}{8\,\pi} \left[\frac{1}{c^2} \left(\frac{\partial \underline{A}}{\partial t} \right)^2 - \left(\underline{\nabla} \wedge \underline{A} \right)^2 \right] \tag{568}$$

91

The Lagrangian is given by the space integral of the Lagrangian density

$$L = \int d^3\underline{r} \, \mathcal{L} \tag{569}$$

On substituting $\underline{A}(\underline{r}, t)$ in the form of eqn(563) and integrating over $\underline{r}$ and using the identity

$$\frac{1}{V} \int d^3\underline{r} \, \exp\left[i \, (\, \underline{k} \, + \, \underline{k}' \,) \, . \, \underline{r} \, \right] = \delta_{\underline{k}+\underline{k}'} \tag{570}$$

one finds the Lagrangian is given by

$$
\begin{aligned}
L &= \frac{1}{8\,\pi} \sum_{\underline{k},\underline{k}'} \sum_{\alpha,\beta} \delta_{\underline{k}+\underline{k}'} \\
&\quad \times \left[\, \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\epsilon}_\beta(\underline{k}') \, \frac{1}{c^2} \left(\frac{\partial \Phi_\alpha(\underline{k})}{\partial t} \right) \left(\frac{\partial \Phi_\beta(\underline{k}')}{\partial t} \right) \right. \\
&\quad \left. + \, (\, \underline{k} \, \wedge \, \epsilon_\alpha(\underline{k}) \,) \, . \, (\, \underline{k}' \, \wedge \, \epsilon_\beta(\underline{k}') \,) \, \Phi_\alpha(\underline{k}) \, \Phi_\beta(\underline{k}') \, \right] \\
&= \frac{1}{8\,\pi} \sum_{\underline{k},\alpha} \left[\frac{1}{c^2} \left(\frac{\partial \Phi_\alpha^*(\underline{k})}{\partial t} \right) \left(\frac{\partial \Phi_\alpha(\underline{k})}{\partial t} \right) - \underline{k}^2 \, \Phi_\alpha^*(\underline{k}) \, \Phi_\alpha(\underline{k}) \right]
\end{aligned}
\tag{571}
$$

In the above expression, the summation over $\underline{k}$ is unrestricted. If the Lagrangian is to be expressed in terms of the independent components, then the summation over $\underline{k}$ must be restricted to half the set of allowed values. With this restriction,
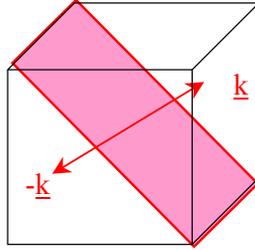
Figure 10: A possible partition of $\underline{k}$-space, which does not contain both $\underline{k}$ and its inverse $-\underline{k}$.

one obtains

$$L = \frac{2}{8\,\pi} \sideset{}{'}\sum_{\underline{k},\alpha} \left[\frac{1}{c^2} \left(\frac{\partial \Phi_\alpha^*(\underline{k})}{\partial t} \right) \left(\frac{\partial \Phi_\alpha(\underline{k})}{\partial t} \right) - \underline{k}^2 \, \Phi_\alpha^*(\underline{k}) \, \Phi_\alpha(\underline{k}) \right] \tag{572}$$

where the prime over the summation denotes the restriction of $\underline{k}$ to values in the "positive" half volume of $\underline{k}$-space. Since there are half the number of independent normal modes, their contributions are twice as big. The Lagrangian is a function of the six generalized variables $\Phi_\alpha(\underline{k})$ and $\Phi_\alpha^*(\underline{k})$ for the independent $\underline{k}$ values. The generalized momenta variables are found as

$$
\begin{aligned}
\Pi_\alpha(\underline{k}) &= \frac{2}{8\,\pi\,c^2}\left(\frac{\partial \Phi_\alpha^*(\underline{k})}{\partial t}\right) \\
\Pi_\alpha^*(\underline{k}) &= \frac{2}{8\,\pi\,c^2}\left(\frac{\partial \Phi_\alpha(\underline{k})}{\partial t}\right)
\end{aligned}
\tag{573}
$$

The Lagrangian equations of motion of the field are given by

$$
\frac{\partial}{\partial t}\left[\frac{1}{8\,\pi\,c^2}\left(\frac{\partial \Phi_\alpha(\underline{k})}{\partial t}\right)\right] = -\frac{k^2}{8\,\pi}\,\Phi_\alpha(\underline{k})
\tag{574}
$$

or

$$
\left(\frac{\partial^2 \Phi_\alpha(\underline{k})}{\partial t^2}\right) = -\,\omega_k^2\,\Phi_\alpha(\underline{k})
\tag{575}
$$

where $\omega_k = c\,k$. Thus, the classical field $\Phi_\alpha(\underline{k})$ has a time-dependent amplitude which resembles that of a harmonic oscillator with frequency $\omega_k = c\,k$. The Hamiltonian can be obtained from the Lagrangian, via the Legendre Transformation

$$
H = {\sum_{\underline{k},\alpha}}'\left[\Pi_\alpha^*(\underline{k})\,\frac{\partial \Phi_\alpha^*(\underline{k})}{\partial t} + \Pi_\alpha(\underline{k})\,\frac{\partial \Phi_\alpha(\underline{k})}{\partial t}\right] - L
\tag{576}
$$

which leads to the explicit expression for the Hamiltonian

$$
H = {\sum_{\underline{k},\alpha}}'\left[\frac{8\,\pi\,c^2}{2}\,\Pi_\alpha^*(\underline{k})\,\Pi_\alpha(\underline{k}) + \frac{2}{8\,\pi}\,k^2\,\Phi_\alpha^*(\underline{k})\,\Phi_\alpha(\underline{k})\right]
\tag{577}
$$

where the summation over $(\underline{k},\alpha)$ runs over the independent normal modes. Hence, the $\underline{k}$ summation only runs over the set of points in $\underline{k}$ space which are not related via the inversion operator. The Hamiltonian is related to the energy of the electromagnetic field, as shall be seen below.

The energy density $\mathcal{H}$ for the electromagnetic field can be expressed as

$$
\mathcal{H} = \frac{1}{8\,\pi}\left[\underline{E}^2 + \underline{B}^2\right]
\tag{578}
$$

in the Coulomb gauge. The energy density can be written in terms of the vector potential as

$$
\mathcal{H} = \frac{1}{8\,\pi}\left[\frac{1}{c^2}\left(\frac{\partial \underline{A}}{\partial t}\right)^2 + \left(\underline{\nabla} \wedge \underline{A}\right)^2\right]
\tag{579}
$$

93

The energy is the integral of the energy density over all space

$$H = \int d^3\underline{r}\, \mathcal{H} \tag{580}$$

When expressed in terms of the generalized coordinates and the generalized momenta, the energy reduces to the expression

$$H = \sum_{\underline{k},\alpha} \left[\frac{8\,\pi\,c^2}{4} \Pi_\alpha(\underline{k})\, \Pi_\alpha^*(\underline{k}) + \frac{1}{8\,\pi}\, \underline{k}^2\, \Phi_\alpha^*(\underline{k})\, \Phi_\alpha(\underline{k}) \right] \tag{581}$$

in which the summation over $\underline{k}$ is unrestricted. Thus, the above expression for the energy is identical to the Hamiltonian for the electromagnetic field. Furthermore, the Hamiltonian has been expressed in terms of a set of the normal modes labeled by $(\underline{k}, \alpha)$.

10.2 Quantizing the Normal Modes

The quantized Hamiltonian is obtained from the classical Hamiltonian by replacing the field components and their canonically conjugate momenta

$$\Phi_\alpha(\underline{k})\ ,\ \Pi_\alpha(\underline{k}) \tag{582}$$

by the operators

$$\hat{\Phi}_\alpha(\underline{k})\ ,\ \hat{\Pi}_\alpha(\underline{k}) \tag{583}$$

and their complex conjugates are replaced by the Hermitean conjugate operators. The canonically conjugate coordinates and momenta operators satisfy the commutation relations

$$
\begin{aligned}
\left[\, \hat{\Phi}_\alpha(\underline{k})\ ,\ \hat{\Pi}_\beta(\underline{k}')\,\right] &= i\,\hbar\,\delta_{\alpha,\beta}\,\delta_{\underline{k},\underline{k}'} \\
\left[\, \hat{\Pi}_\alpha(\underline{k})\ ,\ \hat{\Pi}_\beta(\underline{k}')\,\right] &= 0 \\
\left[\, \hat{\Phi}_\alpha(\underline{k})\ ,\ \hat{\Phi}_\beta(\underline{k}')\,\right] &= 0
\end{aligned}
\tag{584}
$$

The quantized Hamiltonian for the electromagnetic field is given by

$$\hat{H} = \sum_{\underline{k},\alpha} \left[\frac{8\,\pi\,c^2}{4} \hat{\Pi}_\alpha(\underline{k})\, \hat{\Pi}_\alpha^\dagger(\underline{k}) + \frac{1}{8\,\pi}\, \underline{k}^2\, \hat{\Phi}_\alpha^\dagger(\underline{k})\, \hat{\Phi}_\alpha(\underline{k}) \right] \tag{585}$$

The Hamiltonian can be factorized by introducing the annihilation operators

$$\hat{a}_{\underline{k},\alpha} = \frac{1}{\sqrt{2}} \left[i\,\sqrt{\frac{8\,\pi\,c^2}{2\,\hbar\,\omega_k}}\, \hat{\Pi}_\alpha(\underline{k}) + \sqrt{\frac{2\,k^2}{8\,\pi\,\hbar\,\omega_k}}\, \hat{\Phi}_\alpha^\dagger(\underline{k}) \right] \tag{586}$$

and the Hermitean conjugate operators

$$\hat{a}_{\underline{k},\alpha}^\dagger = \frac{1}{\sqrt{2}} \left[-i\,\sqrt{\frac{8\,\pi\,c^2}{2\,\hbar\,\omega_k}}\, \hat{\Pi}_\alpha^\dagger(\underline{k}) + \sqrt{\frac{2\,k^2}{8\,\pi\,\hbar\,\omega_k}}\, \hat{\Phi}_\alpha(\underline{k}) \right] \tag{587}$$

known as creation operators. The commutation relations for the creation and annihilation operators can be obtained directly from the commutation relations of the field operators $\hat{\Phi}_\alpha(\underline{k})$ and $\hat{\Pi}_\alpha(\underline{k})$ which are shown in eqn(584). It can be shown that the creation and annihilation operators satisfy the commutation relations

$$
\begin{array}{rcl}
[\ \hat{a}_{\underline{k},\alpha}\ ,\ \hat{a}^\dagger_{\underline{k}',\beta}\ ] & = & \delta_{\alpha,\beta}\ \delta_{\underline{k},\underline{k}'} \\[2mm]
[\ \hat{a}^\dagger_{\underline{k},\alpha}\ ,\ \hat{a}^\dagger_{\underline{k}',\beta}\ ] & = & 0 \\[2mm]
[\ \hat{a}_{\underline{k},\alpha}\ ,\ \hat{a}_{\underline{k}',\beta}\ ] & = & 0
\end{array}
\tag{588}
$$

The field operators can be expressed in terms of the creation and annihilation operators. Starting with

$$
\hat{a}_{\underline{k},\alpha} \;=\; \frac{1}{\sqrt{2}}\ \left[\ i\ \sqrt{\frac{8\ \pi\ c^2}{2\ \hbar\ \omega_k}}\ \hat{\Pi}_\alpha(\underline{k})\ +\ \sqrt{\frac{2\ k^2}{8\ \pi\ \hbar\ \omega_k}}\ \hat{\Phi}^\dagger_\alpha(\underline{k})\ \right]
\tag{589}
$$

transforming $\underline{k}\ \rightarrow\ -\underline{k}$ and then by noting that $\hat{\Pi}_\alpha(-\underline{k}) = \hat{\Pi}^\dagger_\alpha(\underline{k})$ and $\hat{\Phi}^\dagger_\alpha(-\underline{k}) = \hat{\Phi}_\alpha(\underline{k})$, one finds

$$
\hat{a}_{-\underline{k},\alpha} \;=\; \frac{1}{\sqrt{2}}\ \left[\ i\ \sqrt{\frac{8\ \pi\ c^2}{2\ \hbar\ \omega_k}}\ \hat{\Pi}^\dagger_\alpha(\underline{k})\ +\ \sqrt{\frac{2\ k^2}{8\ \pi\ \hbar\ \omega_k}}\ \hat{\Phi}_\alpha(\underline{k})\ \right]
\tag{590}
$$

One can eliminate $\hat{\Pi}^\dagger_\alpha(\underline{k})$ by adding the expression for the creation operator given by eqn(587) and the expression for the annihilation operator with momentum $-\underline{k}$ given by eqn(590). This process yields the expression for the field component operators $\hat{\Phi}_\alpha(\underline{k})$ in the form

$$
\hat{\Phi}_\alpha(\underline{k}) \;=\; \sqrt{\frac{2\ \pi\ \hbar\ \omega_k}{k^2}}\ \left(\ \hat{a}^\dagger_{\underline{k},\alpha}\ +\ \hat{a}_{-\underline{k},\alpha}\ \right)
\tag{591}
$$

and, by an analogous procedure, the Hermitean conjugate operator is found to be given by

$$
\hat{\Phi}^\dagger_\alpha(\underline{k}) \;=\; \sqrt{\frac{2\ \pi\ \hbar\ \omega_k}{k^2}}\ \left(\ \hat{a}_{\underline{k},\alpha}\ +\ \hat{a}^\dagger_{-\underline{k},\alpha}\ \right)
\tag{592}
$$

which is identical to $\hat{\Phi}_\alpha(-\underline{k})$. Likewise, the canonically conjugate momenta operators are given by

$$
\hat{\Pi}_\alpha(\underline{k}) \;=\; i\ \sqrt{\frac{\hbar\ \omega_k}{8\ \pi\ c^2}}\ \left(\ \hat{a}^\dagger_{-\underline{k},\alpha}\ -\ \hat{a}_{\underline{k},\alpha}\ \right)
\tag{593}
$$

and their Hermitean conjugates are

$$
\hat{\Pi}^\dagger_\alpha(\underline{k}) \;=\; -\ i\ \sqrt{\frac{\hbar\ \omega_k}{8\ \pi\ c^2}}\ \left(\ \hat{a}_{-\underline{k},\alpha}\ -\ \hat{a}^\dagger_{\underline{k},\alpha}\ \right)
\tag{594}
$$

as was anticipated.

10.2.1 The Energy of the Field

The Hamiltonian of the electromagnetic field

$$\hat{H} = \sum_{\underline{k},\alpha} \left[\frac{8\,\pi\,c^2}{4}\,\hat{\Pi}_\alpha(\underline{k})\,\hat{\Pi}_\alpha^\dagger(\underline{k}) + \frac{1}{8\,\pi}\,\underline{k}^2\,\hat{\Phi}_\alpha^\dagger(\underline{k})\,\hat{\Phi}_\alpha(\underline{k}) \right] \quad (595)$$

can be expressed in terms of the creation and annihilation operators as

$$
\begin{aligned}
\hat{H} &= \sum_{\underline{k},\alpha} \frac{\hbar\,\omega_k}{4} \left[\left(\hat{a}_{-\underline{k},\alpha}^\dagger - \hat{a}_{\underline{k},\alpha} \right) \left(\hat{a}_{-\underline{k},\alpha} - \hat{a}_{\underline{k},\alpha}^\dagger \right) \right. \\
&\qquad\qquad \left. + \left(\hat{a}_{\underline{k},\alpha}^\dagger + \hat{a}_{-\underline{k},\alpha} \right) \left(\hat{a}_{\underline{k},\alpha} + \hat{a}_{-\underline{k},\alpha}^\dagger \right) \right] \\
&= \sum_{\underline{k},\alpha} \frac{\hbar\,\omega_k}{4} \left[\hat{a}_{\underline{k},\alpha}^\dagger\,\hat{a}_{\underline{k},\alpha} + \hat{a}_{\underline{k},\alpha}\,\hat{a}_{\underline{k},\alpha}^\dagger + \hat{a}_{-\underline{k},\alpha}^\dagger\,\hat{a}_{-\underline{k},\alpha} + \hat{a}_{-\underline{k},\alpha}\,\hat{a}_{-\underline{k},\alpha}^\dagger \right]
\end{aligned}
$$
$$(596)$$

If one sets $\underline{k} \to -\underline{k}$ in the second set of terms, then one finds the Hamiltonian becomes the sum over independent harmonic oscillators for each $\underline{k}$ value and polarization

$$\hat{H} = \sum_{\underline{k},\alpha} \frac{\hbar\,\omega_k}{2} \left[\hat{a}_{\underline{k},\alpha}^\dagger\,\hat{a}_{\underline{k},\alpha} + \hat{a}_{\underline{k},\alpha}\,\hat{a}_{\underline{k},\alpha}^\dagger \right] \quad (597)$$

The number operator for each normal mode is given by

$$\hat{n}_{\underline{k},\alpha} = \hat{a}_{\underline{k},\alpha}^\dagger\,\hat{a}_{\underline{k},\alpha} \quad (598)$$

and has integer eigenvalues denoted by $n_{\underline{k},\alpha}$. Hence, the energy eigenvalues E are given by

$$E = \sum_{\underline{k},\alpha} \hbar\,\omega_k \left(n_{\underline{k},\alpha} + \frac{1}{2} \right) \quad (599)$$

The energy of the electromagnetic field is quantized in units of $\hbar\,\omega_{\underline{k}} = \hbar\,c\,k$. The quanta are known as photons.

It should be noted that the contributions to the total energy from the zero-point energy terms $\frac{\hbar\omega_k}{2}$ diverge. However, in most circumstances, only the excitation energy of the field is measurable, hence the divergence is mainly irrelevant. The zero-point energy does have physical consequences, and can be observed if the volume or boundary conditions of the field are changed. The change in the zero-point energy of the field due to change in volume or boundary conditions is known as the Casimir effect[24].

---

[24]H. B. G. Casimir, Proc. Neth. Aka. Wetenschapen, **51**, 793 (1948), M. J. Sparnaay, Physica **24**, 761 (1959)

10.2.2 The Electromagnetic Field

The quantized vector potential is given by the operator $\underline{\hat{A}}(\underline{r})$, given by

$$\underline{\hat{A}}(\underline{r}) = \sum_{\underline{k},\alpha} \hat{\epsilon}_\alpha(\underline{k}) \sqrt{\frac{2\pi\hbar c^2}{\omega_k V}} \left(\hat{a}^\dagger_{\underline{k},\alpha} + \hat{a}_{-\underline{k},\alpha} \right) \exp\left[-i\underline{k}\cdot\underline{r} \right] \quad (600)$$

In the Heisenberg representation, the time dependence of the vector potential is found from

$$i\hbar \frac{\partial \underline{A}(\underline{r},t)}{\partial t} = [\underline{A}(\underline{r},t), \hat{H}] \quad (601)$$

which has the solution

$$\underline{\hat{A}}(\underline{r},t) = \exp\left[+i\frac{t}{\hbar}\hat{H} \right] \underline{\hat{A}}(\underline{r},0) \exp\left[-i\frac{t}{\hbar}\hat{H} \right] \quad (602)$$

or

$$\underline{\hat{A}}(\underline{r},t) = \sum_{\underline{k},\alpha} \hat{\epsilon}_\alpha(\underline{k}) \sqrt{\frac{2\pi\hbar c^2}{\omega_k V}} \left(\hat{a}^\dagger_{\underline{k},\alpha} \exp\left[i\omega_k t \right] + \hat{a}_{-\underline{k},\alpha} \exp\left[-i\omega_k t \right] \right) \exp\left[-i\underline{k}\cdot\underline{r} \right]$$

$$(603)$$

The above equation was obtained by noting that, in the basis composed of eigenstates of the number operators $|n_{\underline{k},\alpha}>$, one has

$$\hat{a}_{\underline{k},\alpha}(t) |n_{\underline{k},\alpha}> = \exp\left[+i\omega_k t \, (\hat{a}^\dagger_{\underline{k},\alpha}\hat{a}_{\underline{k},\alpha}+1/2) \right] \hat{a}_{\underline{k},\alpha}(0) |n_{\underline{k},\alpha}> \exp\left[-i\omega_k t \,(n_{\underline{k},\alpha}+1/2) \right]$$

$$= \exp\left[+i\omega_k t \,(n_{\underline{k},\alpha}-1/2) \right] \sqrt{n_{\underline{k},\alpha}} |n_{\underline{k},\alpha}-1> \exp\left[-i\omega_k t \,(n_{\underline{k},\alpha}+1/2) \right]$$

$$= \exp\left[-i\omega_k t \right] \hat{a}_{\underline{k},\alpha} |n_{\underline{k},\alpha}> \quad (604)$$

and that the time-dependent creation operator is given by the Hermitean conjugate expression. Thus, the explicit form of time dependence of the vector potential is a consequence of the explicit time dependence of the creation and annihilation operators in the Heisenberg representation. Alternatively, one can find the time dependence of the creation and annihilation operators directly from the Heisenberg equations of motion without invoking a privileged set of basis states. The equation of motion for the creation operator is given by

$$i\hbar \frac{\partial \hat{a}^\dagger_{\underline{k},\alpha}}{\partial t} = [\hat{a}^\dagger_{\underline{k},\alpha}, \hat{H}] \quad (605)$$

and the commutator is evaluated as

$$[\hat{a}^\dagger_{\underline{k},\alpha}, \hat{a}^\dagger_{\underline{k}',\beta}\hat{a}_{\underline{k}',\beta}] = -\hat{a}^\dagger_{\underline{k},\alpha}\, \delta_{\alpha,\beta}\, \delta_{\underline{k},\underline{k}'} \quad (606)$$

so the equation of motion simplifies to

$$i \, \hbar \, \frac{\partial \hat{a}_{\underline{k},\alpha}^{\dagger}}{\partial t} \; = \; - \, \hbar \, \omega_k \, \hat{a}_{\underline{k},\alpha}^{\dagger} \tag{607}$$

Therefore, one finds the result

$$\hat{a}_{\underline{k},\alpha}^{\dagger}(t) \; = \; \hat{a}_{\underline{k},\alpha}^{\dagger} \, \exp \left[\, i \, \omega_k \, t \, \right] \tag{608}$$

Likewise, the annihilation operator satisfies the equation of motion

$$i \, \hbar \, \frac{\partial \hat{a}_{\underline{k},\alpha}}{\partial t} \; = \; [\, \hat{a}_{\underline{k},\alpha} \, , \, \hat{H} \,] \tag{609}$$

and as

$$[\, \hat{a}_{\underline{k},\alpha} \, , \, \hat{a}_{\underline{k}',\beta}^{\dagger} \, \hat{a}_{\underline{k}',\beta} \,] \; = \; + \, \hat{a}_{\underline{k},\alpha} \, \delta_{\alpha,\beta} \, \delta_{\underline{k},\underline{k}'} \tag{610}$$

so the equation of motion simplifies to

$$i \, \hbar \, \frac{\partial \hat{a}_{\underline{k},\alpha}}{\partial t} \; = \; + \, \hbar \, \omega_k \, \hat{a}_{\underline{k},\alpha} \tag{611}$$

Hence, one finds that the time-dependent annihilation operator is given by

$$\hat{a}_{\underline{k},\alpha}(t) \; = \; \hat{a}_{\underline{k},\alpha} \, \exp \left[\, - \, i \, \omega_k \, t \, \right] \tag{612}$$

which is just the Hermitean conjugate of the $\hat{a}_{\underline{k},\alpha}^{\dagger}(t)$ that was found previously. Therefore, the time-dependence of the vector potential is entirely due to the time-dependence of the Heisenberg representation of the creation and annihilation operators.

10.2.3 The Momentum of the Field

The total momentum operator for the electromagnetic field is given by the integral over all space of the Poynting vector

$$\hat{\underline{P}} \; = \; \frac{1}{4 \, \pi \, c} \, \int \, d^3\underline{r} \, \left(\, \hat{\underline{E}} \, \wedge \, \hat{\underline{B}} \, \right) \tag{613}$$

This will be evaluated by expressing the $\hat{\underline{E}}$ and $\hat{\underline{B}}$ field operators in terms of the vector potential $\underline{A}$ operator via

$$\begin{aligned} \hat{\underline{E}} \; &= \; - \, \frac{1}{c} \, \frac{\partial \hat{\underline{A}}}{\partial t} \\ \hat{\underline{B}} \; &= \; \underline{\nabla} \, \wedge \, \hat{\underline{A}} \end{aligned} \tag{614}$$

The vector potential operator can be written in terms of the creation and annihilation operators for the normal modes as

$$\underline{\hat{A}}(\underline{r},t) \;=\; \sum_{\underline{k},\alpha} \hat{\epsilon}_\alpha(\underline{k}) \sqrt{\frac{2\,\pi\,\hbar\,c^2}{\omega_k\,V}} \left(\hat{a}^\dagger_{\underline{k},\alpha}(t) + \hat{a}_{-\underline{k},\alpha}(t) \right) \exp\left[-\,i\,\underline{k}\,.\,\underline{r} \right] \quad (615)$$

then the $\underline{E}$ and $\underline{B}$ field operators are found as

$$\underline{\hat{E}}(\underline{r}) \;=\; -\,i \sum_{\underline{k},\alpha} \hat{\epsilon}_\alpha(\underline{k}) \sqrt{\frac{2\,\pi\,\hbar\,\omega_k}{V}} \left(\hat{a}^\dagger_{\underline{k},\alpha} - \hat{a}_{-\underline{k},\alpha} \right) \exp\left[-\,i\,\underline{k}\,.\,\underline{r} \right] \quad (616)$$

and

$$\underline{\hat{B}}(\underline{r}) \;=\; -\,i \sum_{\underline{k},\alpha} (\,\underline{k} \wedge \hat{\epsilon}_\alpha(\underline{k})\,) \sqrt{\frac{2\,\pi\,\hbar\,c^2}{\omega_k\,V}} \left(\hat{a}^\dagger_{\underline{k},\alpha} + \hat{a}_{-\underline{k},\alpha} \right) \exp\left[-\,i\,\underline{k}\,.\,\underline{r} \right]$$
$$(617)$$

For a fixed $\underline{k}$, the polarization vectors $\hat{\epsilon}_\alpha(\underline{k})$ and $\underline{k}$ are mutually orthogonal. Therefore, one has

$$\hat{\epsilon}_\alpha(\underline{k}) \wedge (\,\underline{k} \wedge \hat{\epsilon}_\beta(\underline{k})\,) \;=\; \underline{k}\,(\,\hat{\epsilon}_\alpha(\underline{k})\,.\,\hat{\epsilon}_\beta(\underline{k})\,) - \hat{\epsilon}_\beta(\underline{k})\,(\,\underline{k}\,.\,\hat{\epsilon}_\alpha(\underline{k})\,)$$
$$=\; \underline{k}\,\delta_{\alpha,\beta} \quad (618)$$

Hence, the total momentum of the electromagnetic field is determined from

$$\underline{\hat{P}} \;=\; \frac{\hbar}{2} \sum_{\underline{k},\alpha} \hat{\epsilon}_\alpha(\underline{k}) \wedge (\,\underline{k} \wedge \hat{\epsilon}_\alpha(\underline{k})\,) \left(\hat{a}^\dagger_{\underline{k},\alpha} - \hat{a}_{-\underline{k},\alpha} \right) \left(\hat{a}^\dagger_{-\underline{k},\alpha} + \hat{a}_{\underline{k},\alpha} \right)$$
$$=\; \frac{\hbar}{2} \sum_{\underline{k},\alpha} \underline{k} \left(\hat{a}^\dagger_{\underline{k},\alpha} - \hat{a}_{-\underline{k},\alpha} \right) \left(\hat{a}^\dagger_{-\underline{k},\alpha} + \hat{a}_{\underline{k},\alpha} \right) \quad (619)$$

It should be noted that the momentum from each normal mode of the field is parallel to its direction of propagation. Since the creation operators commute

$$\hat{a}^\dagger_{\underline{k},\alpha}\,\hat{a}^\dagger_{-\underline{k},\alpha} \;=\; \hat{a}^\dagger_{-\underline{k},\alpha}\,\hat{a}^\dagger_{\underline{k},\alpha} \quad (620)$$

and that the annihilation operators also commute

$$\hat{a}_{-\underline{k},\alpha}\,\hat{a}_{\underline{k},\alpha} \;=\; \hat{a}_{\underline{k},\alpha}\,\hat{a}_{-\underline{k},\alpha} \quad (621)$$

one finds that the part of the momentum represented by the summation over $\underline{k}$ given by

$$\hbar \sum_{\underline{k},\alpha} \underline{k} \left(\hat{a}^\dagger_{\underline{k},\alpha}\,\hat{a}^\dagger_{-\underline{k},\alpha} - \hat{a}_{-\underline{k},\alpha}\,\hat{a}_{\underline{k},\alpha} \right) \;=\; 0 \quad (622)$$

vanishes since the summand is odd under inversion symmetry. Thus, the momentum of the electromagnetic field is given by

$$\underline{\hat{P}} \;=\; \frac{\hbar}{2} \sum_{\underline{k},\alpha} \underline{k} \left(\hat{a}^\dagger_{\underline{k},\alpha}\,\hat{a}_{\underline{k},\alpha} - \hat{a}_{-\underline{k},\alpha}\,\hat{a}^\dagger_{-\underline{k},\alpha} \right)$$
$$=\; \frac{1}{2} \sum_{\underline{k},\alpha} \left(\hbar\,\underline{k}\,\hat{a}^\dagger_{\underline{k},\alpha}\,\hat{a}_{\underline{k},\alpha} - \hbar\,\underline{k}\,\hat{a}^\dagger_{-\underline{k},\alpha}\,\hat{a}_{-\underline{k},\alpha} - \hbar\,\underline{k} \right) \quad (623)$$

where the commutation relations for the creation and annihilation operators were used to obtain the last line. The last term vanishes when summed over $\underline{k}$, due to inversion symmetry. Hence, the momentum of the field is given by the operator

$$\hat{\underline{P}} \;=\; \frac{1}{2} \sum_{\underline{k},\alpha} \left(\hbar \, \underline{k} \, \hat{a}^{\dagger}_{\underline{k},\alpha} \, \hat{a}_{\underline{k},\alpha} \;-\; \hbar \, \underline{k} \, \hat{a}^{\dagger}_{-\underline{k},\alpha} \, \hat{a}_{-\underline{k},\alpha} \right) \qquad (624)$$

Finally, on transforming $-\underline{k}$ to $\underline{k}$ in the last term of the summand, one finds the total momentum of the field is carried by the excitations since

$$\hat{\underline{P}} \;=\; \sum_{\underline{k},\alpha} \hbar \, \underline{k} \, \hat{a}^{\dagger}_{\underline{k},\alpha} \, \hat{a}_{\underline{k},\alpha} \qquad (625)$$

Thus, each quantum excitation of wave vector $\underline{k}$ has momentum $\hbar \, \underline{k}$.

Since a photon has an energy $\hbar \, c \, k$ and momentum $\hbar \, k$, these quanta are massless because the mass of the quanta are defined as the relativistic invariant length of the momentum four-vector

$$\left(\frac{E}{c} \right)^{2} \;-\; p^{2} \;=\; m^{2} \, c^{2} \qquad (626)$$

which yields $m \;=\; 0$. The energy-momentum dispersion relation of the quanta of the electromagnetic field was conclusively demonstrated by A. H. Compton[25]. Compton showed that when quanta are scattered by charged particles, the photon's dispersion relation follows directly by application of conservation laws to the recoiling particle.

10.2.4 The Angular Momentum of the Field

The total angular momentum operator of the electromagnetic field $\hat{\underline{J}}_{EM}$ is given by

$$\hat{\underline{J}}_{EM} \;=\; \frac{1}{4 \, \pi \, c} \int d^3\underline{r} \left(\underline{r} \, \wedge \, (\, \hat{\underline{E}} \, \wedge \, \hat{\underline{B}} \,) \right) \qquad (627)$$

The i-th component is given by

$$\begin{aligned}
\hat{J}^{(i)}_{EM} \;&=\; \frac{1}{4 \, \pi \, c} \int d^3\underline{r} \; \xi^{i,j,k} \left(x^{(j)} \, (\, \hat{\underline{E}} \, \wedge \, \hat{\underline{B}} \,)^{(k)} \right) \\
&=\; \frac{1}{4 \, \pi \, c} \int d^3\underline{r} \; \xi^{i,j,k} \left(x^{(j)} \, \xi^{k,l,m} \, \hat{E}^{(l)} \, \hat{B}^{(m)} \right) \\
&=\; \frac{1}{4 \, \pi \, c} \int d^3\underline{r} \; \xi^{i,j,k} \left(x^{(j)} \, \xi^{k,l,m} \, \hat{E}^{(l)} \, \xi^{m,n,p} \, \frac{\partial \hat{A}^{(p)}}{\partial x^{(n)}} \right)
\end{aligned}$$

$$(628)$$

[25] A. H. Compton, Phys. Rev. *Second Series*, **21**, 483 (1923).

However, due to the identity

$$\xi^{k,l,m} \; \xi^{m,n,p} \;=\; \left(\; \delta^{k,n} \; \delta^{l,p} \;-\; \delta^{k,p} \; \delta^{l,n} \; \right) \tag{629}$$

one finds

$$\hat{J}^{(i)}_{EM} \;=\; \frac{1}{4 \; \pi \; c} \; \int \; d^3\underline{r} \; \xi^{i,j,k} \; \left(\; x^{(j)} \; \hat{E}^{(l)} \; \frac{\partial \hat{A}^{(l)}}{\partial x^{(k)}} \;-\; x^{(j)} \; \hat{E}^{(l)} \; \frac{\partial \hat{A}^{(k)}}{\partial x^{(l)}} \; \right) \tag{630}$$

On integrating by parts in the last term, one has

$$\hat{J}^{(i)}_{EM} \;=\; \frac{1}{4 \; \pi \; c} \; \int \; d^3\underline{r} \; \xi^{i,j,k} \; \left(\; x^{(j)} \; \hat{E}^{(l)} \; \frac{\partial \hat{A}^{(l)}}{\partial x^{(k)}} \;+\; \frac{\partial}{\partial x^{(l)}} \left(\; x^{(j)} \; \hat{E}^{(l)} \; \right) \; \hat{A}^{(k)} \; \right) \tag{631}$$

Since the divergence of the electric field vanishes[26],

$$\frac{\partial \hat{E}^{(l)}}{\partial x^{(l)}} \;=\; 0 \tag{634}$$

and since

$$\frac{\partial x^{(j)}}{\partial x^{(l)}} \;=\; \delta^{j,l} \tag{635}$$

the total angular momentum can be re-written as

$$\hat{J}^{(i)}_{EM} \;=\; \frac{1}{4 \; \pi \; c} \; \int \; d^3\underline{r} \; \xi^{i,j,k} \; \left(\; \hat{E}^{(l)} \; x^{(j)} \; \frac{\partial \hat{A}^{(l)}}{\partial x^{(k)}} \;+\; \hat{E}^{(j)} \; \hat{A}^{(k)} \; \right) \tag{636}$$

The first term can be recognized as the orbital angular momentum of the field. The orbital angular momentum operator $\hat{L}^{(i)}$ is given by

$$\hat{L}^{(i)} \;=\; -\; i \; \hbar \; \xi^{i,j,k} \; x^{(j)} \; \frac{\partial}{\partial x^{(k)}} \tag{637}$$

so the total angular momentum of the field is given by

$$\hat{J}^{(i)}_{EM} \;=\; \frac{i}{4 \; \pi \; \hbar \; c} \; \int \; d^3\underline{r} \; \left(\; \hat{E}^{(l)} \; \hat{L}^{(i)} \; \hat{A}^{(l)} \;-\; i \; \hbar \; \xi^{i,j,k} \; \hat{E}^{(j)} \; \hat{A}^{(k)} \; \right)$$

$$=\; \frac{i}{4 \; \pi \; \hbar \; c} \; \int \; d^3\underline{r} \; \left(\; \hat{E}^{(l)} \; \hat{L}^{(i)} \; \hat{A}^{(l)} \;+\; \hat{E}^{(j)} \; (\; \hat{S}^{(i)} \;)^{j,k} \; \hat{A}^{(k)} \; \right) \tag{638}$$

---

[26] In the presence of a charge density $q \; |\psi(\underline{r})|^2$, the angular momentum of the EM field will contain a term given by

$$-\; \frac{q}{c} \; \int \; d^3\underline{r} \; \psi^*(\underline{r}) \; \underline{r} \; \wedge \; \underline{A}(\underline{r}) \; \psi(\underline{r}) \tag{632}$$

which is gauge-dependent. This term combines with a corresponding term in the expression for the orbital angular momentum of the charged particles to yield the gauge-invariant term

$$\int \; d^3\underline{r} \; \psi^*(\underline{r}) \; \underline{r} \; \wedge \; (\; \hat{\underline{p}} \;-\; \frac{q}{c} \; \underline{A}(\underline{r}) \;) \; \psi(\underline{r}) \tag{633}$$

Hence, although the total angular momentum of a system of charged particles in an electromagnetic field is well-defined and gauge-invariant, there could be some confusion as to how the angular momentum is distributed between the particles and the field.

where the definition

$$(\ \hat{S}^{(i)} \ )^{j,k} \ = \ - \ i \ \hbar \ \xi^{i,j,k} \tag{639}$$

for $\underline{\hat{S}}$, the intrinsic spin operator for the photon, has been used in obtaining the second line. The total vector angular momentum operator can be expressed as

$$\underline{\hat{J}}_{EM} \ = \ \frac{i}{4 \ \pi \ \hbar \ c} \ \int \ d^3\underline{r} \ \hat{E}^{(j)} \ \left(\ \underline{\hat{L}} \ \delta^{j,k} \ + \ (\ \underline{\hat{S}} \ )^{j,k} \ \right) \ \hat{A}^{(k)} \tag{640}$$

which shows that the orbital angular momentum is diagonal with respect to the field components and the spin angular momentum mixes the different field components.

The total spin component of the angular momentum operator for the electromagnetic field is given by

$$
\begin{aligned}
\hat{S}^{(i)}_{EM} \ &= \ \frac{i}{4 \ \pi \ \hbar \ c} \ \int \ d^3\underline{r} \ \left(\ \hat{E}^{(j)} \ (\ \hat{S}^{(i)} \ )^{j,k} \ \hat{A}^{(k)} \ \right) \\
&= \ \frac{1}{4 \ \pi \ c} \ \int \ d^3\underline{r} \ \xi^{i,j,k} \ \hat{E}^{(j)} \ \hat{A}^{(k)} \\
&= \ \frac{1}{4 \ \pi \ c} \ \int \ d^3\underline{r} \ \left(\ \underline{\hat{E}} \ \wedge \ \underline{\hat{A}} \ \right)^{(i)}
\end{aligned} \tag{641}
$$

This can be expressed in terms of the photon creation and annihilation operators as

$$
\begin{aligned}
\hat{S}^{(i)}_{EM} \ = \ &- \ i \ \frac{\hbar}{2} \ \sum_{\underline{k},\alpha,\beta} \ \left(\ \hat{\epsilon}^{(j)}_{\beta}(\underline{k}) \ \xi^{i,j,k} \ \hat{\epsilon}^{(k)}_{\alpha}(\underline{k}) \ \right) \\
&\times \ \left(\ \hat{a}^{\dagger}_{-\underline{k},\beta} \ - \ \hat{a}_{\underline{k},\beta} \ \right) \ \left(\ \hat{a}^{\dagger}_{\underline{k},\alpha} \ + \ \hat{a}_{-\underline{k},\alpha} \ \right)
\end{aligned} \tag{642}
$$

The first term in parenthesis is recognized as the i-th component of the vector product

$$\hat{\epsilon}_{\beta}(\underline{k}) \ \wedge \ \hat{\epsilon}_{\alpha}(\underline{k}) \tag{643}$$

and, therefore, it is antisymmetric in the polarization indices α and β and the non-zero contributions are restricted to the case $\alpha \ \neq \ \beta$. Since the creation and annihilation operators corresponding to different polarizations commute, the product of the two remaining parenthesis can be re-arranged as the sum of two terms

$$
\begin{aligned}
\hat{S}^{(i)}_{EM} \ = \ &- \ i \ \frac{\hbar}{2} \ \sum_{\underline{k},\alpha,\beta} \ \left(\ \hat{\epsilon}_{\beta}(\underline{k}) \ \wedge \ \hat{\epsilon}_{\alpha}(\underline{k}) \ \right)^{(i)} \\
&\times \ \left[\ \left(\ \hat{a}^{\dagger}_{-\underline{k},\beta} \ \hat{a}^{\dagger}_{\underline{k},\alpha} \ - \ \hat{a}_{\underline{k},\beta} \ \hat{a}_{-\underline{k},\alpha} \ \right) \right. \\
&\left. + \ \left(\ \hat{a}^{\dagger}_{-\underline{k},\beta} \ \hat{a}_{-\underline{k},\alpha} \ - \ \hat{a}^{\dagger}_{\underline{k},\alpha} \ \hat{a}_{\underline{k},\beta} \ \right) \ \right]
\end{aligned} \tag{644}
$$

On transforming the summation variable $\underline{k} \to -\underline{k}$ and commuting the operators, one finds that the first term in the square brackets is symmetric under the interchange of α and β whereas the second term is antisymmetric. Hence, on summing over the polarization indices, the contribution from the first term vanishes, as it is the product of a symmetric and the antisymmetric (vector product) term. Therefore, the total spin operator of the electromagnetic field is expressed as

$$
\hat{S}_{EM}^{(i)} \;=\; i\,\frac{\hbar}{2}\, \sum_{\underline{k},\alpha,\beta} \left(\hat{\epsilon}_\beta(\underline{k}) \;\wedge\; \hat{\epsilon}_\alpha(\underline{k}) \right)^{(i)} \left(\hat{a}_{\underline{k},\alpha}^\dagger\, \hat{a}_{\underline{k},\beta} \;-\; \hat{a}_{-\underline{k},\beta}^\dagger\, \hat{a}_{-\underline{k},\alpha} \right)
$$

(645)

On defining the sense of the polarization vectors relative to $\hat{\underline{k}}$ ($\equiv \hat{e}_3(\underline{k})$ the unit vector in the direction of propagation) via

$$
\left(\hat{\epsilon}_1(\underline{k}) \;\wedge\; \hat{\epsilon}_2(\underline{k}) \right) \;=\; \hat{\underline{k}}
$$

(646)

so that $\hat{\underline{k}}$ corresponds to the z-direction, one finds that

$$
\begin{aligned}
\hat{\underline{S}}_{EM} \;=\; & i\,\frac{\hbar}{2}\, \sum_{\underline{k}} \left[\hat{\underline{k}} \left(\hat{a}_{\underline{k},2}^\dagger\, \hat{a}_{\underline{k},1} \;-\; \hat{a}_{\underline{k},1}^\dagger\, \hat{a}_{\underline{k},2} \right) \right. \\
& \left. -\; \hat{\underline{k}} \left(\hat{a}_{-\underline{k},2}^\dagger\, \hat{a}_{-\underline{k},1} \;-\; \hat{a}_{-\underline{k},1}^\dagger\, \hat{a}_{-\underline{k},2} \right) \right]
\end{aligned}
$$

(647)

On setting $-\underline{k} \to \underline{k}$ in the second part of the summation, the spin of the electromagnetic field is found as

$$
\hat{\underline{S}}_{EM} \;=\; i\,\hbar\, \sum_{\underline{k}} \hat{\underline{k}} \left(\hat{a}_{\underline{k},2}^\dagger\, \hat{a}_{\underline{k},1} \;-\; \hat{a}_{\underline{k},1}^\dagger\, \hat{a}_{\underline{k},2} \right)
$$

(648)

It should be noted that in this expression, the indices $(1,2)$ refer to directions in three-dimensional space and do not refer to the z-component of the spin angular momentum. Therefore, the above equation shows that a plane-polarized photon is not an eigenstate of the single-particle spin operator quantized along the k-axis[27].

In our Cartesian component basis, the eigenstates of the component of the spin operator parallel to the direction of propagation $\hat{S}^{(3)}$, where

$$
\hat{S}^{(3)} \;=\; \hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
$$

(649)

---

[27]Strictly speaking, this quantum number corresponds to the helicity as it is the spin eigenvalue which is quantized along the direction of propagation.

are given by $\widetilde{\underline{\Phi}}_m(\underline{k})$, where

$$
\widetilde{\underline{\Phi}}_{+1}(\underline{k}) \;=\; -\,\frac{1}{\sqrt{2}}\,\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}
$$

$$
\widetilde{\underline{\Phi}}_{0}(\underline{k}) \;=\; \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
$$

$$
\widetilde{\underline{\Phi}}_{-1}(\underline{k}) \;=\; \frac{1}{\sqrt{2}}\,\begin{pmatrix} 1 \\ -\,i \\ 0 \end{pmatrix} \tag{650}
$$

and where the subscript m refers to the eigenvalue of $\hat{S}^{(3)}$, in units of $\hbar$. From this, it follows that an arbitrary transverse vector wave function $\underline{\Phi}(\underline{k})$ can only be expressed as a linear superposition of states involving $m = \pm 1$, and that the $m = 0$ component is absent. On expressing an arbitrary (non-transverse) vector wave function $\underline{\Phi}(\underline{k})$ with components $\Phi^{(1)}(\underline{k})$, $\Phi^{(2)}(\underline{k})$ and $\Phi^{(3)}(\underline{k})$ in terms of its components referred to the helicity eigenstates $\Phi_m(\underline{k})$ one has

$$
\begin{pmatrix} \Phi_{+1}(\underline{k}) \\ \Phi_{0}(\underline{k}) \\ \Phi_{-1}(\underline{k}) \end{pmatrix} \;=\; \frac{1}{\sqrt{2}}\,\begin{pmatrix} -\,1 & i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & i & 0 \end{pmatrix}\begin{pmatrix} \Phi^{(1)}(\underline{k}) \\ \Phi^{(2)}(\underline{k}) \\ \Phi^{(3)}(\underline{k}) \end{pmatrix} \tag{651}
$$

This relation between the two bases can be expressed in the alternate form

$$
\underline{\Phi}(\underline{k}) \;=\; \sum_{m=-1}^{m=1} \hat{e}_m\,\Phi_m(\underline{k}) \tag{652}
$$

where the circularly-polarized unit vectors are introduced via

$$
\hat{e}_{+1} \;=\; -\,\frac{1}{\sqrt{2}}\,\big(\,\hat{\epsilon}_1(\underline{k}) \;+\; i\,\hat{\epsilon}_2(\underline{k})\,\big)
$$

$$
\hat{e}_0 \;=\; \hat{\epsilon}_3(\underline{k})
$$

$$
\hat{e}_{-1} \;=\; \frac{1}{\sqrt{2}}\,\big(\,\hat{\epsilon}_1(\underline{k}) \;-\; i\,\hat{\epsilon}_2(\underline{k})\,\big) \tag{653}
$$

The circularly-polarized unit vectors are associated with photons which have definite helicity eigenvalues. It should be noted that these complex unit vectors are orthogonal, and satisfy

$$
\hat{e}_{m'}^{*} \cdot \hat{e}_m \;=\; \delta_{m,m'} \tag{654}
$$

The above relations allow one to define the circularly-polarized creation and annihilation operators via their relation to the quantum fields. This procedure yields

$$
\sum_{i=1}^{i=3} \hat{\epsilon}_i(\underline{k})\,\hat{a}_{\underline{k},i} \;=\; \sum_{m=-1}^{m=1} \hat{e}_m(\underline{k})\,\hat{a}_{\underline{k},m} \tag{655}
$$

104

Hence, the photon annihilation operators corresponding to a definite helicity are related to the annihilation operators for plane-polarized photons via

$$\hat{a}_{\underline{k},m=+1} = -\frac{1}{\sqrt{2}} (\hat{a}_{\underline{k},1} - i \hat{a}_{\underline{k},2})$$
$$\hat{a}_{\underline{k},m=0} = \hat{a}_{\underline{k},3}$$
$$\hat{a}_{\underline{k},m=-1} = \frac{1}{\sqrt{2}} (\hat{a}_{\underline{k},1} + i \hat{a}_{\underline{k},2}) \tag{656}$$

and the inverse relations are given by

$$\hat{a}_{\underline{k},1} = -\frac{1}{\sqrt{2}} (\hat{a}_{\underline{k},m=1} - \hat{a}_{\underline{k},m=-1})$$
$$\hat{a}_{\underline{k},2} = -\frac{i}{\sqrt{2}} (\hat{a}_{\underline{k},m=1} + \hat{a}_{\underline{k},m=-1})$$
$$\hat{a}_{\underline{k},3} = \hat{a}_{\underline{k},m=0} \tag{657}$$

When expressed in terms of the circularly-polarized unit vectors, the spin operator for the electromagnetic field becomes

$$\hat{\underline{S}}_{EM} = \hbar \sum_{\underline{k}} \hat{\underline{k}} \left(\hat{a}^{\dagger}_{\underline{k},m=1} \hat{a}_{\underline{k},m=1} - \hat{a}^{\dagger}_{\underline{k},m=-1} \hat{a}_{\underline{k},m=-1} \right) \tag{658}$$

which is expressed in terms of photons with definite helicity. Within the manifold of single-photon states with momentum $\hbar \underline{k}$, the spin operator has eigenvalues of $\pm\hbar$ when measured along the direction $\hat{\underline{k}}$. It is seen that the photon has helicity $m = \pm1$ but does not involve the helicity state with $m = 0$ since the electromagnetic field is transverse. The transverse nature of the field is due to the photon being massless. In general, a massive particle with spin S should have $(2S + 1)$ helicity states. However, a massless particle can only have the two helicity states corresponding to $m = \pm S$.

The angular momentum of the elementary excitation of the electromagnetic field was inferred from experiments in which beams of circularly-polarized light were absorbed by a sensitive torsional pendulum[28]. Quantum electromagnetic theory shows that the angular momentum density of left circularly-polarized light is just $\hbar$ times the photon density or, equivalently, is just ω^{-1} times the energy density which is also the case for classical electromagnetism[29]. Hence, the net increase of angular momentum per unit time can easily be calculated from the excess of the angular momentum flux flowing into the pendulum over that flowing out. Beth's experiments verified that the net torque on the pendulum was consistent with the theoretical prediction. Thus, the quantized electromagnetic field has been shown to be related to a massless particle with spin $\hbar$ and energy-momentum given by the four-vector $(\hbar\omega_k/c, \hbar\underline{k})$. This particle is the photon. Every quantized field is to be associated with a type of particle.

---

[28] R. A. Beth, Phys. Rev. **50**, 115 (1936).
[29] J. H. Poynting, Proc. Roy. Soc. A**82**, 560 (1909).
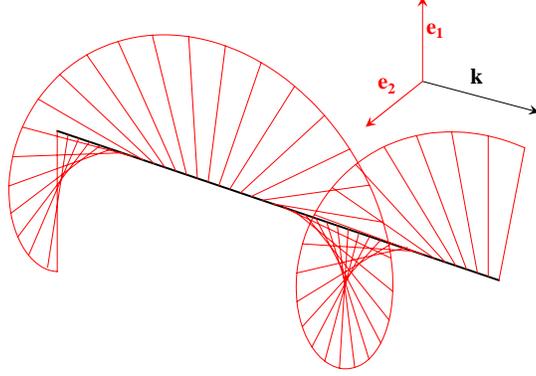
Figure 11: The circularly-polarized normal modes of a classical electromagnetic field are composed of two plane-polarized waves which are out of phase, are mutually orthogonal, and are transverse to the direction of propagation $\underline{k}$. The resulting electric field spirals along the direction of propagation. The left circularly-polarized wave shown in the diagram corresponds to a helicity of $+\hbar$.

10.3 Uncertainty Relations

The eigenstates of the field operators such as $\underline{\hat{A}}(\underline{r}, t)$ do not correspond to eigenstates of the photon number operators.

Consider the electric field

$$
\underline{\hat{E}} \;=\; -\; \frac{1}{c}\; \frac{\partial \underline{\hat{A}}}{\partial t} \tag{659}
$$

Although the expectation value of $\underline{\hat{E}}$ vanishes for any eigenstate of the set of occupation numbers $\mid \{ n_{\underline{k}', \beta} \} \;>$

$$
<\; \{ n_{\underline{k}', \beta} \} \mid \underline{\hat{E}} \mid \{ n_{\underline{k}', \beta} \} \;>\; =\; 0 \tag{660}
$$

since

$$
<\; \{ n_{\underline{k}', \beta} \} \mid a_{\underline{k}, \alpha} \mid \{ n_{\underline{k}', \beta} \} \;>\; =\; 0 \tag{661}
$$

the fluctuation in the field is given by

$$
\begin{aligned}
<\; \{ n_{\underline{k}', \beta} \} \mid \underline{\hat{E}} \cdot \underline{\hat{E}} \mid \{ n_{\underline{k}', \beta} \} \;>\; &-\; \mid <\; \{ n_{\underline{k}', \beta} \} \mid \underline{\hat{E}} \mid \{ n_{\underline{k}', \beta} \} \;>\; \mid^2 \\
&=\; <\; \{ n_{\underline{k}', \beta} \} \mid \underline{\hat{E}} \cdot \underline{\hat{E}} \mid \{ n_{\underline{k}', \beta} \} \;> \\
&=\; \frac{4\,\pi}{V} \sum_{\underline{k}, \alpha} \hbar \omega_k \left(n_{\underline{k}, \alpha} + \frac{1}{2} \right) \\
&\rightarrow\; \infty \tag{662}
\end{aligned}
$$

The fluctuations in the field diverge because the zero-point energy fluctuations diverge.

106

The commutation relations between the x-component of the $\underline{E}$ field and the $\underline{B}$ field at the same instant of time are non-zero[30]. That is,

$$
\begin{aligned}
[\, \hat{E}_x(\underline{r}) \, , \, \hat{B}_y(\underline{r}') \,] \quad &= \quad \frac{2 \, \pi}{V} \, \sum_{\underline{k}, \alpha} \, \hbar \, \omega_k \, \hat{\epsilon}_\alpha(\underline{k})_x \, (\, \hat{\underline{k}} \, \wedge \, \hat{\epsilon}_\alpha(\underline{k}) \,)_y \, \exp \left[\, i \, \underline{k} \, . \, (\, \underline{r}' \, - \, \underline{r} \,) \, \right] \\
&\quad - \frac{2 \, \pi}{V} \, \sum_{\underline{k}, \alpha} \, \hbar \, \omega_k \, \hat{\epsilon}_\alpha(\underline{k})_x \, (\, \hat{\underline{k}} \, \wedge \, \hat{\epsilon}_\alpha(\underline{k}) \,)_y \, \exp \left[\, - \, i \, \underline{k} \, . \, (\, \underline{r}' \, - \, \underline{r} \,) \, \right] \\
&= \quad - \frac{4 \, \pi \, \hbar \, c}{V} \, \sum_{\underline{k}} \, k_z \, \exp \left[\, - \, i \, \underline{k} \, . \, (\, \underline{r}' \, - \, \underline{r} \,) \, \right] \\
&= \quad i \, \frac{4 \, \pi \, c \, \hbar}{V} \, \frac{\partial}{\partial z} \, \sum_{\underline{k}} \, \exp \left[\, - \, i \, \underline{k} \, . \, (\, \underline{r}' \, - \, \underline{r} \,) \, \right] \\
&= \quad i \, \frac{c \, \hbar}{2 \, \pi^2} \, \frac{\partial}{\partial z} \, \delta^3(\underline{r}' - \underline{r})
\end{aligned}
\tag{663}
$$

The fact that the two polarizations are transverse to the unit vector $\hat{k}$ has been used to obtain the third line. Since $\hat{\underline{E}}$ and $\hat{\underline{B}}$ do not commute, it follows that $\underline{E}$ and $\underline{B}$ obey an uncertainty relation in that the values of $\underline{E}$ and $\underline{B}$ cannot both be specified to arbitrary accuracy at the same point.

However, if two points in space time x and x' are not causally related, i.e.

$$
| \, \underline{r}' \, - \, \underline{r} \, | \, \neq \, c \, | \, t' \, - \, t \, |
\tag{664}
$$

then the operators commute

$$
[\, \hat{E}_x(\underline{r}, t) \, , \, \hat{B}_y(\underline{r}', t') \,] \, = \, 0
\tag{665}
$$

Thus, if the two points in space-time are not connected by the propagation of light, then the E_x and B_y fields can both be determined to arbitrary accuracy.

10.4 Coherent States

We shall focus our attention on one normal mode of the electromagnetic field, and shall drop the indices $(\underline{k}, \alpha)$ labelling the normal mode. A coherent state $| \, a_\varphi \, >$ is defined as an eigenstate of the annihilation operator

$$
\hat{a} \, | \, a_\varphi \, > \, = \, a_\varphi \, | \, a_\varphi \, >
\tag{666}
$$

For example, the vacuum state or ground state is an eigenstate of the annihilation operator, in which case $a_\varphi \, = \, 0$.

---

[30]P. Jordan and W. Pauli Jr. **47**, 151 (1927).

The coherent state[31] can be found as a linear superposition of eigenstates of the number operator with eigenvalues n

$$| a_\varphi > = \sum_{n=0}^{\infty} C_n | n > \tag{667}$$

On substituting this form in the definition of the coherent state

$$\begin{aligned}
\hat{a} | a_\varphi > &= \sum_n C_n \hat{a} | n > \\
&= a_\varphi \sum_n C_n | n >
\end{aligned} \tag{668}$$

and using the property of the annihilation operator, one has

$$\sum_n C_n \sqrt{n} | n - 1 > = a_\varphi \sum_n C_n | n > \tag{669}$$

On taking the matrix elements of this equation with the state $< m |$, and using the orthonormality of the eigenstates of the number operator, one finds

$$C_{m+1} \sqrt{m + 1} = a_\varphi C_m \tag{670}$$

Hence, on iterating downwards, one finds

$$C_m = \left(\frac{a_\varphi^m}{\sqrt{m!}} \right) C_0 \tag{671}$$

and the coherent state can be expressed as

$$| a_\varphi > = C_0 \sum_{n=0}^{\infty} \left(\frac{a_\varphi^n}{\sqrt{n!}} \right) | n > \tag{672}$$

The normalization constant C_0 can be found from

$$1 = C_0^* C_0 \sum_{n=0}^{\infty} \left(\frac{a_\varphi^{n*} a_\varphi^n}{n!} \right) \tag{673}$$

by noting that the sum exponentiates to yield

$$1 = C_0^* C_0 \exp \left[a_\varphi^* a_\varphi \right] \tag{674}$$

so, on choosing the phase of C_0, one has

$$C_0 = \exp \left[- \frac{1}{2} a_\varphi^* a_\varphi \right] \tag{675}$$
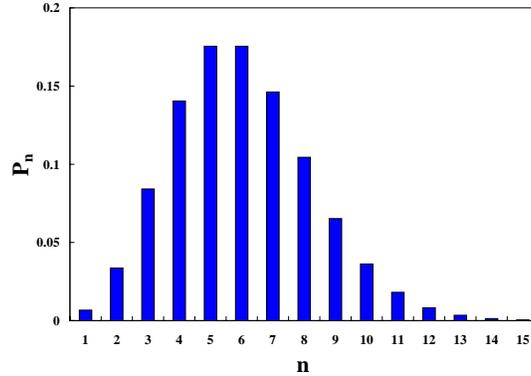
Figure 12: The probability of finding n photons $P(n)$ in a normal mode represented by a coherent state.

From this, it can be shown that if the number of photons in a coherent state are measured, the result n will occur with a probability given by

$$P(n) \;=\; \frac{(\, a_\varphi^* \, a_\varphi \,)^n}{n!} \; \exp\left[\, - \, a_\varphi^* \, a_\varphi \, \right] \tag{676}$$

Thus, the photon statistics are governed by a Poisson distribution. Furthermore, the quantity $a_\varphi^* \, a_\varphi$ is the average number of photons $\bar{n}$ present in the coherent state.

The coherent states can be written in a more compact form. Since the state with occupation number n can be written as

$$\mid n \;>\; = \; \frac{(\, \hat{a}^\dagger \,)^n}{\sqrt{n!}} \mid 0 \;> \tag{677}$$

the coherent state can also be expressed as

$$\mid a_\varphi \;>\; = \; \exp\left[\, - \, \frac{1}{2} \, a_\varphi^* \, a_\varphi \, \right] \sum_{n=0}^{\infty} \frac{(\, a_\varphi \, \hat{a}^\dagger \,)^n}{n!} \mid 0 \;> \tag{678}$$

or on summing the series as an exponential

$$\mid a_\varphi \;>\; = \; \exp\left[\, - \, \frac{1}{2} \, a_\varphi^* \, a_\varphi \, \right] \exp\left[\, a_\varphi \, \hat{a}^\dagger \, \right] \mid 0 \;> \tag{679}$$

Thus the coherent state is an infinite linear superposition of states with different occupation numbers, each coefficient in the linear superposition has a specific phase relation with every other coefficient.

---

[31] R. J. Glauber, Phys. Rev. Lett. **10**, 84 (1963).

The above equation represents a transformation between number operator states and the coherent states. The inverse transformation can be found by expressing a_φ as a magnitude a and a phase φ

$$a_\varphi = a \, \exp\left[\, i \, \varphi \, \right] \qquad (680)$$

The number states can be expressed in terms of the coherent states via the inverse transformation

$$\mid n > \;\; = \;\; \frac{\sqrt{n!}}{a^n} \, \exp\left[\, + \, \frac{1}{2} \, a^2 \, \right] \, \int_0^{2\pi} \frac{d\varphi}{2\pi} \, \exp\left[\, - \, i \, n \, \varphi \, \right] \mid a_\varphi > \qquad (681)$$

by integrating over the phase φ of the coherent state. Since the set of occupation number states is complete, the set of coherent states must also span Hilbert space. In fact, the set of coherent states is over-complete.

The coherent state $\mid a_\varphi >$ can be represented by the point a_φ in the Argand plane. The overlap matrix elements between two coherent states is calculated
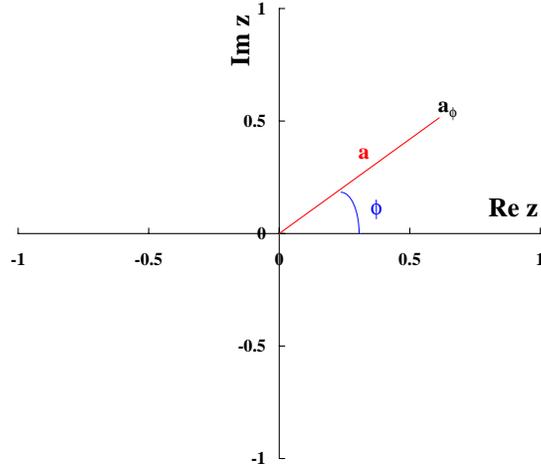
Figure 13: Since a coherent state $\mid a_\phi >$ is completely determined by a complex number a_φ, it can be represented by a point in the complex plane.

as

$$\mid < a'_{\varphi'} \mid a_\varphi > \mid^2 \;\; = \;\; \exp\left[\, - \, \mid a_\varphi \, - \, a'_{\varphi'} \mid^2 \, \right] \qquad (682)$$

Hence, coherent states corresponding to different points are not orthogonal. The coherent states form an over complete basis set. The over completeness relation can be expressed as

$$\int \frac{d \, \Re e \, a_\varphi \, d \, \Im m \, a_\varphi}{\pi} \, \mid a_\varphi > < a_\varphi \mid \;\; = \;\; \hat{I} \qquad (683)$$

110

This relation can be proved by taking the matrix elements between the occupation number states $< n' \mid$ and $\mid n >$, which leads to

$$\int \frac{d \, \Re e \, a_\varphi \, d \, \Im m \, a_\varphi}{\pi} \quad < n' \mid a_\varphi > < a_\varphi \mid n > \; = \; \delta_{n',n} \tag{684}$$

which can be evaluated as

$$= \; \int \frac{d \, \Re e \, a_\varphi \, d \, \Im m \, a_\varphi}{\pi} \quad < n' \mid a_\varphi > < a_\varphi \mid n >$$

$$= \; \int_0^\infty da \; a \int_0^{2\pi} \frac{d\varphi}{\pi} \frac{a_\varphi^{* \, n'} \, a_\varphi^n}{\sqrt{n'! \, n!}} \exp \left[\; - \mid a_\varphi \mid^2 \; \right]$$

$$= \; \int_0^\infty da \; a \frac{a^{n+n'}}{\sqrt{n'! \, n!}} \exp \left[\; - \mid a_\varphi \mid^2 \; \right] \int_0^{2\pi} \frac{d\varphi}{\pi} \exp \left[\; i \, (\, n \; - \; n' \,) \, \varphi \; \right]$$

$$= \; \int_0^\infty da \; a \frac{a^{n+n'}}{\sqrt{n'! \, n!}} \exp \left[\; - \, a^2 \; \right] 2 \, \delta_{n,n'} \tag{685}$$

On changing variable to $s = a^2$, one proves the completeness relation by noting that

$$\int_0^\infty ds \; s^n \; \exp \left[\; - \, s \; \right] \; = \; n! \tag{686}$$

Hence, the coherent states form a complete basis set.

The effect of the creation operator on the coherent state can be expressed as

$$\hat{a}^\dagger \mid a_\varphi > \; = \; \hat{a}^\dagger \exp \left[\; - \; \frac{1}{2} \, a_\varphi^* \, a_\varphi \; \right] \exp \left[\; a_\varphi \, \hat{a}^\dagger \; \right] \mid 0 >$$

$$= \; \exp \left[\; - \; \frac{1}{2} \, a_\varphi^* \, a_\varphi \; \right] \hat{a}^\dagger \exp \left[\; a_\varphi \, \hat{a}^\dagger \; \right] \mid 0 >$$

$$= \; \exp \left[\; - \; \frac{1}{2} \, a_\varphi^* \, a_\varphi \; \right] \frac{\partial}{\partial a_\varphi} \exp \left[\; a_\varphi \, \hat{a}^\dagger \; \right] \mid 0 >$$

$$= \; \exp \left[\; - \; \frac{1}{2} \, a_\varphi^* \, a_\varphi \; \right] \frac{\partial}{\partial a_\varphi} \exp \left[\; + \; \frac{1}{2} \, a_\varphi^* \, a_\varphi \; \right] \mid a_\varphi >$$

$$\tag{687}$$

The coherent state is not an eigenstate of the creation operator, since the resulting state does not include the zero-photon state.

The expectation value of the field operators between the coherent states yields the classical value, since

$$< \; a_\varphi \mid (\; \hat{a}^\dagger \; + \; \hat{a} \;) \mid a_\varphi > \; = \; (\; a_\varphi^* \; + \; a_\varphi \;) \tag{688}$$

In deriving the above equation, the definition

$$\hat{a} \mid a_\varphi > \; = \; a_\varphi \mid a_\varphi > \tag{689}$$

has been used in the term involving the annihilation operator and the term originating from the creation operator is evaluated using the Hermitean conjugate equation

$$< a_\varphi \mid \hat{a}^\dagger = < a_\varphi \mid a_\varphi^* \tag{690}$$

One also finds that that the expectation value of the number operator is given by

$$< a_\varphi \mid \hat{a}^\dagger \, \hat{a} \mid a_\varphi > = a_\varphi^* \, a_\varphi \tag{691}$$

so the magnitude of a_φ is related to the average number of photons in the coherent state $\overline{n}$. This identification is consistent with the Poisson distribution of eqn(676) which governs the probability of finding n photons in the coherent state. The coherent state is not an eigenstate of the number operator since there are fluctuations in any measurement of the number of photons. The rms fluctuation Δn can be evaluated by noting that

$$
\begin{aligned}
< a_\varphi \mid \hat{n}^2 \mid a_\varphi > \ &= \ < a_\varphi \mid \hat{a}^\dagger \, \hat{a} \, \hat{a}^\dagger \, \hat{a} \mid a_\varphi > \\
&= \ < a_\varphi \mid \hat{a}^\dagger \, \hat{a}^\dagger \, \hat{a} \, \hat{a} \mid a_\varphi > \ + \ < a_\varphi \mid \hat{a}^\dagger \, \hat{a} \mid a_\varphi > \\
&= \ (a_\varphi^*)^2 \, (a_\varphi)^2 \ + \ a_\varphi^* \, a_\varphi
\end{aligned}
\tag{692}
$$

where the boson commutation relations have been used in the second line. Thus, the mean squared fluctuation in the number operator is given by

$$< a_\varphi \mid \Delta \hat{n}^2 \mid a_\varphi > \ = \ a_\varphi^* \, a_\varphi \tag{693}$$

The rms fluctuation of the photon number is only negligible when compared to the average value if a_φ has a large magnitude

$$a_\varphi^* \, a_\varphi \ \gg \ 1 \tag{694}$$

The expectation values of coherent states almost behave completely classically. The deviation from the classical expectation values can be seen by examining

$$< a_\varphi \mid \hat{a} \, \hat{a}^\dagger \mid a_\varphi > = a_\varphi^* \, a_\varphi \ + \ 1 \tag{695}$$

which is evaluated by using the commutation relations. It is seen that the expectation values can be approximated by the classical values, if the magnitude of a_φ is much greater than unity.

Exercise:

Determine the expectation values for the electric and magnetic field operators in a coherent state which represents a plane-polarized electromagnetic wave.

Exercise:

Determine the expectation values for the electric and magnetic field operators in a coherent state which represents a left circularly-polarized electromagnetic wave composed of photons with a helicity of $+1$.

10.4.1   The Phase-Number Uncertainty Relation

From the discussion of coherent states, it is seen that the coherent state has a definite phase, but does not have a definite number of quanta. In general, it is impossible to know both the phase of a state and the number of a state. This is formalized as a phase - number uncertainty relation.

The phase and amplitude of a state is related to the annihilation operator. Since the annihilation operator is non-Hermitean, one can construct the annihilation operator as a function of Hermitean operators. Formally, the amplitude can be related to the square root operator, and the phase to a phase operator[32]. Hence, one can write

$$\hat{a}_{\underline{k},\alpha} \;=\; \exp\left[\; + \; i \; (\hat{\varphi}_{\underline{k},\alpha} - \omega_k \; t) \;\right] \; \sqrt{\hat{n}_{\underline{k},\alpha}} \tag{696}$$

and the Hermitean conjugate operator, the creation operator can be expressed as

$$\hat{a}^{\dagger}_{\underline{k},\alpha} \;=\; \sqrt{\hat{n}_{\underline{k},\alpha}} \; \exp\left[\; - \; i \; (\hat{\varphi}_{\underline{k},\alpha} - \omega_k \; t) \;\right] \tag{697}$$

since it has been required that $\sqrt{\hat{n}}$ and $\hat{\varphi}$ are Hermitean. Furthermore, the operator $\sqrt{\hat{n}}$ must have the property

$$\sqrt{\hat{n}_{\underline{k},\alpha}} \; \sqrt{\hat{n}_{\underline{k},\alpha}} \;=\; \hat{n}_{\underline{k},\alpha} \tag{698}$$

On substituting the expressions for the creation and annihilation operators, in terms of the phase and amplitude, into boson commutation relations

$$[\; \hat{a}_{\underline{k},\alpha} \; , \; \hat{a}^{\dagger}_{\underline{k}',\beta} \;] \;=\; \delta_{\underline{k},\underline{k}'} \; \delta_{\alpha,\beta} \tag{699}$$

etc, one finds

$$\begin{aligned}
\delta_{\underline{k},\underline{k}'} \; \delta_{\alpha,\beta} \;=\; & \exp\left[\; + \; i \; (\hat{\varphi}_{\underline{k},\alpha} - \omega_k \; t) \;\right] \; \sqrt{\hat{n}_{\underline{k},\alpha}} \; \sqrt{\hat{n}_{\underline{k}',\beta}} \; \exp\left[\; - \; i \; (\hat{\varphi}_{\underline{k}',\beta} - \omega_{k'} \; t) \;\right] \\
& - \; \sqrt{\hat{n}_{\underline{k}',\beta}} \; \exp\left[\; - \; i \; (\hat{\varphi}_{\underline{k}',\beta} - \omega_{k'} \; t) \;\right] \; \exp\left[\; + \; i \; (\hat{\varphi}_{\underline{k},\alpha} - \omega_k \; t) \;\right] \; \sqrt{\hat{n}_{\underline{k},\alpha}}
\end{aligned} \tag{700}$$

Thus, for $\underline{k} = \underline{k}'$ and $\alpha = \beta$, one has

$$\exp\left[\; + \; i \; \varphi_{\underline{k},\alpha} \;\right] \; \hat{n}_{\underline{k},\alpha} \; - \; \hat{n}_{\underline{k},\alpha} \; \exp\left[\; + \; i \; \varphi_{\underline{k},\alpha} \;\right] \;=\; \exp\left[\; + \; i \; \varphi_{\underline{k},\alpha} \;\right] \tag{701}$$

---

[32]P. A. M. Dirac, Proc. Roy. Soc. **A 114**, 243 (1927).

This relationship is satisfied, if the phase and number operators satisfy the commutation relation

$$[\ \hat{n}_{\underline{k},\alpha}\ ,\ \varphi_{\underline{k},\alpha}\ ]\ =\ i \tag{702}$$

If one can construct the Hermitean operators that satisfy this commutation relation, then one can show that the rms uncertainties phase and number must satisfy the inequality

$$(\Delta\varphi_{\underline{k},\alpha})_{rms}\ (\Delta n_{\underline{k},\alpha})_{rms}\ \geq\ 1 \tag{703}$$

It should be noted that only the relative phase can be measured[33]. Thus, if the phase difference of any two components $(\underline{k},\alpha)$ and $(\underline{k}',\alpha')$ is specified precisely, then the occupation number of either component can not be specified.

Exercise:

Express the vector potential and the electric and magnetic field operators in terms of the amplitude and phase operators.

10.4.2   Argand Representation of Coherent States

The coherent state $\mid a_{\varphi}\ >$ can be represented by the point a_{φ} in the Argand plane. The overlap matrix elements between two coherent states is calculated as

$$\mid\ <\ a'_{\varphi'}\ \mid\ a_{\varphi}\ >\ \mid^2\ =\ \exp\left[\ -\ \mid a_{\varphi}\ -\ a'_{\varphi'}\ \mid^2\ \right] \tag{704}$$

Hence, coherent states are not orthogonal. In fact, their overlap decreases exponentially with large "separations" between the points a_{φ} and $a'_{\varphi'}$ in the Argand plane. We shall denote $\mid a_{\varphi}\mid$ by a. Two states separated by distances $a\ \Delta\varphi$ or Δa such that $a\ \Delta\varphi \geq 1$ and $\Delta a \geq 1$ are effectively orthogonal or independent. However, states within an area given by $\Delta a \times a\ \Delta\varphi \approx 1$ have significant overlap and so can represent the same state. Therefore, the minimum uncertainty state occupies an area $\Delta a \times a\ \Delta\varphi \approx 1$. We note that $2\ a\ \Delta a$ can be interpreted as a measure of the uncertainty Δn_{φ} in the particle number for the state, and $\Delta\varphi$ is the uncertainty in the phase of the state. Hence, the phase - number uncertainty relation sets the area of the Argand diagram that can be associated with a single state as

$$a\ \Delta a\ \Delta\varphi\ \sim\ 1 \tag{705}$$

---

[33]L. Susskind and J. Glogower, Physics **1**, 49 (1964).
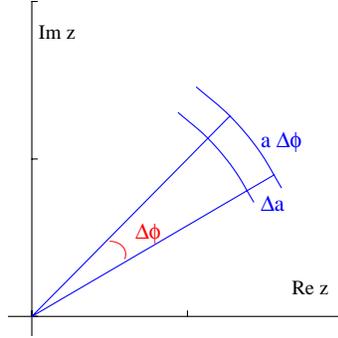
Figure 14: Due to the phase-number uncertainty principle, the minimum area of the Argand diagram needed to represent a minimum uncertainty state has dimensions such that $a \, \Delta a \, \Delta\varphi \sim 1$.

11   Non-Relativistic Quantum Electrodynamics

The non-relativistic Hamiltonian for a particle with charge q and mass m interacting with a quantized electromagnetic field can be expressed as

$$\hat{H} \;=\; \frac{\hat{\underline{p}}^2}{2\,m} \;+\; q\,\phi(\underline{r}) \;-\; \frac{q}{2\,m\,c}\left(\hat{\underline{p}}\cdot\hat{\underline{A}}(\underline{r}) \;+\; \hat{\underline{A}}(\underline{r})\cdot\hat{\underline{p}}\right) \;+\; \frac{q^2}{2\,m\,c^2}\,\hat{\underline{A}}^2(\underline{r})$$

$$+\; \int \, d^3\underline{r}' \left(\frac{\hat{\underline{E}}^2(\underline{r}') \;+\; \hat{\underline{B}}^2(\underline{r}')}{8\,\pi}\right) \tag{706}$$

when the vector potential is chosen to satisfy the Coulomb gauge. The second, third and fourth terms are to be evaluated at the location of the charged point particle, $\underline{r}$, and the last term is evaluated at all points in space. The Hamiltonian can be expressed as

$$\hat{H} \;=\; \hat{H}_0 \;+\; \hat{H}_{rad} \;+\; \hat{H}_{int} \tag{707}$$

where $\hat{H}_0$ is the Hamiltonian for the charged particle in the electrostatic potential ϕ

$$\hat{H}_0 \;=\; \frac{\hat{p}^2}{2\,m} \;+\; q\,\phi(\underline{r}) \tag{708}$$

and $\hat{H}_{rad}$ is the Hamiltonian for the electromagnetic radiation and H_{int} is the interaction

$$\hat{H}_{int} \;=\; -\,\frac{q}{2\,m\,c}\left(\hat{\underline{p}}\cdot\hat{\underline{A}} \;+\; \hat{\underline{A}}\cdot\hat{\underline{p}}\right) \;+\; \frac{q^2}{2\,m\,c^2}\,\hat{\underline{A}}^2 \tag{709}$$

The interaction term is composed of a paramagnetic interaction which is linearly proportional to the vector potential and the diamagnetic interaction which is proportional to the square of the vector potential. When the electromagnetic

115

field is quantized, the radiation Hamiltonian has the form

$$\hat{H}_{rad} \;=\; \sum_{\underline{k},\alpha} \frac{\hbar\,\omega_k}{2} \left(\; \hat{a}^{\dagger}_{\underline{k},\alpha}\,\hat{a}_{\underline{k},\alpha} \;+\; \hat{a}_{\underline{k},\alpha}\,\hat{a}^{\dagger}_{\underline{k},\alpha} \;\right) \tag{710}$$

Since the quantized vector potential is given by

$$\underline{\hat{A}}(\underline{r},t) \;=\; \frac{1}{\sqrt{V}} \sum_{\underline{k},\alpha} \sqrt{\frac{2\,\pi\,\hbar\,c^2}{\omega_k}} \, \hat{e}_{\alpha}(\underline{k}) \left(\; \hat{a}^{\dagger}_{\underline{k},\alpha} + \hat{a}_{-\underline{k},\alpha} \;\right) \exp\left[\,-\,i\,\underline{k}\cdot\underline{r}\,\right] \tag{711}$$

the paramagnetic interaction can be expressed as

$$\hat{H}_{para} \;=\; -\,\frac{q}{m\,c} \sum_{\underline{k},\alpha} \sqrt{\frac{2\,\pi\,\hbar\,c^2}{V\,\omega_k}} \, \hat{\underline{p}}\cdot\hat{\epsilon}_{\alpha}(\underline{k}) \left(\; \hat{a}^{\dagger}_{\underline{k},\alpha} + \hat{a}_{-\underline{k},\alpha} \;\right) \exp\left[\,-\,i\,\underline{k}\cdot\underline{r}\,\right]$$
$$\tag{712}$$

in which the transverse gauge condition $\underline{\nabla}\cdot\underline{A} \;=\; 0$ has also been used. The
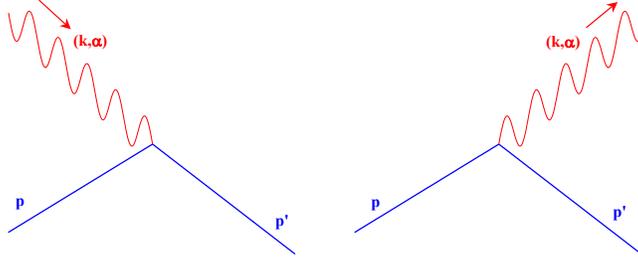
Figure 15: The paramagnetic interaction leads to scattering of an electron from $\underline{p}$ to $\underline{p}'$ by either (a) absorbing a photon, or (b) by emitting a photon.

diamagnetic interaction is expressed as

$$\hat{H}_{dia} \;=\; \frac{q^2}{2\,m\,c^2} \sum_{\underline{k},\underline{k}',\alpha,\beta} \left(\frac{2\,\pi\,\hbar\,c^2}{\sqrt{\omega_{\underline{k}}\,\omega_{\underline{k}'}}\,V}\right) \hat{\epsilon}_{\beta}(\underline{k}')\cdot\hat{\epsilon}_{\alpha}(\underline{k}) \, \exp\left[\,-\,i\,(\,\underline{k}\,+\,\underline{k}'\,)\cdot\underline{r}\,\right]$$
$$\times \left(\; \hat{a}^{\dagger}_{\underline{k}',\beta}\,\hat{a}^{\dagger}_{\underline{k},\alpha} \;+\; \hat{a}^{\dagger}_{\underline{k}',\beta}\,\hat{a}_{-\underline{k},\alpha} \;+\; \hat{a}_{-\underline{k}',\beta}\,\hat{a}^{\dagger}_{\underline{k},\alpha} \;+\; \hat{a}_{-\underline{k}',\beta}\,\hat{a}_{-\underline{k},\alpha} \;\right) \tag{713}$$

For charged particles with spin one-half, analysis of the non-relativistic Pauli equation shows that there is another interaction term involving the particles' spins. This interaction can be described by the anomalous Zeeman interaction

$$\hat{H}_{Zeeman} \;=\; -\,\frac{q\,\hbar}{2\,m\,c} \left(\; \underline{\sigma}\cdot\underline{B}\;\right) \tag{714}$$

where

$$\underline{B} \;=\; \underline{\nabla} \wedge \underline{A}(\underline{r}) \tag{715}$$

and σ_i are the three Pauli matrices.

Generally, the paramagnetic interaction has a greater strength than the Zeeman interaction. This can be seen by examining the magnitudes of the interactions. The paramagnetic interaction has a magnitude given by

$$\frac{e}{m \, c} \, \underline{p} \cdot \underline{A} \tag{716}$$

and for an atom of size a , the uncertainty principle yields

$$p \sim \frac{\hbar}{a} \tag{717}$$

The Zeeman interaction has a magnitude given by

$$\frac{e \, \hbar}{m \, c} \, \underline{\sigma} \cdot (\, \underline{k} \, \wedge \, \underline{A} \,) \tag{718}$$

but since k is the wavelength of light

$$k \sim \frac{1}{\lambda} \tag{719}$$

Hence, since the wave length of light is larger than the linear dimension of an atom, $\lambda > a$, one finds the inequality between the magnitude of the paramagnetic interaction and the Zeeman interaction

$$\frac{e \, \hbar}{m \, c} \, \frac{1}{a} \, A \, > \, \frac{e \, \hbar}{m \, c} \, \frac{1}{\lambda} \, A \tag{720}$$

Both the paramagnetic and Zeeman coupling strengths are proportional to the magnitude of the vector potential A, hence the ratio of the strengths of the interactions are independent of A. Therefore, there magnitudes satisfy the inequality

$$\frac{1}{a} \, > \, \frac{1}{\lambda} \tag{721}$$

so the Zeeman interaction can frequently be neglected in comparison with the paramagnetic interaction.

11.1   Emission and Absorption of Photons

An atom in an electromagnetic field has its constituent charges perturbed by the oscillating field, and those perturbations may lead to either to the absorption of radiation or emission of further radiation. However, thermal equilibrium between matter and radiation can only be reached if, in addition to these induced processes, there exists also a spontaneous process in which an excited atom emits radiation even in the absence of any measurable radiation. This spontaneous emission process may be considered as being induced by the zero-point fluctuations of the electromagnetic field.

11.1.1 The Emission of Radiation

We shall consider a state $| (nlm) \{n_{\underline{k}',\beta}\} >$ which is an energy eigenstate of the unperturbed Hamiltonian $\hat{H}_0$ and the radiation Hamiltonian $\hat{H}_{rad}$. The interaction $\hat{H}_{int}$ causes the system to make a transition from the initial state to a final state. In the initial state, the electron is in an energy state designated by the quantum numbers (n, l, m) and the electromagnetic field is in a state specified by the number of photons in each normal mode. That is, the photon field is in an initial state which is specified by the set of photon quantum numbers, $\{n_{\underline{k}',\beta'}\}$. We shall consider the transition in which the electron makes a transition from the initial state to a final state denoted by (n', l', m'). Since the
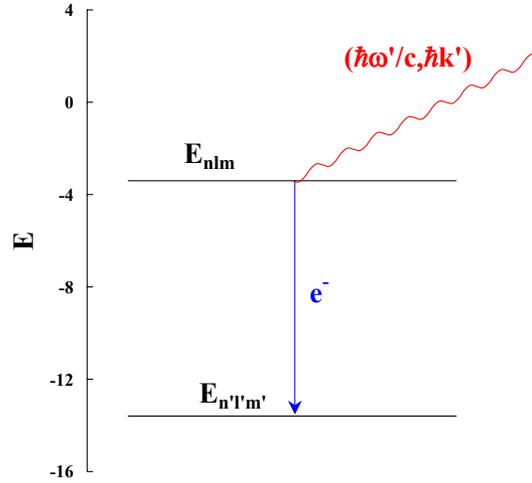
Figure 16: An electron in the initial atomic state with energy $E_{n,l,m}$ makes a transition to the final atomic state with energy $E_{n',l',m'}$, by emitting a photon with energy $\hbar\omega_{k'}$.

photon is emitted, the final state of the photon field described by the set $\{n'_{\underline{k}',\beta}\}$ where

$$n'_{\underline{k}',\beta} \;=\; n_{\underline{k}',\beta} \qquad \text{for} \quad (\underline{k}',\beta) \;\neq\; (\underline{k},\alpha) \tag{722}$$

and the number of photons in a normal mode $(\underline{k},\alpha)$ is increased by one

$$n'_{\underline{k},\alpha} \;=\; n_{\underline{k},\alpha} \;+\; 1 \tag{723}$$

The transition rate for the electron to make a transition from (n, l, m) to (n', l', m') can be calculated[34] from the Fermi-Golden rule expression

$$\left(\frac{1}{\tau}\right) = \left(\frac{2\,\pi}{\hbar}\right) \sum_{\underline{k},\alpha} | \; < n'l'm' \, \{n'_{\underline{k}',\beta}\} \, | \, \hat{H}_{int} \, | \, nlm \, \{n_{\underline{k}',\beta}\} \; > \; |^2 \, \delta(\, E_{nlm} - E_{n'l'm'} - \hbar\,\omega_{\underline{k},\alpha} \,) \tag{724}$$

---

[34]P. A. M. Dirac, Proc. Roy. Soc. **A 112**, 661 (1926), **A 114**, 243 (1927).

The delta function expresses the conservation of energy. The energy of the initial state is given by

$$
E_{nlm} + \sum_{\underline{k}',\beta} \hbar\,\omega_{\underline{k}',\beta}\,(\,n_{\underline{k}',\beta} + \frac{1}{2}\,) \tag{725}
$$

and the final state has energy

$$
E_{n'l'm'} + \sum_{\underline{k}',\beta} \hbar\,\omega_{\underline{k}',\beta}\,(\,n'_{\underline{k}',\beta} + \frac{1}{2}\,) \tag{726}
$$

The difference in the energy of the initial state and final state is evaluated as

$$
E_{nlm} - E_{n'l'm'} - \hbar\,\omega_{\underline{k},\alpha} \tag{727}
$$

which is the argument of the delta function and must vanish if energy is conserved. The sum over $\underline{k}$ can be evaluated by assuming that the radiation field is confined to a volume V. The allowed $\underline{k}$ values for the normal modes are determined by the boundary conditions. In this case, the sum over $\underline{k}$ is transformed to an integral over $\underline{k}$-space via

$$
\sum_{\underline{k}} \rightarrow \frac{V}{(\,2\,\pi\,)^3} \int d^3\underline{k} \tag{728}
$$

The matrix elements of the interaction Hamiltonian between photon energy eigenstates is evaluated as

$$
< \{n'_{\underline{k}',\beta}\} \mid \hat{H}_{int} \mid \{n_{\underline{k}',\beta}\} > \quad = \quad -\frac{q}{m\,c}\sum_{\underline{k},\alpha}\sqrt{\frac{2\,\pi\,\hbar\,c^2}{V\,\omega_k}}
$$

$$
\times \quad < \{n'_{\underline{k}',\beta}\} \mid \hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\underline{p}}\,(\,\hat{a}^\dagger_{\underline{k},\alpha} + \hat{a}_{-\underline{k},\alpha}\,)\,\exp\left[\,-\,i\,\underline{k}\cdot\underline{r}\,\right] \mid \{n_{\underline{k}',\beta}\} > \tag{729}
$$

since only the paramagnetic part of the interaction has non-zero matrix elements. For the photon emission process, the matrix elements of the creation operator between the initial and final states of the electromagnetic cavity is evaluated as

$$
< \{n'_{\underline{k}',\beta}\} \mid \hat{a}^\dagger_{\underline{k},\alpha} \mid \{n_{\underline{k}',\beta}\} > = \sqrt{n_{\underline{k},\alpha} + 1} \tag{730}
$$

hence, the matrix elements of the interaction are given by

$$
< n'l'm'\,\{n'_{\underline{k}',\beta}\} \mid \hat{H}_{int} \mid nlm\,\{n_{\underline{k}',\beta}\} > \quad = \quad -\frac{q}{m\,c}\sqrt{\frac{2\,\pi\,\hbar\,c^2}{V\,\omega_k}}\,\sqrt{n_{\underline{k},\alpha} + 1}
$$

$$
\times \quad < n'l'm' \mid \hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\underline{p}}\,\exp\left[\,-\,i\,\underline{k}\cdot\underline{r}\,\right] \mid nlm > \tag{731}
$$

Therefore, the transition rate for photon emission can be expressed as

$$\frac{1}{\tau} = \frac{2\pi}{\hbar} \left(\frac{q}{mc}\right)^2 \frac{V}{(2\pi)^3} \int d^3\underline{k} \left(\frac{2\pi\hbar c^2}{V\omega_k}\right) \sum_\alpha (n_{\underline{k},\alpha} + 1)$$

$$\times \; | < n'l'm' | \hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\underline{p}} \; \exp\left[-i\underline{k}\cdot\underline{r} \right] | nlm > |^2 \; \delta(E_{nlm} - E_{n'l'm'} - \hbar\omega_k)$$

(732)

The above expression shows that the rate for emitting a photon into state $(\underline{k}, \alpha)$ is proportional to a factor of $n_{\underline{k},\alpha} + 1$, which depends on the state of occupation of the normal mode. The term proportional to the photon occupation number describes stimulated emission. However, if there are no photons initially present in this normal mode, one still has a non-zero transition rate corresponding to spontaneous emission. These factors are the result of the rigorous calculations[35] based on Dirac's quantization of the electromagnetic field, but were previously derived by Einstein[36] using a different argument. From the above expression, it is seen that the number of photons emitted into state $(\underline{k}, \alpha)$ increases in proportional to the number of photons present in that normal mode. This stimulated emission increases the number of photons and can lead to amplification of the number of quanta in the normal mode, and leads to the phenomenon of Light Amplification by Stimulated Emission of Radiation (LASER).

11.1.2 The Dipole Approximation

The dipole approximation is justified by noting that in an emission process, the typical energy of the photon is of the order of 10 eV. Hence, a typical wave length of the photon is given by

$$\lambda = \frac{2\pi c\hbar}{\hbar\omega} \sim 3000 \; \mathring{A}$$

(733)

whereas the typical length scale r for the electronic state is of the order of an Angstrom. Therefore the product $k\,r \sim 10^{-3}$, so the exponential factor in the vector potential can be Taylor expanded as

$$\exp\left[-i\underline{k}\cdot\underline{r} \right] \sim 1 - i\underline{k}\cdot\underline{r} + \dots$$

(734)

The first term in the expression produces results that are equivalent to the radiation from an oscillating classical electric dipole. If only the first term in the expansion is retained, the resulting approximation is known as the dipole approximation. The second term in the expansion yields results equivalent to the radiation from an electric quadrupole. The dipole approximation, where only the first term in the expansion is retained, is justified for transitions where the

[35]P. A. M. Dirac, Proc. Roy. Soc. **A 114**, 243 (1927).
[36]A. Einstein, Verh. Deutsche Phys. Ges. **18**, 318 (1916), Phys. Z. **18**, 121 (1917).
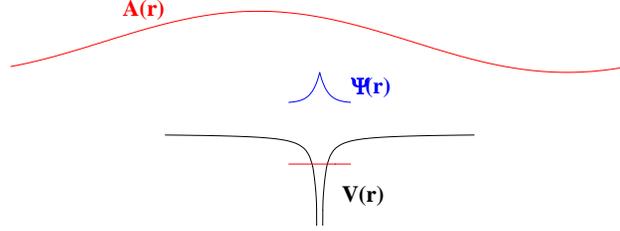
Figure 17: A cartoon depicting the relative length-scales assumed in the dipole approximation.

successive terms in the expansion are successively smaller by factors of the order of 10^{-3}. The dipole approximation crudely restricts consideration to the case where the emitted photon can only have zero orbital angular momentum. This follows from the dipole approximation's requirement that the size of the atom is negligible compared with the scale over which the vector potential varies. Then the vector potential in the spatial region where the electron is located only describes photons with zero orbital angular momentum.

In the dipole approximation, the transition rate for single photon emission is given by

$$
\frac{1}{\tau} \approx \frac{2\,\pi}{\hbar} \left(\frac{q}{m\,c} \right)^2 \frac{V}{(\,2\,\pi\,)^3} \int d^3\underline{k} \left(\frac{2\,\pi\,\hbar\,c^2}{V\,\omega_k} \right) \sum_\alpha (\,n_{\underline{k},\alpha}\,+\,1\,)
$$
$$
\times\,|\,\hat{\epsilon}_\alpha(\underline{k})\,.\, < \,n'l'm'\,|\,\hat{\underline{p}}\,|\,nlm\, > \,|^2\,\delta(\,E_{nlm}\,-\,E_{n'l'm'}\,-\,\hbar\,\omega_k\,)
$$

$$(735)$$

The matrix elements of the momentum can be evaluated by noting that the states $|\,nlm\,>$ are eigenstates of the unperturbed electronic Hamiltonian so

$$
\hat{H}_0\,|\,nlm\,> \,=\,E_{nlm}\,|\,nlm\,> \tag{736}
$$

where the unperturbed Hamiltonian is given by

$$
\hat{H}_0\,=\,\frac{\hat{\underline{p}}^2}{2\,m}\,+\,V(r) \tag{737}
$$

The electronic momentum operator $\hat{\underline{p}}$ can be expressed in terms of the commutator of the Hamiltonian $\hat{H}_0$ and $\underline{r}$ through the relation

$$
[\,\underline{r}\,,\,\hat{H}_0\,]\,=\,i\,\frac{\hbar}{m}\,\hat{\underline{p}} \tag{738}
$$

On using this relation, the matrix elements of the momentum operator can be written in terms of the matrix elements of the electron's position operator $\underline{r}$ by

$$
<\,n'l'm'\,|\,\hat{\underline{p}}\,|\,nlm\,> \quad=\quad -\,i\,\frac{m}{\hbar}\quad <\,n'l'm'\,|\,[\,\underline{r}\,,\,\hat{H}_0\,]\,|\,nlm\,>
$$

$$
= i \frac{m}{\hbar} (E_{nlm} - E_{n'l'm'}) < n'l'm' \mid \underline{r} \mid nlm >
$$

$$(739)$$

Therefore, in the dipole approximation, the transition rate is given by

$$
\frac{1}{\tau} \approx \frac{q^2}{(2\pi)} \int d^3\underline{k} \, \omega_k \sum_{\alpha} (n_{\underline{k},\alpha} + 1)
$$
$$
\times \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, < n'l'm' \mid \underline{r} \mid nlm > \mid^2 \delta(E_{nlm} - E_{n'l'm'} - \hbar\,\omega_k)
$$

$$(740)$$

where the property of the delta function has been used to set

$$
\frac{(E_{nlm} - E_{n'l'm'})^2}{\hbar^2} \delta(E_{nlm} - E_{n'l'm'} - \hbar\,\omega_k) = \omega_k^2 \, \delta(E_{nlm} - E_{n'l'm'} - \hbar\,\omega_k)
$$

$$(741)$$

It is seen that the volume of the electromagnetic cavity has dropped out of the expression of eqn(740) for the transition rate. We shall assume that the number of photons $n_{\underline{k},\alpha}$ in the initial state is zero. The (complex) factor

$$
\underline{d}_{nlm,n'l'm'} = q < n'l'm' \mid \underline{r} \mid nlm > \tag{742}
$$

is defined as the electric dipole moment, and the electronic energy difference is denoted by the frequency

$$
E_{nlm} - E_{n'l'm'} = \hbar\,\omega_{nl,n'l'} \tag{743}
$$

With this notation, the transition rate can be expressed as

$$
\frac{1}{\tau} \approx \frac{1}{2\pi} \int d^3\underline{k} \left(\frac{\omega_{nl,n'l'}^2}{\omega_k} \right) \sum_{\alpha} \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, \underline{d}_{nlm,n'l'm'} \mid^2 \delta(\hbar\,\omega_{nl,n'l'} - \hbar\,\omega_k)
$$

$$(744)$$

The integration over $d^3\underline{k}$ can be performed by separating the integration over the direction $d\Omega_k$ of the outgoing photon and an integration over the magnitude of $\underline{k}$. The integration over the magnitude of $\underline{k}$ can be performed by noting that the integrand is proportional to a Dirac delta function, so the transition rate can be evaluated as

$$
\frac{1}{\tau} = \frac{\omega_{nl,n'l'}^2}{2\pi\,\hbar} \int d\Omega_k \int_0^\infty dk \left(\frac{k^2}{\omega_k} \right) \sum_{\alpha} \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, \underline{d}_{nlm,n'l'm'} \mid^2 \delta(\omega_{nl,n'l'} - c\,k)
$$

$$
= \frac{\omega_{nl,n'l'}^3}{2\pi\,\hbar\,c^3} \int d\Omega_k \sum_{\alpha} \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, \underline{d}_{nlm,n'l'm'} \mid^2 \tag{745}
$$

The above expression yields the rate at which an electron makes a transition between the initial and final electronic state, in which one photon of any polarization is emitted in any direction.

If one is only interested in the decay rate of the electronic state via the emission of a photon, one should sum over all polarizations and integrate over all directions of the emitted photon. The direction of the emitted photon $\hat{k}$ is expressed in terms of polar coordinates defined with respect to an arbitrarily chosen polar axis. The direction of the photon's wave vector $\hat{k}$ is defined as
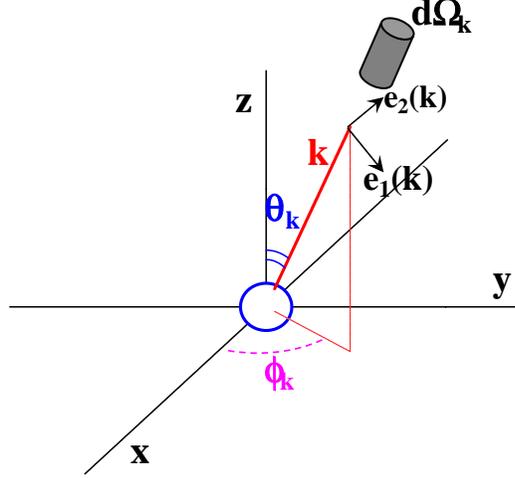
Figure 18: A photon is emitted with wave vector $\underline{k}$ with a direction denoted by the polar coordinates (θ_k, φ_k). The polarization vector $\hat{e}_1(\underline{k})$ is chosen to be in the plane containing the polar-axis and $\underline{k}$, therefore, $\hat{e}_2(\underline{k})$ is parallel to the $x - y$ plane.

$\hat{k} = (\sin\theta_k \cos\varphi_k, \sin\theta_k \sin\varphi_k, \cos\theta_k)$. The directions of the two transverse polarizations α are defined as

$$
\begin{aligned}
\hat{\epsilon}_1(\underline{k}) &= (\cos\theta_k \cos\varphi_k, \cos\theta_k \sin\varphi_k, -\sin\theta_k) \\
\hat{\epsilon}_2(\underline{k}) &= (-\sin\varphi_k, \cos\varphi_k, 0)
\end{aligned}
\tag{746}
$$

The scalar product between the polarization vectors and the dipole moment can be expressed in terms of the Cartesian components via

$$
\hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{d}_{nlm,n'l'm'} = \sum_i \hat{\epsilon}_\alpha^{(i)}(\underline{k}) \cdot (\, \underline{d}_{nlm,n'l'm'}^{(i)} \,)
\tag{747}
$$

As neither the polarization nor the direction of the outgoing photon are measured, the transition rates is determined as an integral over all directions

$$
\left(\frac{1}{\tau} \right) = \frac{\omega_{nl,n'l'}^3}{2\,\pi\,\hbar\,c^3} \sum_{i,j} \sum_\alpha \int d\Omega_k \; \hat{\epsilon}_\alpha^{(j)}(\underline{k}) \; \hat{\epsilon}_\alpha^{(i)}(\underline{k}) \; (\, \underline{d}_{nlm,n'l'm'}^{(j)} \,)^* \; (\, \underline{d}_{nlm,n'l'm'}^{(i)} \,)
\tag{748}
$$

123

On using the identity

$$\frac{1}{4\pi} \sum_\alpha \int d\Omega_k \; \hat{\epsilon}^{(j)}_\alpha(\underline{k}) \; \hat{\epsilon}^{(i)}_\alpha(\underline{k}) \; = \; \frac{2}{3} \; \delta_{i,j} \tag{749}$$

one finds that the transition rate is given by the scalar product of complex vectors

$$\left(\frac{1}{\tau} \right) \; = \; \frac{4 \; \omega^3_{nl,n'l'} \; \underline{d}^*_{nlm,n'l'm'} \cdot \underline{d}_{nlm,n'l'm'}}{3 \; \hbar \; c^3} \tag{750}$$

The electric dipole matrix elements can be shown to vanish between most pairs of states. The selection rules determine which matrix elements are non-zero and, therefore, which electric dipole transitions are allowed.

11.1.3 Electric Dipole Radiation Selection Rules

Electric dipole induced transitions obey the selection rules $\Delta l = \pm 1$ and either $\Delta m = \pm 1$ or 0, where l is the quantum number for electron's orbital angular momentum and m is the z-component. The dipole selection rules can be derived by writing the wave functions for the one-electron states as

$$\psi_{n,l,m}(\underline{r}) \; = \; R_{nl}(r) \; Y^l_m(\theta, \varphi) \tag{751}$$

where $R_{n,l}(r)$ is the radial wave function, and $Y^l_m(\theta, \varphi)$ is the spherical harmonic function quantized along the z-direction. The components of an arbitrarily oriented electric dipole matrix elements involve matrix elements of the quantities

$$\begin{aligned} x &= r \; \sin\theta \; \cos\varphi \\ y &= r \; \sin\theta \; \sin\varphi \\ z &= r \; \cos\theta \end{aligned} \tag{752}$$

Since the above expressions are the components of a vector, they can be re-written as combinations of the spherical harmonics with angular momentum $l = 1$, via

$$\begin{aligned} x &= r \; \frac{1}{2} \sqrt{\frac{8\pi}{3}} \left(Y^1_{-1}(\theta, \varphi) - Y^1_1(\theta, \varphi) \right) \\ y &= r \; \frac{i}{2} \sqrt{\frac{8\pi}{3}} \left(Y^1_{-1}(\theta, \varphi) + Y^1_1(\theta, \varphi) \right) \\ z &= r \; \sqrt{\frac{4\pi}{3}} \; Y^1_0(\theta, \varphi) \end{aligned} \tag{753}$$

Hence, the components of the vector $\underline{r}$ can be written as

$$\underline{r} \; = \; \sqrt{\frac{4\pi}{3}} \; r \left[\frac{\hat{e}_x + i \, \hat{e}_y}{\sqrt{2}} \; Y^1_{-1} \; + \; \hat{e}_z \, Y^1_0 \; - \; \frac{\hat{e}_x - i \, \hat{e}_y}{\sqrt{2}} \; Y^1_1 \right] \tag{754}$$

124

The circular polarization vectors are given by

$$
\begin{aligned}
\hat{e}_{m=-1} &= \frac{\hat{e}_x - i\,\hat{e}_y}{\sqrt{2}} \\
\hat{e}_{m=0} &= \hat{e}_z \\
\hat{e}_{m=+1} &= -\frac{\hat{e}_x + i\,\hat{e}_y}{\sqrt{2}}
\end{aligned}
\tag{755}
$$

which are orthogonal

$$
\hat{e}_{m'}^* \cdot \hat{e}_m = \delta_{m,m'}
\tag{756}
$$

Hence, the vector $\underline{r}$ can be written in the alternate forms

$$
\begin{aligned}
\underline{r} &= \sqrt{\frac{4\pi}{3}}\, r \sum_m \hat{e}_m^*\, Y_m^1(\theta,\varphi) \\
&= \sqrt{\frac{4\pi}{3}}\, r \sum_m \hat{e}_m\, Y_m^1(\theta,\varphi)^*
\end{aligned}
\tag{757}
$$

This illustrates that through the dipole approximation coupling term

$$
\underline{r} \cdot \left(\hat{e}_m\, \hat{a}_{\underline{k},m} + \hat{e}_m^*\, \hat{a}_{\underline{k},m}^\dagger \right)
\tag{758}
$$

(where $k \approx 0$), an electron with angular momentum quantized along the z-direction most naturally couples to circularly-polarized light with the same quantization axis. The electric dipole matrix elements involve the three factors

$$
\int_0^{2\pi} d\varphi \int_0^\pi d\theta\, \sin\theta\, Y_{m'}^{l'}(\theta,\varphi)^*\, Y_{\pm 1}^1(\theta,\varphi)\, Y_m^l(\theta,\varphi)
$$

$$
\int_0^{2\pi} d\varphi \int_0^\pi d\theta\, \sin\theta\, Y_{m'}^{l'}(\theta,\varphi)^*\, Y_0^1(\theta,\varphi)\, Y_m^l(\theta,\varphi)
\tag{759}
$$

which come from the angular integrations. Conservation of angular momentum leads to the dipole-transition selection rules

$$
l' = l \pm 1
\tag{760}
$$

and

$$
\begin{aligned}
m' &= m \pm 1 \\
m' &= m
\end{aligned}
\tag{761}
$$

because one unit of angular momentum is carried away by the photon in the form of its spin[37].

---

[37] In the dipole approximation, the photon is restricted to have zero orbital angular momentum. Therefore, the angular momentum is completely transformed to the photon's spin. More generally, the spatial (plane-wave) part of the vector potential should be expanded in terms of spherical harmonics to exhibit the photon's orbital angular momentum components.

The m-selection rules for electric dipole transitions.

The selection rules on the z-component of the angular momentum, m, follow directly from the φ-dependence of the spherical harmonics

$$Y_m^l(\theta, \varphi) \ = \ \Theta_m^l(\theta) \ \frac{1}{\sqrt{2 \pi}} \ \exp \left[i \, m \, \varphi \right] \tag{762}$$

so, the integral over the Cartesian components of the dipole matrix elements involve

$$\frac{1}{2 \pi} \int_0^{2\pi} d\varphi \ \exp \left[i \, (m - m') \, \varphi \right] \left\{ \begin{array}{c} \sin \varphi \\ \cos \varphi \end{array} \right\} = \frac{1}{2} \left\{ \begin{array}{c} - \, i \, \delta_{m+1,m'} \ + \ i \, \delta_{m-1,m'} \\ \delta_{m+1,m'} \ + \ \delta_{m-1,m'} \end{array} \right\} \tag{763}$$

and

$$\frac{1}{2 \pi} \int_0^{2\pi} d\varphi \ \exp \left[i \, (m \, - \, m') \, \varphi \right] \ = \ \delta_{m,m'} \tag{764}$$

The above results lead to the selection rules for the z-component of the electron's orbital angular momentum

$$\begin{aligned} m' \ &= \ m \, \pm \, 1 \\ m' \ &= \ m \end{aligned} \tag{765}$$

An alternate derivation of the selection rules for the z-component of the electron's orbital angular momentum can be found from considerations of the commutation relations

$$\begin{aligned} [\, \hat{L}_z \, , \, x \,] \ &= \ i \, \hbar \, y \\ [\, \hat{L}_z \, , \, y \,] \ &= \ - \, i \, \hbar \, x \\ [\, \hat{L}_z \, , \, z \,] \ &= \ 0 \end{aligned} \tag{766}$$

On taking the matrix elements between states with definite z-components of the angular momenta, one finds

$$\begin{aligned} < \, n'l'm' \, | \, [\, \hat{L}_z \, , \, x \,] \, | \, nlm \, > \ &= \ i \, \hbar \ < \, n'l'm' \, | \, y \, | \, nlm \, > \\ < \, n'l'm' \, | \, [\, \hat{L}_z \, , \, y \,] \, | \, nlm \, > \ &= \ - \, i \, \hbar \ < \, n'l'm' \, | \, x \, | \, nlm \, > \\ < \, n'l'm' \, | \, [\, \hat{L}_z \, , \, z \,] \, | \, nlm \, > \ &= \ 0 \end{aligned} \tag{767}$$

which reduce to

$$\begin{aligned} (\, m' \, - \, m \,) \ < \, n'l'm' \, | \, x \, | \, nlm \, > \ &= \ i \ < \, n'l'm' \, | \, y \, | \, nlm \, > \\ (\, m' \, - \, m \,) \ < \, n'l'm' \, | \, y \, | \, nlm \, > \ &= \ - \, i \ < \, n'l'm' \, | \, x \, | \, nlm \, > \\ (\, m' \, - \, m \,) \ < \, n'l'm' \, | \, z \, | \, nlm \, > \ &= \ 0 \end{aligned} \tag{768}$$

126

From the last equation, it follows that either $m' = m$ or that

$$< n'l'm' \mid z \mid nlm > = 0 \qquad (769)$$

On combining the first two equations, one finds that

$$
\begin{aligned}
(m' - m)^2 < n'l'm' \mid x \mid nlm > &= i (m' - m) < n'l'm' \mid y \mid nlm > \\
&= < n'l'm' \mid x \mid nlm > \qquad (770)
\end{aligned}
$$

The above equation is solved by requiring that either

$$(m' - m)^2 = 1 \qquad (771)$$

or

$$< n'l'm' \mid x \mid nlm > = 0 \qquad (772)$$

Hence, the m-selection rules for the electric dipole transitions are $\Delta m = \pm 1, 0$.

The l-selection rules for electric dipole transitions.

The selection rules for the magnitude of the electron's orbital angular momentum can be found by considering the double commutator

$$[\hat{L}^2 , [\hat{L}^2 , \underline{r}]] = 2 \hbar^2 \left(\underline{r} \, \hat{L}^2 + \hat{L}^2 \, \underline{r} \right) \qquad (773)$$

On taking the matrix elements of this equation between different eigenstates of the magnitude of the orbital angular momentum, one finds

$$\left(l' (l' + 1) - l (l + 1) \right)^2 < n'l'm' \mid \underline{r} \mid nlm > = 2 \left(l' (l' + 1) + l (l + 1) \right) < n'l'm' \mid \underline{r} \mid nlm > \qquad (774)$$

Since

$$\left(l' (l' + 1) - l (l + 1) \right)^2 = (l' + l + 1)^2 (l' - l)^2 \qquad (775)$$

and

$$2 \left(l' (l' + 1) + l (l + 1) \right) = \left((l' + l + 1)^2 + (l' - l)^2 - 1 \right) \qquad (776)$$

the above equation is satisfied if, either

$$< n'l'm' \mid \underline{r} \mid nlm > = 0 \qquad (777)$$

or

$$\left((l' + l + 1)^2 - 1 \right) \left((l' - l)^2 - 1 \right) = 0 \qquad (778)$$

127

The first factor in eqn(778) is always positive when $l' \neq l$, therefore, the electric dipole selection rule becomes $\Delta l = \pm 1$.

The actual values of the matrix elements can be found from explicit calculations. The θ-dependence of the matrix elements is governed by the associated Legendre functions through

$$\Theta_m^l(\theta) = \sqrt{\frac{(2l+1)}{2} \frac{(l-m)!}{(l+m)!}} P_m^l(\cos\theta) \tag{779}$$

which obey the recursion relations

$$\sin\theta \, P_{m-1}^l(\cos\theta) = \frac{P_m^{l+1}(\cos\theta) - P_m^{l-1}(\cos\theta)}{2l+1} \tag{780}$$

and

$$\sin\theta \, P_{m+1}^l(\cos\theta) = \frac{(l+m)(l+m+1) P_m^{l-1}(\cos\theta) - (l-m)(l-m+1) P_m^{l+1}(\cos\theta)}{2l+1} \tag{781}$$

appropriate for the $\Delta m = \pm 1$ transitions and

$$\cos\theta \, P_m^l(\cos\theta) = \frac{(l-m+1) P_m^{l+1}(\cos\theta) + (l+m) P_m^{l-1}(\cos\theta)}{2l+1} \tag{782}$$

for the constant m transition, $\Delta m = 0$. Using the recursion relations, one finds that

$$\sin\theta \, \Theta_m^l(\theta) = \sqrt{\frac{(l+m+2)(l+m+1)}{(2l+1)(2l+3)}} \, \Theta_{m+1}^{l+1}$$
$$- \sqrt{\frac{(l-m)(l-m-1)}{(2l-1)(2l+1)}} \, \Theta_{m+1}^{l-1} \tag{783}$$

for $\Delta m = 1$, while for $\Delta m = -1$ one finds

$$\sin\theta \, \Theta_m^l(\theta) = \sqrt{\frac{(l+m)(l+m-1)}{(2l+1)(2l-1)}} \, \Theta_{m-1}^{l-1}$$
$$- \sqrt{\frac{(l+2-m)(l+1-m)}{(2l+1)(2l+3)}} \, \Theta_{m-1}^{l+1} \tag{784}$$

and for constant m

$$\cos\theta \, \Theta_m^l(\theta) = \sqrt{\frac{(l+1+m)(l+1-m)}{(2l+1)(2l+3)}} \, \Theta_m^{l+1}$$
$$+ \sqrt{\frac{(l+m)(l-m)}{(2l-1)(2l+1)}} \, \Theta_m^{l-1} \tag{785}$$

128

The coefficients in the above equation have a similar form to the Clebsch-Gordon coefficients. The dipole matrix elements can be evaluated by taking the matrix elements of the above set of relations with $\Theta^{l'}_{m'}(\theta)^*$ and then using the orthogonality properties. The above three relations give rise to the selection rules for the magnitude of the orbital angular momentum l

$$l' = l \pm 1 \tag{786}$$

Hence, not only have the selection rules on l been re-derived but the angular integrations have also been evaluated.

What the above mathematics describes is how the spin angular momentum of the emitted photon is combined with the orbital angular momentum of the electron in the final state, so that total angular momentum is conserved. This implies the selection rules which leads to the magnitude of the initial and final electronic angular momentum l having to satisfy the triangular inequality

$$l' + 1 \geq l \geq |l' - 1| \tag{787}$$

as required by the rules of combination of angular momentum. The evaluation of the dipole matrix elements is an explicit example of the Wigner-Eckart theorem. For this example, the irreducible tensor is the vector $\underline{V}$ with components V^μ given by

$$V^\pm = \mp \frac{(x \pm i y)}{\sqrt{2}} = r \sqrt{\frac{4\pi}{3}} Y^1_{\pm 1}$$

$$V^0 = z = r \sqrt{\frac{4\pi}{3}} Y^1_0 \tag{788}$$

Then, since the electric dipole carries angular momentum $(1, \mu)$, the Wigner-Eckart theorem reduces to

$$< n'l'm' \,|\, V^\mu \,|\, nlm > = \frac{1}{\sqrt{2 l' + 1}} \ < l, m; 1, \mu \,|\, l'm' > \ < n'l' \,|\,|V|\,|\, nl > \tag{789}$$

where the first term which represents the angular integration is a Clebsch-Gordon coefficient and the second factor is the reduced matrix element which does depend on the form of the particular vector, but is independent of any choice of coordinate system. Furthermore, the Wigner-Eckart theorem yields the selection rules for the electric dipole transition

$$l + l' \geq 1 \geq |l - l'| \tag{790}$$

Exercise:

Table 1: Matrix Elements of the Components of the Dipole Moment

l'	m'	x	y	z
	$m' = m+1$	$\frac{1}{2}\sqrt{\frac{(l+2+m)(l+1+m)}{(2l+1)(2l+3)}}$	$-\frac{i}{2}\sqrt{\frac{(l+2+m)(l+1+m)}{(2l+1)(2l+3)}}$	-
$l' = l+1$	$m' = m$	-	-	$\sqrt{\frac{(l+1+m)(l+1-m)}{(2l+1)(2l+3)}}$
	$m' = m-1$	$-\frac{1}{2}\sqrt{\frac{(l+2-m)(l+1-m)}{(2l+1)(2l+3)}}$	$-\frac{i}{2}\sqrt{\frac{(l+2-m)(l+1-m)}{(2l+1)(2l+3)}}$	-
	$m' = m+1$	$-\frac{1}{2}\sqrt{\frac{(l-m)(l-1-m)}{(2l-1)(2l+1)}}$	$\frac{i}{2}\sqrt{\frac{(l-m)(l-1-m)}{(2l-1)(2l+1)}}$	-
$l' = l-1$	$m' = m$	-	-	$\sqrt{\frac{(l+m)(l-m)}{(2l-1)(2l+1)}}$
	$m' = m-1$	$\frac{1}{2}\sqrt{\frac{(l+m)(l-1+m)}{(2l-1)(2l+1)}}$	$\frac{i}{2}\sqrt{\frac{(l+m)(l-1+m)}{(2l-1)(2l+1)}}$	-

Using the commutation relations for the j-th component of a vector $\hat{V}^j$ with the i-th component of the orbital angular momentum $\hat{L}^i$,

$$[\ \hat{L}^i\ ,\ \hat{V}^j\ ]\ =\ i\ \hbar\ \sum_k\ \xi^{i,j,k}\ \hat{V}^k \tag{791}$$

where $\xi^{i,j,k}$ is the antisymmetric Levi-Civita symbol, show that

$$[\ \hat{L}^2\ ,\ \underline{\hat{V}}\ ]\ =\ -\ i\ \hbar\ \left(\ \underline{\hat{L}}\ \wedge\ \underline{\hat{V}}\ -\ \underline{\hat{V}}\ \wedge\ \underline{\hat{L}}\ \right)$$
$$=\ -\ 2\ i\ \hbar\ \left(\ \underline{\hat{L}}\ \wedge\ \underline{\hat{V}}\ -\ i\ \hbar\ \underline{\hat{V}}\ \right) \tag{792}$$

From the above equation, derive the double commutation relation

$$[\ \hat{L}^2\ ,\ [\ \hat{L}^2\ ,\ \underline{\hat{V}}\ ]\ ]\ =\ 2\ \hbar^2\ \left(\ \underline{\hat{V}}\ \hat{L}^2\ +\ \hat{L}^2\ \underline{\hat{V}}\ \right)\ -\ 4\ \hbar^2\ \underline{\hat{L}}\ \left(\ \underline{\hat{L}}\ .\ \underline{\hat{V}}\ \right) \tag{793}$$

and that the last term of the above expression is zero if $\underline{\hat{V}}\ =\ \underline{r}$.

The parity selection rule.

In addition to the electronic orbital angular momentum selection rules, there is a parity selection rule. The parity operation is an inversion through the origin given by $\underline{r} \to -\ \underline{r}$. The parity operator $\hat{\mathcal{P}}$ has the effect

$$\hat{\mathcal{P}}\ \psi(\underline{r})\ =\ \psi(-\underline{r}) \tag{794}$$

The parity operator is its own inverse since for any state $\psi(\underline{r})$

$$
\begin{aligned}
\hat{\mathcal{P}}^2 \ \psi(\underline{r}) \ &= \ \hat{\mathcal{P}} \ \psi(-\underline{r}) \\
&= \ \psi(\underline{r})
\end{aligned}
\tag{795}
$$

Therefore, the parity operator has eigenvalues $p \ = \ \pm 1$ for the eigenstates which are defined by

$$
\hat{\mathcal{P}} \ \phi_p(\underline{r}) \ = \ p \ \phi_p(\underline{r})
\tag{796}
$$

so

$$
\hat{\mathcal{P}}^2 \ \phi_p(\underline{r}) \ = \ p^2 \ \phi_p(\underline{r}) \ = \ \phi_p(\underline{r})
\tag{797}
$$

which yields $p^2 \ = \ 1$ or $p \ = \ \pm \ 1$. In polar coordinates, the parity operation is equivalent to a reflection

$$
\theta \ \to \ \pi \ - \ \theta
\tag{798}
$$

followed by a rotation

$$
\varphi \ \to \ \varphi \ + \ \pi
\tag{799}
$$

In electromagnetic processes, parity is conserved since the Coulomb potential is
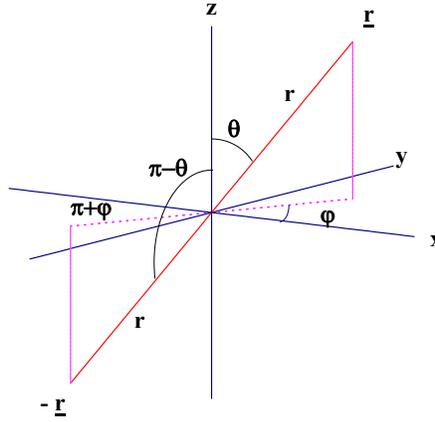
Figure 19: The effect of the parity operator on a displacement vector $\underline{r}$, in spherical polar coordinates.

symmetric under reflection[38]. Therefore, the parity operator $\hat{\mathcal{P}}$ commutes with the Hamiltonian

$$
[\ \hat{\mathcal{P}} \ , \ \hat{H} \ ] \ = \ 0
\tag{800}
$$

and so one can find states $\mid \phi_n >$ that are simultaneous eigenstates of $\hat{H}$ and $\hat{\mathcal{P}}$.

$$
\begin{aligned}
\hat{H} \mid \phi_n > \ &= \ E_n \mid \phi_n > \\
\hat{\mathcal{P}} \mid \phi_n > \ &= \ p_n \mid \phi_n >
\end{aligned}
\tag{801}
$$

---

[38]The weak interaction does not conserve parity.

Inversion transforms vector operators according to

$$\hat{\mathcal{P}} \, \underline{r} \, \hat{\mathcal{P}}^{-1} \; = \; - \, \underline{r} \tag{802}$$

Hence, for any matrix elements of $\underline{r}$ between any eigenstates of the parity operator, one has

$$
\begin{aligned}
< \, \phi_{n'} \mid \underline{r} \mid \phi_n \, > \;\; &= \;\; - \, < \, \phi_{n'} \mid \hat{\mathcal{P}} \, \underline{r} \, \hat{\mathcal{P}}^{-1} \mid \phi_n \, > \\
&= \;\; - \, p_{n'} \, p_n \; < \, \phi_{n'} \mid \underline{r} \mid \phi_n \, >
\end{aligned} \tag{803}
$$

Therefore, the parity must change in an electric dipole transition

$$p_{n'} \, p_n \; = \; - \, 1 \tag{804}$$

This is known as the Laporte selection rule for electric dipole transitions[39]. The validity of this selection follows from the fact that inversion commutes with the orbital angular momentum operator. The spherical harmonics are eigenstates of the parity operator since

$$\hat{\mathcal{P}} \, Y_m^l(\theta, \varphi) \; = \; (\, - \, 1 \,)^l \, Y_m^l(\theta, \varphi) \tag{805}$$

This is proved by examining

$$Y_l^l(\theta, \varphi) \; \propto \; \sin^l \theta \, \exp \left[\, i \, l \, \varphi \, \right] \tag{806}$$

which is seen to be an eigenfunction of the parity operator with eigenvalue $(\, - \, 1 \,)^l$. All other spherical harmonics with the same value of l have the same eigenvalue since the lowering operator (like any component of the angular momentum) commutes with the parity operator. Therefore, one can use the angular momentum selection rule to show that parity does change in an electric dipole transition since

$$(\, - \, 1 \,)^{l+l'} \; = \; - \, 1 \tag{807}$$

The Laporte selection rule is satisfied since $\Delta l \; = \; 1$.

11.1.4 Angular Distribution of Dipole Radiation

We shall assume that the initial state is polarized so that the electron is in an electronic state labelled by m, where the axis of quantization is fixed in space. The decay rate in which a photon of polarization α is emitted into the solid angle $d\Omega_k$ is given by

$$\frac{1}{\tau_{d\Omega_k}} \; = \; \frac{\omega_{nl,n'l'}^3}{2 \, \pi \, \hbar \, c^3} \; d\Omega_k \; \sum_\alpha \; \mid \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{d}_{nlm,n'l'm'} \mid^2 \tag{808}$$

---

[39]O. Laporte, Z. Physik, **23**, 135 (1924).

For a photon emitted in the direction $\hat{k}$

$$\hat{k} \;=\; (\sin\theta_k \cos\varphi_k, \sin\theta_k \sin\varphi_k, \cos\theta_k) \qquad (809)$$

the polar polarization vectors are given by

$$
\begin{aligned}
\hat{\epsilon}_1(\underline{k}) &= (\cos\theta_k \cos\varphi_k, \cos\theta_k \sin\varphi_k, -\sin\theta_k) \\
\hat{\epsilon}_2(\underline{k}) &= (-\sin\varphi_k, \cos\varphi_k, 0)
\end{aligned} \qquad (810)
$$

Therefore, the scalar products of the transition matrix elements of $\underline{r}$ with the polarizations are given by

$$
\begin{aligned}
\hat{\epsilon}_1(\underline{k}) \,.\; <\, n'l'm' \mid \underline{r} \mid nlm\, > \;=\;\; &\frac{1}{2}\; \cos\theta_k\; \exp[\,-\,i\,\varphi_k\,]\; <\, n'l'm' \mid (x+iy) \mid nlm\, > \\
&+\,\frac{1}{2}\; \cos\theta_k\; \exp[\,+\,i\,\varphi_k\,]\; <\, n'l'm' \mid (x-iy) \mid nlm\, > \\
&-\,\sin\theta_k\;\; <\, n'l'm' \mid z \mid nlm\, >
\end{aligned} \qquad (811)
$$

and

$$
\begin{aligned}
\hat{\epsilon}_2(\underline{k}) \,.\; <\, n'l'm' \mid \underline{r} \mid nlm\, > \;=\;\; &-\,\frac{i}{2}\; \exp[\,-\,i\,\varphi_k\,]\; <\, n'l'm' \mid (x+iy) \mid nlm\, > \\
&+\,\frac{i}{2}\; \exp[\,+\,i\,\varphi_k\,]\; <\, n'l'm' \mid (x-iy) \mid nlm\, >
\end{aligned}
$$
$$\qquad (812)$$

Due to the m-selection rules

$$
\begin{aligned}
<\, n'l'm' \mid (x+iy) \mid nlm\, > &\;\propto\; \delta_{m'-m-1} \\
<\, n'l'm' \mid (x-iy) \mid nlm\, > &\;\propto\; \delta_{m'-m+1}
\end{aligned} \qquad (813)
$$

and

$$<\, n'l'm' \mid z \mid nlm\, > \;\propto\; \delta_{m'-m} \qquad (814)$$

the cross-terms in the square of the matrix elements are zero. Hence, on summing over the polarizations, one finds that the (θ_k, φ_k) dependence of the decay is governed by the dipole matrix elements through

$$
\begin{aligned}
\sum_\alpha \mid \hat{\epsilon}_\alpha(\underline{k}) \,.\; \underline{r}_{nlm,n'l'm'} \mid^2 \;=\;\; &\frac{1}{4}\left(\,1\,+\cos^2\theta_k\,\right)\mid\, <\, n'l'm' \mid (x+iy) \mid nlm\, >\mid^2 \\
&+\,\frac{1}{4}\left(\,1\,+\cos^2\theta_k\,\right)\mid\, <\, n'l'm' \mid (x-iy) \mid nlm\, >\mid^2 \\
&+\,\sin^2\theta_k\mid\, <\, n'l'm' \mid z \mid nlm\, >\mid^2
\end{aligned} \qquad (815)
$$

For $l' = l+1$, the above sum is found to depend on the angular factors

$$
\begin{aligned}
I^{m'}_{l'=l+1}(\theta_k,\varphi_k) \;=\;\; &\left(\,1\,+\,\cos^2\theta_k\,\right)\frac{1}{4}\frac{(l+2+m)(l+1+m)}{(2l+1)(2l+3)}\,\delta_{m'-m-1} \\
&+\left(\,1\,+\,\cos^2\theta_k\,\right)\frac{1}{4}\frac{(l+2-m)(l+1-m)}{(2l+1)(2l+3)}\,\delta_{m'-m+1} \\
&+\,\sin^2\theta_k\,\frac{(l+1+m)(l+1-m)}{(2l+1)(2l+3)}\,\delta_{m'-m}
\end{aligned} \qquad (816)
$$

Since the z-component of the final electron's orbital angular momentum is not measured, m' should be summed over. The angular distribution of the emitted radiation for the $l' = l+1$ transition when neither the polarization nor the final state m' value are measured is given by

$$\sum_{m'} I^{m'}_{l'=l+1}(\theta_k, \varphi_k) \; = \; \frac{1}{2}\,(\,1 + \cos^2 \theta_k\,)\,\frac{(l+2)(l+1) \,+\, m^2}{(2l+1)(2l+3)} + \sin^2 \theta_k \,\frac{(l+1)^2 \,-\, m^2}{(2l+1)(2l+3)}$$

$$(817)$$

This factor determines the angular dependence of the emitted electromagnetic radiation, which clearly depends on the value of m specifying the initial electronic state. On rearranging the expression, one finds that the anisotropy is governed by the factor

$$\sum_{m'} I^{m'}_{l'=l+1}(\theta_k, \varphi_k) \; = \; \frac{(l+2)(l+1) \,+\, m^2}{(2l+1)(2l+3)} \; + \; \frac{1}{2}\,\frac{l(l+1) \,-\, 3\,m^2}{(2l+1)(2l+3)}\,\sin^2 \theta_k$$

$$(818)$$

which shows that for $m = 0$ the photons are preferentially emitted perpendicular to the direction of quantization axis since this maximizes the overlap between the polarization and the dipole matrix element. In the opposite case of large values of $m^2\,[\,3\,m^2 \,>\, l\,(l+1)\,]$, one finds that the photons are preferentially emitted parallel (or anti-parallel) to the axis of quantization. On integrating over all directions of the emitted photon, one obtains

$$\frac{1}{4\,\pi}\,\int\,d\Omega_k\,\sum_{m'} I^{m'}_{l'=l+1}(\theta_k, \varphi_k) \; = \; \frac{2}{3}\,\frac{(l+1)}{(2l+1)}$$

$$(819)$$

The independence of the result on m follows since, in this case, there are no angular correlations and the choice of direction of quantization of m is completely arbitrary. The total decay rate for an electron in a state with fixed m due to an $l' = l+1$ transition is given by

$$\frac{1}{\tau_{l'=l+1}} \; = \; \frac{4\,e^2\,\omega^3_{nl,n'l'}}{3\,\hbar\,c^3}\,\frac{(l+1)}{(2l+1)}\,\left|\,\int_0^\infty\,dr\,r^2\,R^*_{n'l+1}(r)\,r\,R_{nl}(r)\,\right|^2$$

$$(820)$$

for $l' = l+1$. This decay rate would be measured in experiments in which neither the final state of the electron nor the final photon state is measured.

However, if the initial electronic state is unpolarized, then one should statistically average over the initial m. In this case, the emitted radiation becomes isotropic

$$\frac{1}{2l+1}\,\sum_{m=-l}^{l}\,\sum_{m'} I^{m'}_{l'=l+1}(\theta_k, \varphi_k) \; = \; \frac{2}{3}\,\frac{(l+1)}{(2l+1)}$$

$$(821)$$

since

$$\frac{1}{2l+1}\,\sum_{m=-l}^{l}\,m^2 \; = \; \frac{1}{3}\,l(l+1)$$

$$(822)$$

134

Hence, if the initial electronic state is unpolarized, the electromagnetic radiation is isotropic. The decay rate for the $l' = l+1$ transition starting with a statistical distribution of m values is given by

$$\frac{1}{\tau_{l'=l+1}} = \frac{4 \ e^2 \ \omega_{nl,n'l'}^3}{3 \ \hbar \ c^3} \ \frac{(l+1)}{(2l+1)} \left| \int_0^\infty dr \ r^2 \ R_{n'l+1}^*(r) \ r \ R_{nl}(r) \right|^2 \quad (823)$$

for $l' = l+1$. This is the same result that was previously obtained for the decay rate of a level with a specific m value, when the m' value of the final state and the polarization or direction of the emitted photon are not measured.

For the case where $l' = l - 1$, one finds that the decay rate involves the angular factor

$$\begin{aligned} I_{l'=l-1}^{m'}(\theta_k, \varphi_k) &= \left(1 \ + \ \cos^2 \theta_k \right) \frac{1}{4} \frac{(l-m)(l-1-m)}{(2l-1)(2l+1)} \ \delta_{m'-m-1} \\ &+ \left(1 \ + \ \cos^2 \theta_k \right) \frac{1}{4} \frac{(l-1+m)(l+m)}{(2l-1)(2l+1)} \ \delta_{m'-m+1} \\ &+ \ \sin^2 \theta_k \ \frac{(l+m)(l-m)}{(2l-1)(2l+1)} \ \delta_{m'-m} \quad (824) \end{aligned}$$

which on summing over the final values of m' yields the angular dependence of the radiation field

$$\sum_{m'} I_{l'=l-1}^{m'}(\theta_k, \varphi_k) = \frac{1}{2} \left(1 + \cos^2 \theta_k \right) \frac{l(l-1) \ + \ m^2}{(2l-1)(2l+1)} + \sin^2 \theta_k \ \frac{l^2 - m^2}{(2l-1)(2l+1)} \quad (825)$$

The anisotropy of the emitted radiation is determined by the factor

$$\sum_{m'} I_{l'=l-1}^{m'}(\theta_k, \varphi_k) = \frac{l(l-1) \ + \ m^2}{(2l-1)(2l+1)} \ + \ \frac{1}{2} \frac{l(l+1) \ - \ 3 \ m^2}{(2l-1)(2l+1)} \ \sin^2 \theta_k \quad (826)$$

which shows that for $m = 0$ the photons are preferentially emitted perpendicular to the direction of quantization axis since this maximizes the overlap between the polarization and the dipole matrix element. In the opposite case of larger m^2 [$3 \ m^2 \ > \ l \ (l+1)$], one finds that the photons are preferentially emitted parallel (or anti-parallel) to the axis of quantization.

Again it is noted that if the initial state is unpolarized, so that m has to be averaged over, then the radiation field is isotropic since

$$\frac{1}{2l+1} \sum_{m=-l}^{l} \sum_{m'} I_{l'=l-1}^{m'}(\theta_k, \varphi_k) = \frac{2}{3} \frac{l}{(2l+1)} \quad (827)$$

Therefore, the decay rate in which the photon is emitted in any direction is given by the expression

$$\frac{1}{\tau_{l'=l-1}} = \frac{4 \ e^2 \ \omega_{nl,n'l'}^3}{3 \ \hbar \ c^3} \ \frac{l}{(2l+1)} \left| \int_0^\infty dr \ r^2 \ R_{n'l-1}^*(r) \ r \ R_{nl}(r) \right|^2 \quad (828)$$

135

for $l' = l - 1$.

Classical Interpretation.

The quantum mechanical results for the angular distribution of the radiation can be understood in terms of a simple classical model of the atom. In Bohr's model, a single electron orbits a central nucleus to which it is bound by the attractive Coulomb potential. We shall assume that the radius of the orbit is a and that the electron is performing a circular orbit in the $x - y$ plane. Since the direction of the electron's orbital angular momentum is aligned with the z-axis, it corresponds to the case where $m \approx l$ and $l \gg 1$. In this case, the electron has an oscillating dipole moment given by

$$
\begin{aligned}
\underline{d}(t) &= q \, a \left(\cos \omega \, t \, \hat{e}_x + \sin \omega \, t \, \hat{e}_y \right) \\
&= q \, a \, \Re e \left(\hat{e}_x - i \, \hat{e}_y \right) \exp \left[i \, \omega \, t \right]
\end{aligned} \tag{829}
$$

This rotating dipole moment can be decomposed into two orthogonal linear dipole moments which oscillate out of phase with each other. It should be
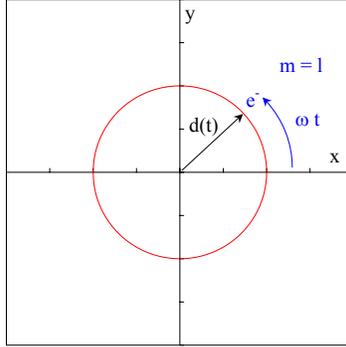
Figure 20: A classical electron orbiting in the x-y plane $(m = l)$ can be considered as producing two perpendicular linearly-oscillating electric-dipole moments.

recalled that a classical oscillating (linear) electric dipole moment radiates power $P(\omega)$ into a solid angle $d\Omega_k$ with a distribution given by

$$
\left(\frac{dP}{d\Omega_k} \right)_{\text{linear}} = \frac{c}{8 \, \pi} \left(\frac{\omega}{c} \right)^4 | \, \underline{d} \, |^2 \, \sin^2 \Theta_{kd} \tag{830}
$$

where Θ_{kd} is the angle between the detector and the direction of the electric dipole. On considering the radiation from the atom to be generated from two orthogonal linear oscillating dipoles, one finds

$$
\left(\frac{dP}{d\Omega_k} \right)_{\text{dipole}} = \frac{c}{8 \, \pi} \left(\frac{\omega}{c} \right)^4 | \, \underline{d} \, |^2 \left(\sin^2 \Theta_{kx} + \sin^2 \Theta_{ky} \right) \tag{831}
$$
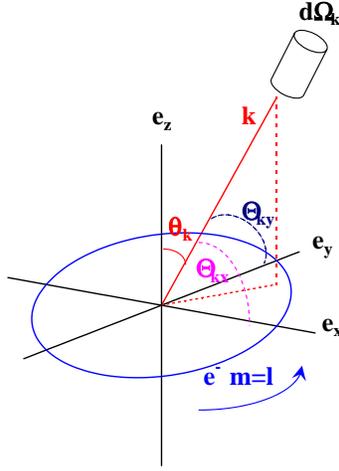
Figure 21: The polarization of the radiated electromagnetic field for an electron orbiting in the x-y plane ($m = l$) can be comprehended in terms of the classical radiation emanating from two linearly oscillating electric-dipole moments. The angles Θ_{kx} and Θ_{ky}, respectively, are the angles between the emitted radiation and the x-axis and the angle subtended by the emitted radiation and the y-axis.

which on using

$$
\begin{aligned}
\cos\Theta_{kx} &= \sin\theta_k \, \cos\varphi_k \\
\cos\Theta_{ky} &= \sin\theta_k \, \sin\varphi_k
\end{aligned}
\tag{832}
$$

becomes

$$
\left(\frac{dP}{d\Omega_k} \right)_{\text{dipole}} = \frac{c}{8\,\pi} \left(\frac{\omega}{c} \right)^4 |\,\underline{d}\,|^2 \left(1 + \cos^2\theta_k \right)
\tag{833}
$$

Since the energy of the emitted photon is given by $\hbar\,\omega$, one finds the angular dependence of the semi-classical prediction of the decay rate is given by

$$
\frac{1}{\tau_{d\Omega_k}} = \frac{e^2}{8\,\pi\,\hbar\,a} \left(\frac{\omega\,a}{c} \right)^3 \left(1 + \cos^2\theta_k \right) d\Omega_k
\tag{834}
$$

The polarization vector is parallel to the direction of the electric field, which in turn is given by the direction of the oscillating dipole that produced it. Hence, a detector which is arranged to accept radiation travelling in the direction $\hat{k}$ will detect polarizations that are found by projecting the electron's orbit onto the plane perpendicular to $\hat{k}$. For example, in this case where the electron's orbit is in the $x - y$ plane, so radiation along the z-axis will be circularly-polarized, whereas radiation in the $x - y$ plane will be linearly-polarized.
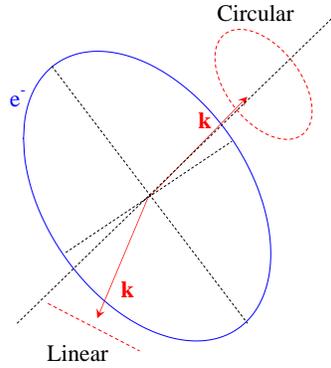
Figure 22: The polarization at a field point can be determined by considering the projection of the electrons orbit on a plane perpendicular to the direction of emission $\underline{k}$. The polarization vector of the classical EM wave follows the projected orbit of the dipole moment.

The angular dependence of the decay rate follows directly from the expressions of eqn(817) and eqn(825) by setting $m \approx l \gg 1$, replacing the radial matrix elements of r by a, adding the expressions and inserting them into eqn(808). The analysis shows that quantum mechanics reproduces the classical limit correctly, as is expected from the correspondence principle.

11.1.5 The Decay Rate from Dipole Transitions.

The decay rate due to dipole transitions includes processes in which photons of all polarizations are emitted in all directions. Accordingly, the decay rate is found by summing over all polarizations and integrating over the directions of the emitted photon. For a spherically symmetric system, the energy will be independent of the z-component of the orbital angular momentum. In this case, one should sum over all values of m' corresponding to the degenerate final states. On summing over all final states corresponding to a specific l' value, that is on summing over m' where $m' = m, \ m \pm 1$, one finds that the transition rate can be expressed as

$$\left(\frac{1}{\tau} \right) \;=\; \frac{4 \ e^2 \ \omega_{nl,n'l'}^3}{3 \ \hbar \ c^3} \left\{ \begin{array}{c} \frac{(l+1)}{(2l+1)} \\ \frac{l}{(2l+1)} \end{array} \right\} \left| \int_0^\infty dr \ r^2 \ R_{n'l'}(r) \ r \ R_{nl}(r) \ \right|^2 (835)$$

for

$$l' \;=\; \left\{ \begin{array}{c} l+1 \\ l-1 \end{array} \right. \tag{836}$$

It should be noted that, for a fixed l', the lifetime of the state $\mid nlm >$ is independent of the value of m. This is expected since the choice of the quantization

138

Table 2: Radial wave functions $R_{nl}(\rho)$ for a Hydrogenic-like atom, where $\rho = \frac{Zr}{a}$. The functions are normalized so that $\int_0^\infty d\rho\ \rho^2\ R_{nl}\ R_{nl}\ =\ 1$.

$n\ =\ 1$	$l\ =\ 0$	$2\ \exp\left[\ -\ \rho\ \right]$
$n\ =\ 2$	$l\ =\ 0$	$\frac{1}{\sqrt{2}}\left(1\ -\ \frac{\rho}{2}\right)\exp\left[\ -\ \frac{\rho}{2}\right]$
	$l\ =\ 1$	$\frac{1}{2\sqrt{6}}\ \rho\ \exp\left[\ -\ \frac{\rho}{2}\right]$
$n\ =\ 3$	$l\ =\ 0$	$\frac{2}{3^{\frac{3}{2}}}\left(1\ -\ \frac{2}{3}\rho\ +\ \frac{2}{27}\rho^2\right)\exp\left[\ -\ \frac{\rho}{3}\right]$
	$l\ =\ 1$	$\frac{2^{\frac{5}{2}}}{3^{\frac{7}{2}}}\left(1\ -\ \frac{\rho}{6}\right)\rho\ \exp\left[\ -\ \frac{\rho}{3}\right]$
	$l\ =\ 2$	$\frac{2^{\frac{3}{2}}}{3^{\frac{9}{2}}\sqrt{5}}\ \rho^2\ \exp\left[\ -\ \frac{\rho}{3}\right]$

Table 3: Values of $|\int_0^\infty dr\ r^2\ R_{nl}\ r\ R_{n'l-1}|^2$ in atomic units.

n,l	$n',l-1$	
np	$1s$	$2^8 n^7 (n-1)^{2n-5}(n+1)^{-2n-5}$
	$2s$	$2^{17} n^7 (n^2-1)(n-2)^{2n-6}(n+2)^{-2n-6}$
	$3s$	$2^8 3^7 n^7 (n^2-1)(n-3)^{2n-8}(7n^2-27)^2(n+3)^{-2n-8}$
nd	$2p$	$2^{19} 3^{-1} n^9 (n^2-1)(n-2)^{2n-7}(n+2)^{-2n-7}$
	$3p$	$2^{11} 3^9 n^9 (n^2-1)(n^2-4)(n-3)^{2n-8}(n+3)^{-2n-8}$
nf	$3d$	$2^{13} 3^9 5^{-1} n^{11}(n^2-1)(n^2-4)(n-3)^{2n-9}(n+3)^{-2n-9}$

direction is completely arbitrary.

There are no selection rules associated with the radial integration in the dipole matrix elements

$$\int_0^\infty dr\ r^2\ R_{n'l-1}\ r\ R_{nl} \tag{837}$$

The radial part of the dipole matrix element can be expressed in terms of the hypergeometric function $F(a,b,c)$ via

$$
\begin{aligned}
&\int_0^\infty dr\ r^2\ R_{n'l-1}\ r\ R_{nl} \\
=\ &\frac{a\ (-1)^{n'-l}}{4(2l-1)}\ \sqrt{\frac{(n+l)!(n'+l-1)!}{(n-l-1)!(n'-l)!}}\left[\frac{(4n'n)^{l+1}(n-n')^{n'+n-2l-2}}{(n'+n)^{n'+n}}\right] \\
&\times\left[F(l+1-n,l-n',2l,-\frac{4n'n}{(n'-n)^2})-\left(\frac{n'-n}{n'+n}\right)^2 F(l+1-n,l-n',-\frac{4n'n}{(n'-n)^2})\right]
\end{aligned}
\tag{838}
$$

Simple analytic expressions for the squares of the matrix elements for small values of (n',l) are shown in Table(3).

The radial integrations were evaluated by Schrödinger[40] using the generating

---

[40] E. Schrödinger, Ann. der Phys. **79**, 362 (1926).

140

function expansion for the Laguerre polynomials. Eckart[41] and Gordon[42] have calculated these dipole matrix elements by other means. In general, the lifetime of the hydrogenic states increases with increasing n, varying roughly as n^3 for a fixed value of l. The decrease in the dipole matrix elements with increasing n is simply due to the increasing numbers of nodes in the radial wave functions.

The magnitude of the decay rate is estimated as

$$\frac{1}{\tau} \sim \frac{c}{a} \left(\frac{\omega \, a}{c} \right)^3 \left(\frac{e^2}{\hbar \, c} \right) \tag{839}$$

where the magnitude of the dipole matrix element is estimated as $e \, a$ where a is the Bohr radius. On setting $\hbar\omega$ equal to the electrostatic energy of hydrogen, the remaining factor is estimated to have the magnitude

$$\left(\frac{\omega \, a}{c} \right) \sim \left(\frac{e^2}{\hbar \, c} \right) \tag{840}$$

where the length scale a has dropped out. Hence, as

$$\left(\frac{e^2}{\hbar \, c} \right) \approx \frac{1}{137.0359979} \tag{841}$$

one finds that the decay rate is given by

$$\frac{1}{\tau} \sim \frac{c}{a} \left(\frac{e^2}{\hbar \, c} \right)^4 \tag{842}$$

so the decay time is approximately eight orders of magnitude larger than the time taken for the photon to cross the atom. When averaged over l, the electric dipole decay rate is given by

$$\frac{1}{\tau_n} \propto \sum_l \frac{(2l+1)}{n^5} \sim n^{-\frac{9}{2}} \tag{843}$$

so, as seen in Table(4), the decay is slower for the higher energy levels.

11.1.6 The $2p \rightarrow 1s$ Electric Dipole Transition Rate.

Consider the decay of the $2p$ state (with $m = 0$) to the $1s$ state in the hydrogen atom. As can be seen from Table(2), the initial state is described by an electronic wave function

$$\psi_{2p}(\underline{r}) = \frac{1}{2 \sqrt{6} \, a^3} \left(\frac{r}{a} \right) \exp \left[- \frac{1}{2} \frac{r}{a} \right] \sqrt{\frac{3}{4 \, \pi}} \cos \theta \tag{844}$$

---

[41] C. Eckart, Phys. Rev. **28**, 927 (1926).
[42] W. Gordon, Ann. der Phys. **2**, 1031 (1929).

Table 4: Electric Dipole Transition Rates for Hydrogen, in units of 10^8sec^{-1}.

Initial	Final	n=1	n=2	n=3
2p	ns	6.25	-	-
3s	np	-	0.063	-
3p	ns	1.64	0.22	-
3d	np	-	0.64	-
4s	np	-	0.025	0.018
4p	ns	0.68	0.095	0.030
4p	nd	-	-	0.003
4d	np	-	0.204	0.070
4f	nd	-	-	0.137

and the final state electronic is given by

$$\psi_{1s}(\underline{r}) \;=\; \frac{2}{\sqrt{a^3}} \; \exp\left[-\frac{r}{a} \right] \frac{1}{\sqrt{4\pi}} \tag{845}$$

where the length scale a is the Bohr radius

$$a \;=\; \frac{\hbar^2}{m\, e^2} \tag{846}$$

The decay rate in the Fermi-Golden rule, evaluated in the dipole approximation, is given by

$$\frac{1}{\tau} \;=\; \frac{4\,\omega_{1,2}^3\,\underline{d}^*_{1s,2p}\cdot\underline{d}_{1s,2p}}{3\,\hbar\,c^3} \tag{847}$$

The frequency is evaluated from

$$\begin{aligned}
\hbar\,\omega_{12} \;&=\; E_{2p} \;-\; E_{1s} \\
&=\; \frac{m\,e^4}{2\,\hbar^2}\left(1 - \frac{1}{4} \right) \\
&=\; \frac{3}{8}\,\frac{e^2}{a}
\end{aligned} \tag{848}$$

Hence,

$$\frac{1}{\tau} \;=\; \frac{4}{3}\left(\frac{3}{8}\right)^3 \left(\frac{e^2}{\hbar\,a}\right)^3 \frac{e^2\,a^2}{\hbar\,c^3}\left|\frac{\underline{d}_{1s,2p}}{e\,a}\right|^2$$

142

$$
= \quad \frac{9}{128} \left(\frac{e^2}{\hbar\, c} \right)^4 \frac{c}{a} \left| \frac{\underline{d}_{1s,2p}}{e\, a} \right|^2 \tag{849}
$$

Therefore, the scattering rate is determined by the ratio $\frac{c}{a}$ but also is modified by the fourth power of the dimensionless electromagnetic coupling strength

$$
\left(\frac{e^2}{\hbar\, c} \right) \approx \frac{1}{137.0359979} \tag{850}
$$

The smallness of this factor allows us to only consider the Fermi-Golden rule expression for the decay rate. The dimensionless dipole matrix elements are expected to be non-zero, since they obey the selection rules. They are non-zero, as can be directly verified by performing an integration. The only non-zero dipole matrix element originates from the z-component of the dipole

$$
\underline{d}_{1s,2p} = e \int d^3\underline{r}\ \psi_{1s}^*(\underline{r})\ \underline{r}\ \psi_{2p}(\underline{r}) \tag{851}
$$

since only the z-component satisfies the $\Delta m = 0$ selection rule. The angular integration is evaluated as

$$
\frac{\sqrt{3}}{4\,\pi} \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta\ \sin\theta\ \cos^2\theta \quad = \quad \frac{\sqrt{3}}{4\,\pi}\ 2\,\pi\ \frac{2}{3}
$$
$$
= \quad \frac{1}{\sqrt{3}} \tag{852}
$$

and the radial integration yields

$$
\int_0^{\infty} dr\ r^2\ R_{1s}^*(r)\ r\ R_{2p}(r) \quad = \quad \frac{2}{2\,\sqrt{6}} \int_0^{\infty} dr\ \frac{r^4}{a^3}\ \exp\left[-\frac{3}{2}\frac{r}{a} \right]
$$
$$
= \quad \frac{a}{\sqrt{6}} \left(\frac{2}{3} \right)^5 \int_0^{\infty} dx\ x^4\ \exp\left[-x \right]
$$
$$
= \quad a\ \frac{4!}{\sqrt{6}} \left(\frac{2}{3} \right)^5 = 4\,a\,\sqrt{6} \left(\frac{2}{3} \right)^5 \tag{853}
$$

Hence, the magnitude of the dipole matrix element is evaluated as

$$
\frac{d_{1s,2p}}{e\, a} = 4\,\sqrt{2} \left(\frac{2}{3} \right)^5 \tag{854}
$$

Therefore, the dipole allowed decay rate is given by

$$
\frac{1}{\tau} = \left(\frac{2}{3} \right)^8 \left(\frac{e^2}{\hbar\, c} \right)^4 \frac{c}{a} \tag{855}
$$

Hence, the time scale τ is of the order of 10^{-10} seconds. The exact value of the decay time is calculated to be 1.6×10^{-9} seconds.

11.1.7 Electric Quadrupole and Magnetic Dipole Transitions.

Consider decays such as the $3d$ state (with $m = 0$) to the $1s$ state in the hydrogen atom. Since, in this transition, the change in the electron's angular momentum is two units, the transition is forbidden in the dipole approximation. Therefore, the transition rate is evaluated by keeping the next order term in the expansion

$$\exp\left[-\, i\, \underline{k} \cdot \underline{r} \right] \;\approx\; 1 \,-\, i\, \underline{k} \cdot \underline{r} \,+\, \ldots \tag{856}$$

The second term in the expansion describes electric quadrupole and magnetic dipole transitions.

The matrix elements that have to be evaluated are of the form

$$<\, n'l'm' \mid (\, \underline{k} \cdot \underline{r} \,)\, (\, \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{\hat{p}} \,) \mid nlm \,> \tag{857}$$

This shall be written as the sum of two terms, with different symmetries with respect to interchange of $\underline{r}$ and $\underline{p}$. These two terms will describe electric quadrupole and magnetic dipole transitions. The matrix elements are written as the sum of a term symmetric under the interchange of $\underline{r}$ and $\underline{\hat{p}}$ and a term that is anti-symmetric

$$
\begin{aligned}
(\, \underline{k} \cdot \underline{r} \,)\, (\, \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{\hat{p}} \,) \;=\;\; & \frac{1}{2}\Big(\, (\, \underline{k} \cdot \underline{r} \,)\, (\, \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{\hat{p}} \,) \,+\, (\, \underline{k} \cdot \underline{\hat{p}} \,)\, (\, \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{r} \,) \,\Big) \\
& +\, \frac{1}{2}\Big(\, (\, \underline{k} \cdot \underline{r} \,)\, (\, \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{\hat{p}} \,) \,-\, (\, \underline{k} \cdot \underline{\hat{p}} \,)\, (\, \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{r} \,) \,\Big)
\end{aligned}
\tag{858}
$$

The first term represents the matrix elements for the electric quadrupole transitions[43], and the second term represents the matrix elements for the magnetic dipole transitions. The first term can be written as the scalar products of a symmetric dyadic

$$\underline{k} \cdot \left(\, \underline{r}\ \underline{\hat{p}} \,+\, \underline{\hat{p}}\ \underline{r}\, \right) \cdot \hat{\epsilon}_\alpha(\underline{k}) \tag{859}$$

The scalar products are organized such that the left most vector outside the parenthesis forms a scalar product with the left most vectors within the parenthesis, and likewise with the right most vectors. The electronic matrix elements only involve the dyadic operator, as the wave vector and polarization vectors are properties of the photon. The matrix elements

$$<\, n'l'm' \mid \left(\, \underline{r}\ \underline{\hat{p}} \,+\, \underline{\hat{p}}\ \underline{r}\, \right) \mid nlm \,> \tag{860}$$

are evaluated by first noting that

$$[\, \underline{r} \,,\, \underline{\hat{p}}^2 \,] \;=\; 2\, i\, \hbar\, \underline{\hat{p}} \tag{861}$$

---

[43] J. A. Gaunt and W. H. McCrea, Proc. Camb. Phil. Soc. **23**, 930 (1927).

which allows the momentum operator to be written as

$$\hat{\underline{p}} = \frac{i\,m}{\hbar}\,[\,\hat{H}\,,\,\underline{r}\,] \tag{862}$$

Therefore, the matrix elements of the dyadic can be expressed in the form of the matrix elements of the commutator with the dyadic

$$
\begin{aligned}
< n'l'm' \,|\, \Big(\; \underline{r}\;\hat{\underline{p}} \,+\, \hat{\underline{p}}\;\underline{r}\;\Big)\,|\, nlm > \;
&=\; \frac{i\,m}{\hbar}\; < n'l'm'\,|\,\underline{r}\,[\,\hat{H}\,,\,\underline{r}\,] \,+\, [\,\hat{H}\,,\,\underline{r}\,]\,\underline{r}\,|\, nlm > \\[4pt]
&=\; \frac{i\,m}{\hbar}\; < n'l'm'\,|\,[\,\hat{H}\,,\,\underline{r}\,\underline{r}\,]\,|\, nlm > \\[4pt]
&=\; i\,m\,\frac{(\,E_{n'l'm'} \,-\, E_{nlm}\,)}{\hbar}\; < n'l'm'\,|\,\underline{r}\,\underline{r}\,|\, nlm > \\[4pt]
&=\; i\,m\,\omega_{n',n}\; < n'l'm'\,|\,\underline{r}\,\underline{r}\,|\, nlm > \tag{863}
\end{aligned}
$$

The decay rate in the Fermi-Golden rule, evaluated in the electric quadrupole approximation, is given by

$$
\begin{aligned}
\frac{1}{\tau} \;&=\; \Big(\frac{e}{m\,c}\Big)^2 \int d^3\underline{k}\,\frac{m^2\,c^2\,\omega_k^2}{8\,\pi\,\omega_k}\,\sum_\alpha \,|\,\underline{k}\;.\; < n'l'm'\,|\,\underline{r}\,\underline{r}\,|\, nlm > \;.\; \hat{\epsilon}_\alpha(\underline{k})\,|^2 \\
&\quad\times\, \delta(\,E_{nlm} \,-\, E_{n'l'm'} \,-\, \hbar\,\omega_k\,) \\[4pt]
&=\; \frac{e^2}{8\,\pi\,\hbar}\,\Big(\frac{\omega_{nl,n'l'}}{c}\Big)^5 \int d\Omega_{\underline{k}}\,\sum_\alpha \,|\,\hat{k}\;.\; < n'l'm'\,|\,\underline{r}\,\underline{r}\,|\, nlm > \;.\; \hat{\epsilon}_\alpha(\underline{k})\,|^2
\end{aligned}
\tag{864}
$$

where, in the second line $\underline{k}$ is restricted to have the magnitude

$$k \;=\; \frac{\omega_{nl,n'l'}}{c} \tag{865}$$

The frequency is evaluated from

$$
\begin{aligned}
\hbar\,\omega_{nl,n'l'} \;&=\; E_{nl} \,-\, E_{n'l'} \\[4pt]
&\sim\; \frac{m\,e^4}{\hbar^2} \;=\; m\,c^2\,\Big(\frac{e^2}{\hbar\,c}\Big)^2 \\[4pt]
&\sim\; \frac{e^2}{a} \tag{866}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{1}{\tau} \;&\sim\; \Big(\frac{e^2}{\hbar\,a\,c}\Big)^5 \frac{e^2}{8\,\pi\,\hbar} \int d\Omega_{\underline{k}}\,\sum_\alpha \,|\,\hat{k}\;.\; < n'l'm'\,|\,\underline{r}\,\underline{r}\,|\, nlm > \;.\; \hat{\epsilon}_\alpha(\underline{k})\,|^2 \\[4pt]
&\sim\; \Big(\frac{e^2}{\hbar\,c}\Big)^6 \frac{c}{a} \tag{867}
\end{aligned}
$$

Therefore, the scattering rate is determined by the ratio $\frac{c}{a}$ but also is modified by the sixth power of the dimensionless electromagnetic coupling strength

$$\left(\frac{e^2}{\hbar\, c} \right) \approx \frac{1}{137.0359979} \tag{868}$$

The smallness of this factor allows us to only consider the Fermi-Golden rule expression for the decay rate. The dimensionless quadrupole matrix elements are expected to be non-zero, since they obey the selection rules which involve the exchange of two units of angular momentum. They are non-zero, as can be directly verified by performing an integration. Therefore, the quadrupole allowed decay rate is given by

$$\frac{1}{\tau} \sim \left(\frac{e^2}{\hbar\, c} \right)^6 \frac{c}{a} \tag{869}$$

Hence, the time scale τ is expected to be of the order of 10^{-6} seconds[44].

This type of transition is known as an electric quadrupole transition. Because of the transversality condition

$$\hat{k} \,.\, \hat{\epsilon}_\alpha(\underline{k}) \;=\; 0 \tag{870}$$

one can add a diagonal term to the dyadic without affecting the result. A diagonal term with a magnitude that makes the resulting dyadic traceless is added to the dyadic, leading to the expression

$$Q_{i,j} \;=\; e \left(x_i\, x_j \;-\; \frac{1}{3}\, \delta_{i,j} \mid \underline{r} \mid^2 \right) \tag{871}$$

Therefore, the transition rate is governed by the electric quadrupole tensor

$$< \; n'l'm' \mid Q_{i,j} \mid nlm \; > \tag{872}$$

The symmetric dyadic $Q_{i,j}$ has six inequivalent components, which because of the restriction that the dyadic is traceless, can be reduced to five independent components. Due to the transformational properties of the dyadic under rotation, it can be expressed as a linear combination of the spherical harmonics $Y_m^2(\theta, \varphi)$ and nothing else[45]. This can be seen from rewriting the quadrupole tensor $\tilde{Q}$

$$\frac{\tilde{Q}}{e} \;=\; \begin{pmatrix} xx - \frac{r^2}{3} & xy & xz \\ yx & yy - \frac{r^2}{3} & yz \\ zx & zy & zz - \frac{r^2}{3} \end{pmatrix} \tag{873}$$

---

[44]This estimate will be modified upwards by several orders of magnitude, due to the presence of a large dimensionless factor that was not accounted for.

[45]The transformational properties of the dyadic follow immediately from the transformational properties of the vector $\underline{r}$

in terms of spherical polar coordinates

$$\frac{\tilde{Q}}{e\ r^2} = \begin{pmatrix} \frac{1}{2}\sin^2\theta\ \cos 2\varphi - \frac{1}{6}(3\cos^2\theta - 1) & \frac{1}{2}\sin^2\theta\sin 2\varphi & \sin\theta\cos\theta\cos\varphi \\ \frac{1}{2}\sin^2\theta\sin 2\varphi & -\frac{1}{2}\sin^2\theta\ \cos 2\varphi - \frac{1}{6}(3\cos^2\theta - 1) & \sin\theta\cos\theta\sin\varphi \\ \sin\theta\cos\theta\cos\varphi & \sin\theta\cos\theta\sin\varphi & \frac{1}{3}(3\cos^2\theta - 1) \end{pmatrix}$$

$$(874)$$

The presence of states with orbital angular momentum of only two makes the dyadic an irreducible second rank tensor. Application of the Wigner-Eckart theorem to an irreducible second rank tensor results in the electric quadrupole selection rules

$$l + l' \geq 2 \geq |l - l'| \tag{875}$$

The angular momentum carried away by the photon consists of the spin-one carried away by the photon in addition to the component of the photon's wave function described by the spherical Bessel function $j_1(kr) \sim k\,r$ which carries off one unit of orbital angular momentum. In addition to the angular momentum selection rules, there are parity selection rules for the electric quadrupole transitions. Since the parity operator satisfies

$$\hat{\mathcal{P}}\ \underline{r}\ \hat{\mathcal{P}}^{-1} = -\underline{r} \tag{876}$$

then the electric quadrupole matrix elements satisfy

$$\begin{aligned} < n' \mid \underline{r}\,\underline{r} \mid n > &= < n' \mid \hat{\mathcal{P}}\ \underline{r}\,\underline{r}\ \hat{\mathcal{P}}^{-1} \mid n > \\ &= p_{n'}\ p_n\ \ < n' \mid \underline{r}\,\underline{r} \mid n > \end{aligned} \tag{877}$$

Therefore, the parity does not change in an electric quadrupole transition as $p_{n'}\ p_n = 1$.

The magnetic dipole matrix elements are given by

$$\frac{1}{2}\Bigg((\ \underline{k}\cdot\underline{r}\ )(\ \hat{\epsilon}_\alpha(\underline{k})\cdot\underline{\hat{p}}\ ) - (\ \underline{k}\cdot\underline{\hat{p}}\ )(\ \hat{\epsilon}_\alpha(\underline{k})\cdot\underline{r}\ )\Bigg) \tag{878}$$

which can be re-written as

$$\frac{1}{2}\Bigg((\ \underline{k}\wedge\hat{\epsilon}_\alpha(\underline{k})\ )\cdot(\ \underline{r}\wedge\underline{\hat{p}}\ )\Bigg) \tag{879}$$

The first term is of the form

$$\underline{B} = \underline{\nabla}\wedge\underline{A} \rightarrow \underline{k}\wedge\underline{A} \rightarrow \underline{\nabla}\wedge\hat{\epsilon}_\alpha(\underline{k}) \tag{880}$$

and the second term

$$\underline{r}\wedge\underline{p} \tag{881}$$

is the orbital angular momentum. The orbital angular momentum produces the orbital magnetic moment given by

$$\underline{\mu} = \frac{e}{2\ m\ c}\ (\ \underline{r}\wedge\underline{p}\ ) \tag{882}$$

147

The magnetic dipole transition should be extended from orbital angular momentum to include the spin magnetic moment which is of the same order

$$\left(\frac{e\,\hbar}{2\,m\,c} \right) \underline{\sigma} \cdot (\underline{k} \wedge \hat{\epsilon}_\alpha(\underline{k}))$$

$$\left(\frac{e}{2\,m\,c} \right) (\underline{r} \wedge \underline{p}) \cdot (\underline{k} \wedge \hat{\epsilon}_\alpha(\underline{k})) \tag{883}$$

since orbital angular momentum is quantized in units of $\hbar$. The angular momentum selection rule for the magnetic dipole transition is given by

$$\Delta l = 0 \tag{884}$$

and

$$1 \geq | \Delta m | \tag{885}$$

also parity does not change.

Terms with higher-order orbital angular momentum that occur in the expansion of the photon's wave function $\exp[\, i\, \underline{k} \cdot \underline{r} \,]$ can be found by using the Rayleigh expansion. The terms with orbital angular momentum l are proportional to the spherical harmonics $j_l(kr)$ which vary as $(kr)^l$ when $kr \to 0$, as is found from the expansion of the exponential term. The presence of the extra factors k^l in the matrix element has the result that the electric 2^s-th multi-pole transition rates are found to vary as

$$\frac{1}{\tau} \propto \left(\frac{\omega_{n,n'}}{c} \right)^{2s+1} a^{2s} \tag{886}$$

where s is the magnitude of the change in the electronic orbital angular momentum, which satisfies the inequality

$$(l + l') \geq s \geq | l' - l | \tag{887}$$

The extra factors from the photon's angular momentum results in an overall decrease in the electric multi-pole transition rate by a factor of $(\frac{e^2}{\hbar\,c})^{2s}$. It should also be noted that the relative strength of the higher-order electric multi-pole transitions increase more rapidly with Z than the electric dipole transitions. Therefore, it is frequently found that the quadrupole transitions cannot be neglected for the heavy elements. Alternatively, higher-order multipole transitions do become important in the x-ray region, since in this region the wavelength of the radiation is comparable to the spatial extent of the charged particle's wave function.

11.1.8   The $3d \to 1s$ Electric Quadrupole Transition Rate

The transition $3d \to 1s$ is forbidden to occur via the dipole process, since it involves a change of l by two units. It may occur as an electric quadrupole

transition. The electric quadrupole transition rate can be expressed as

$$\frac{1}{\tau} = \frac{1}{8\,\pi\,\hbar} \left(\frac{\omega}{c}\right)^5 \int d\Omega_k \sum_{\alpha} |\ \hat{k}\ .\ <\ 1s\ |\ \tilde{Q}\ |\ 3d\ >\ .\ \hat{\epsilon}_{\alpha}(\underline{k})\ |^2 \qquad (888)$$

where $\tilde{Q}$ represents the quadrupole tensor. The frequency factor can be evaluated as

$$\left(\frac{\omega}{c}\right) = \frac{4}{9\,a} \left(\frac{e^2}{\hbar\,c}\right) \qquad (889)$$

hence, the rate can be expressed in the form

$$\frac{1}{\tau} = \frac{c}{8\,\pi\,a} \left(\frac{4}{9}\right)^5 \left(\frac{e^2}{\hbar\,c}\right)^6 \int d\Omega_k \sum_{\alpha} \frac{1}{e^2\,a^4} |\ \hat{k}\ .\ <\ 1s\ |\ \tilde{Q}\ |\ 3d\ >\ .\ \hat{\epsilon}_{\alpha}(\underline{k})\ |^2$$
$$(890)$$

We shall consider the transition from the $m = 0$ state of the $3d$ level to the $1s$ state. As can be easily shown, the matrix elements of quadrupole tensor for this transition are diagonal and are given by

$$<\ 1s\ |\ \tilde{Q}\ |\ 3d\ > = <\ 1s\ | \begin{pmatrix} -\frac{Q_{zz}}{2} & 0 & 0 \\ 0 & -\frac{Q_{zz}}{2} & 0 \\ 0 & 0 & Q_{zz} \end{pmatrix} |\ 3d\ > \qquad (891)$$

Therefore, the transition matrix elements are of the form

$$\int d\Omega_k \sum_{\alpha} \frac{1}{e^2\,a^4} |\ \hat{k}\ .\ <\ 1s\ |\ \tilde{Q}\ |\ 3d\ >\ .\ \hat{\epsilon}_{\alpha}(\underline{k})\ |^2$$
$$= \int d\Omega_k \sum_{\alpha} \frac{|\ <\ 1s\ |\ Q_{zz}\ |\ 3d\ >\ |^2}{e^2\,a^4} \left(\ \hat{k}_z\ \hat{\epsilon}_{\alpha}(\underline{k})_z\ -\ \frac{1}{2}\ \hat{k}_x\ \hat{\epsilon}_{\alpha}(\underline{k})_x\ -\ \frac{1}{2}\ \hat{k}_y\ \hat{\epsilon}_{\alpha}(\underline{k})_y\ \right)^2$$
$$(892)$$

The direction of the emitted photon $\hat{k}$ is expressed as

$$\hat{k} = (\sin\theta_k \cos\varphi_k, \sin\theta_k \sin\varphi_k, \cos\theta_k) \qquad (893)$$

and the polarization vectors are given by

$$\hat{\epsilon}_1(\underline{k}) = (\cos\theta_k \cos\varphi_k, \cos\theta_k \sin\varphi_k, -\sin\theta_k)$$
$$\hat{\epsilon}_2(\underline{k}) = (-\sin\varphi_k, \cos\varphi_k, 0) \qquad (894)$$

Thus for the $m = 0$ level, one finds that the integral over the angular distribution is given by

$$\int d\Omega_k \sum_{\alpha} \frac{1}{e^2\,a^4} |\ \hat{k}\ .\ <\ 1s\ |\ \tilde{Q}\ |\ 3d\ >\ .\ \hat{\epsilon}_{\alpha}(\underline{k})\ |^2$$
$$= \int d\Omega_k \frac{|\ <\ 1s\ |\ Q_{zz}\ |\ 3d\ >\ |^2}{e^2\,a^4} \left(\ -\ \frac{3}{2}\ \sin\theta_k\ \cos\theta_k\ \right)^2$$
$$= \frac{3}{10}\ 4\,\pi\ \frac{|\ <\ 1s\ |\ Q_{zz}\ |\ 3d\ >\ |^2}{e^2\,a^4} \qquad (895)$$

The scattering rate becomes

$$
\frac{1}{\tau} = \frac{3}{20} \left(\frac{4}{9}\right)^5 \frac{c}{a} \left(\frac{e^2}{\hbar\, c}\right)^6 \frac{|<1s\,|\,Q_{zz}\,|\,3d>|^2}{e^2\,a^4} \tag{896}
$$

The dimensionless quadrupole matrix elements are evaluated as the product of an angular integral and a radial integral

$$
\begin{aligned}
\frac{<1s\,|\,Q_{zz}\,|\,3d>}{e\,a^2} &= \int d\Omega\; Y_0^0(\theta,\varphi)^* \frac{1}{3}\,(\,3\,\cos^2\theta\,-\,1\,)\, Y_0^2(\theta,\varphi) \\
&\quad \times \int_0^\infty dr\; r^2\; R_{1s}^*(r)\; \frac{r^2}{a^2}\; R_{3d}(r) \\
&= -\frac{1}{3}\,\sqrt{\frac{4}{5}}\,4\left(\frac{3}{4}\right)^{\frac{9}{2}} \tag{897}
\end{aligned}
$$

Finally, one finds the resulting expression for the quadrupole decay rate of the $3d$ state with $m = 0$

$$
\frac{1}{\tau} = \frac{1}{3600} \frac{c}{a} \left(\frac{e^2}{\hbar\, c}\right)^6 \tag{898}
$$

which is evaluated as $228\ \mathrm{sec}^{-1}$. From the above analysis, it is seen that angular distribution for the emitted photon is governed by the factor

$$
\cos^2\theta_k\;\sin^2\theta_k \tag{899}
$$

and the intensity is largest for the cone with $\theta_k \approx 0.28\,\pi$ or $0.72\,\pi$. This angular dependence of the emitted radiation is the same as found by considering the radiation from an oscillating classical quadrupole, for which the radiated power is given by

$$
\left(\frac{dP}{d\Omega_k}\right)_{\mathrm{quad}} = \frac{c}{288\,\pi} \left(\frac{\omega}{c}\right)^6 Q^2\;\cos^2\theta_k\;\sin^2\theta_k \tag{900}
$$

However, the angular distribution of the emitted quadrupole radiation is to be contrasted with the $\Delta m = 0$ decay of a $2p$ electron, that has an angular distribution of emitted photons given by

$$
\sin^2\theta_k \tag{901}
$$

which is maximum for $\theta_k = \frac{\pi}{2}$.

The $3d \rightarrow 1s$ quadrupole decay rate will be dwarfed by dipole allowed cascade emission processes such as $3d \rightarrow 2p$ followed by $2p \rightarrow 1s$. Therefore, the intensity of the emitted light corresponding to the $3d \rightarrow 1s$ process is expected to be extremely weak. However, the quadrupole line is expected to be much more readily observed in absorption spectra.
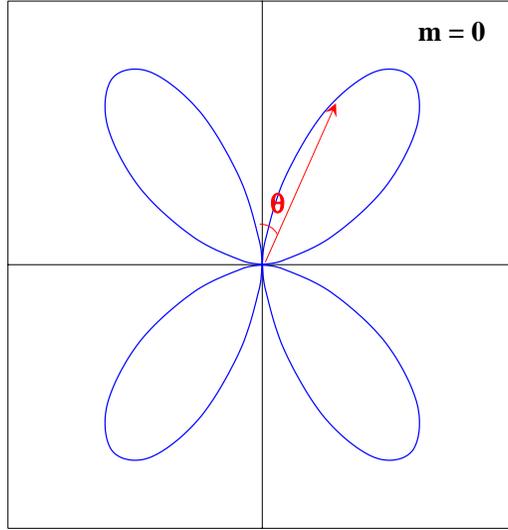
Figure 23: The angular distribution of quadrupole radiation for the $\Delta m = 0$ transition, as a function of polar angle θ_k.

11.1.9 Two-photon decay of the $2s$ state of Hydrogen.

The $2s$ state of the hydrogen atom can not decay via the paramagnetic interaction since, as it can be shown that the matrix elements that govern the emission intensity vanish

$$< 1s \mid \hat{\epsilon}_\alpha(\underline{k}) . \hat{\underline{p}} \, \exp\left[- i \, \underline{k} . \underline{r} \right] \mid 2s > = 0 \qquad (902)$$

First on integrating by parts, the matrix elements can be written as

$$i \, \hbar \quad < 1s \mid \hat{\epsilon}_\alpha(\underline{k}) . \underline{\nabla} \, \exp\left[- i \, \underline{k} . \underline{r} \right] \mid 2s >$$

$$= + i \, \hbar \quad < 2s \mid \quad \exp\left[+ i \, \underline{k} . \underline{r} \right] \hat{\epsilon}_\alpha(\underline{k}) . \underline{\nabla} \mid 1s >^*$$

$$\qquad (903)$$

On utilizing the expression for the $1s$ wave function

$$\psi_{1s}(r) = \frac{1}{\sqrt{\pi \, a^3}} \, \exp\left[- \frac{r}{a} \right] \qquad (904)$$

one finds that

$$\underline{\nabla} \psi_{1s}(r) = - \frac{\underline{r}}{r} \, \frac{\psi_{1s}(r)}{a} \qquad (905)$$

151

Hence, the transition matrix element is given by

$$- i \ \frac{\hbar}{a} \ \int_0^\infty \ d^3\underline{r} \ \psi_{2s}(r) \ \exp \left[\ - \ i \ \underline{k} \ . \ \underline{r} \ \right] \ \frac{\hat{\epsilon}_\alpha(\underline{k}) \ . \ \underline{r}}{r} \ \psi_{1s}^*(r) \qquad (906)$$

The matrix elements can be simply evaluated in spherical polar coordinates if one chooses the direction of $\underline{k}$ as the polar axis. The plane-wave, therefore, only depends on the z-component of $\underline{r}$ and since

$$\hat{\epsilon}_\alpha(\underline{k}) \ . \ \underline{k} \ = \ 0 \qquad (907)$$

the factor $\hat{\epsilon}_\alpha(\underline{k}) \ . \ \underline{r}$ only depends on x and y and is antisymmetric with respect to the transformations $x \ \rightarrow \ - \ x$ and $y \ \rightarrow \ - \ y$. All other factors are even functions of x and y. On integrating over the directions in the $x - y$ plane, one finds that the integral is identically zero.

The above result could have been (partially) anticipated by considering the selection rules. The electric dipole transition is forbidden by parity. The magnetic dipole transition is zero in this non-relativistic treatment. All magnetic and electric quadrupole and higher multipole transitions are forbidden by angular momentum conservation.

The $2s$ state decays via two-photon emission which is described by the diamagnetic interaction and by the effect of the paramagnetic interaction taken to second-order in time-dependent perturbation theory. Since only the part of the paramagnetic interaction that creates a photon is involved, for our purposes the paramagnetic interaction can be replaced by

$$\hat{H}_{para} \ \rightarrow \ - \ \frac{q}{m \ c} \ \sum_{\underline{k},\alpha} \ \left(\frac{2 \ \pi \ \hbar \ c^2}{V \ \omega_k} \right)^{\frac{1}{2}} \ \hat{\underline{p}} \ . \ \hat{\epsilon}_\alpha(\underline{k}) \ a_{\underline{k},\alpha}^\dagger \ \exp \left[\ - \ i \ \underline{k} \ . \ \underline{r} \ \right] \qquad (908)$$

Likewise, the diamagnetic interaction can be replaced by
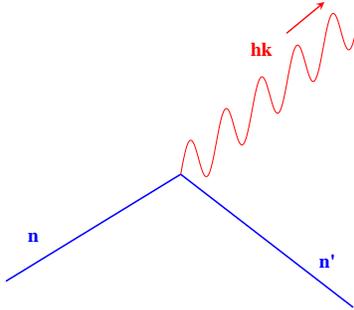
Figure 24: The one-photon emission part of the paramagnetic interaction.

$$
\hat{H}_{dia} \rightarrow \frac{q^2}{2\,m\,c^2} \sum_{\underline{k},\alpha;\underline{k}',\alpha'} \left(\frac{2\,\pi\,\hbar\,c^2}{V} \right) \frac{\hat{\epsilon}_\alpha(\underline{k})\,.\,\hat{\epsilon}_{\alpha'}(\underline{k}')}{\sqrt{\omega_k\,\omega_{k'}}}\, a^\dagger_{\underline{k},\alpha}\, a^\dagger_{\underline{k}',\alpha'} \, \exp\left[-\,i\,(\,\underline{k}+\underline{k}'\,)\,.\,\underline{r} \right]
$$

$$(909)$$

for two-photon emission. The system is assumed to be initially in an eigenstate
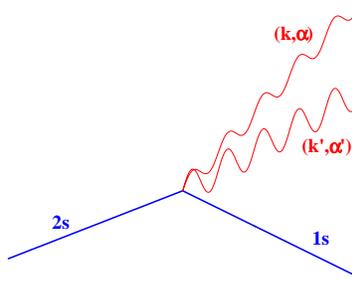
Figure 25: The two-photon emission part of the diamagnetic interaction.

of the unperturbed Hamiltonian $\mid n >$ but, due to the interaction $\hat{H}_{int}$ makes transitions to states $\mid n' >$. These states are to be considered as products of the electronic states and the states of the electromagnetic cavity. Following the usual procedure of time-dependent perturbation theory, the above state $\mid \psi_n >$ can be decomposed in terms of a complete set of non-interacting energy eigenstates $\mid n >$ via

$$
\mid \psi_n > \; = \; \sum_{n'} C_{n'}(t)\,\exp\left[-\,\frac{i}{\hbar}\,E_{n'}\,t \right]\,\mid n' >
$$

$$(910)$$

where $C_{n'}(t)$ are time-dependent coefficients. The probability of finding the system in the final state $\mid n' >$ at time t is then given by $|C_{n'}(t)|^2$. The rate at which the transition $n \rightarrow n'$ occurs is then given by the time-derivative of $|C_{n'}(t)|^2$.

It shall be assumed that the interaction is slowly turned on when $t \rightarrow -\infty$. The interaction can be reduced to zero at large negative times by introducing a multiplicative factor of $\exp[\,+\,\eta\,t\,]$ in the interaction, where η is an infinitesimally small positive constant. To first-order in the diamagnetic interaction, one finds

$$
C^{(1)}_{n'}(t) \;=\; \left(\frac{-\,i}{\hbar} \right)\, <n'\mid \exp[\,-\,i\,(\,\underline{k}'\,+\,\underline{k}\,)\,.\,\underline{r}\,]\mid n >\, \left(\frac{q^2}{2\,m\,c^2} \right) \left(\frac{2\,\pi\,\hbar\,c^2}{\sqrt{\omega_k\,\omega_{k'}}\,V} \right)
$$

$$
\times\,2\,\hat{\epsilon}_\alpha(\underline{k})\,.\,\hat{\epsilon}_{\alpha'}(\underline{k}') \int_{-\infty}^{t} dt'\,\exp\left[\frac{i}{\hbar}\,(\,\hbar\,\omega'\,+\,\hbar\,\omega\,+\,E_{n'}\,-\,E_n\,)\,t' \right]
$$

$$(911)$$

where $\omega \,=\, c\,k$ and $\omega' \,=\, c\,k'$ are the energies of the two photons in the final state. It should be noted that, since we have evaluated the photonic part of the

initial and final state, the labels n and n', now only describe the electronic part of the inital and final states. The small quantity η has been absorbed as a small imaginary part to the initial state energy

$$E_n \ \rightarrow \ E_n \ + \ i \, \eta \, \hbar \tag{912}$$

The paramagnetic interaction is of order of q and the diamagnetic interaction is of order q^2. Thus, to second-order in q, one must include the diamagnetic interaction and the paramagnetic interaction to second-order. There are two second-order terms which represent:

(a) emission of the photon $(\underline{k}, \alpha)$ followed by the emission of a photon $(\underline{k}', \alpha')$.

(b) emission of a photon $(\underline{k}', \alpha')$ followed by the emission of the photon $(\underline{k}, \alpha)$.
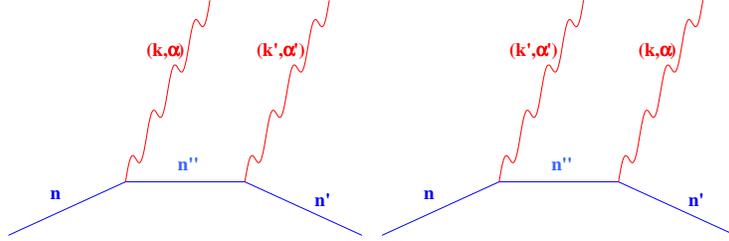
Figure 26: The two-photon emission processes to the paramagnetic interaction to second-order.

The second-order contribution to the transition amplitude is given by

$$
\begin{aligned}
C_{n'}^{(2)}(t) \ = \ & \left(\frac{-i}{\hbar} \right)^2 \left(\frac{q}{m \, c} \right)^2 \left(\frac{2 \, \pi \, \hbar \, c^2}{\sqrt{\omega_k \, \omega_{k'}} \, V} \right) \int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt'' \\
& \left[\sum_{n''} \exp[\, \frac{i}{\hbar} \, (E_{n'} + \hbar \omega' - E_{n''}) \, t' \,] \, \exp[\, \frac{i}{\hbar} \, (E_{n''} + \hbar \omega - E_n) \, t'' \,] \right. \\
& \times \ < \, n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \, . \, \hat{\underline{p}} \mid n''l''m'' \, > \, < \, n''l''m'' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \, . \, \hat{\underline{p}} \mid nlm \, > \\
& + \sum_{n''} \exp[\, \frac{i}{\hbar} \, (E_{n'} + \hbar \omega - E_{n''}) \, t' \,] \, \exp[\, \frac{i}{\hbar} \, (E_{n''} + \hbar \omega' - E_n) \, t'' \,] \\
& \times \ < \, n'l'm' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \, . \, \hat{\underline{p}} \mid n''l''m'' \, > \, < \, n''l''m'' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \, . \, \hat{\underline{p}} \mid nlm \, > \left. \right]
\end{aligned} \tag{913}
$$

The earliest time integration can be evaluated leading to

$$
C_{n'}^{(2)}(t) \ = \ \left(\frac{-i}{\hbar} \right) \left(\frac{q}{m \, c} \right)^2 \left(\frac{2 \, \pi \, \hbar \, c^2}{\sqrt{\omega_k \, \omega_{k'}} \, V} \right) \int_{-\infty}^{t} dt' \sum_{n''} \exp[\, \frac{i}{\hbar} \, (\hbar \omega' + E_{n'} - E_n - \hbar \omega) \, t' \,]
$$

154

$$\left[\frac{< n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \hat{\underline{p}} \mid n''l''m'' > < n''l''m'' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \hat{\underline{p}} \mid nlm >}{(E_n - E_{n''} - \hbar\,\omega)} \right.$$
$$\left. + \frac{< n'l'm' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \hat{\underline{p}} \mid n''l''m'' > < n''l''m'' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \hat{\underline{p}} \mid nlm >}{(E_n - E_{n''} - \hbar\,\omega')} \right] \tag{914}$$

as long as the denominators are non-vanishing.

The coefficients $C_{n'}^{(1)}(t)$ and $C_{n'}^{(2)}(t)$ have the same type of time-dependence. The remaining integration over time yields

$$\left(\frac{-i}{\hbar} \right) \int_{-\infty}^{t} dt' \, \exp\left[\frac{i}{\hbar} \, (\hbar\omega' + \hbar\omega + E_{n'} - E_n - i\hbar\eta) \, t' \right]$$
$$= - \frac{\exp[\frac{i}{\hbar} \, (\hbar\omega' + \hbar\omega + E_{n'} - E_n - i\hbar\eta) \, t]}{(\hbar\omega' + \hbar\omega + E_{n'} - E_n - i\hbar\eta)} \tag{915}$$

The transition rate is given by

$$\frac{1}{\tau} = \frac{\partial}{\partial t}\left(\left| C_{n'}^{(1)}(t) + C_{n'}^{(2)}(t) \right|^2 \right) \tag{916}$$

but the time-dependence of the squared modulus is contained in the common factor

$$\left| \frac{\exp[\frac{i}{\hbar} \, (\hbar\omega' + \hbar\omega + E_{n'} - E_n - i\hbar\eta) \, t]}{(\hbar\omega' + \hbar\omega + E_{n'} - E_n - i\hbar\eta)} \right|^2 = \frac{\exp\left[2\,\eta\,t \right]}{(\hbar\omega' + \hbar\omega + E_{n'} - E_n)^2 + \hbar^2\eta^2} \tag{917}$$

Since the momenta and polarizations of the emitted photons are not measured, the rate is summed over $(\underline{k}, \alpha)$ and $(\underline{k}', \alpha')$. Therefore, the transition rate is given by the expression

$$\frac{1}{\tau} = \sum_{\underline{k},\alpha:\underline{k}',\alpha'} \frac{2\,\eta\,\exp\left[2\,\eta\,t \right]}{(\hbar\omega' + \hbar\omega + E_{n'} - E_n)^2 + \hbar^2\eta^2} \, M^2 \tag{918}$$

where the matrix elements M are due to the combined effect of the diamagnetic interaction and the paramagnetic interaction taken to second-order. That is,

$$M = < 1s \, \underline{k}\alpha, \underline{k}'\alpha' \mid \hat{H}_{dia} \mid 2s >$$
$$+ \sum_{n''l''m''} \frac{< 1s \, \underline{k}\,\alpha, \underline{k}'\alpha' \mid \hat{H}_{para} \mid n''l''m'' \, \underline{k}'\alpha' > < n''l''m'' \, \underline{k}'\alpha' \mid \hat{H}_{para} \mid 2s >}{E_{2s} - E_{n''l''m''} - \hbar\omega_{k'}}$$
$$+ \sum_{n''l''m''} \frac{< 1s \, \underline{k}\,\alpha, \underline{k}'\alpha' \mid \hat{H}_{para} \mid n''l''m'' \, \underline{k}\alpha > < n''l''m'' \, \underline{k}\alpha \mid \hat{H}_{para} \mid 2s >}{E_{2s} - E_{n''l''m''} - \hbar\omega_k}$$
$$\tag{919}$$

These three terms add coherently, and it should be noted that the intermediate state is only a virtual state and it can have a higher-energy than the $2s$ state[46]. In the limit $\eta \rightarrow 0$ the first term in the expression for the transition rate of eqn(918) reduces to a delta function which expresses conservation of energy between the initial and final states.

$$
\lim_{\eta \rightarrow 0} \frac{1}{\pi} \frac{\eta \, \exp\left[2 \, \eta \, t \right]}{(\hbar\omega' + \hbar\omega + E_{n'} - E_n)^2 \, + \, \hbar^2\eta^2} \; = \; \delta(\, E_{2s} \, - \, E_{1s} \, - \, \hbar\omega_{k'} \, - \, \hbar\omega_k \,)
$$
$$(920)$$

In the limit $\eta \rightarrow 0$ the transition rate reduces to the Fermi-Golden rule expression

$$
\frac{1}{\tau} \; = \; \frac{2\pi}{\hbar} \sum_{\underline{k},\alpha:\underline{k}',\alpha'} | \, M \, |^2 \, \delta(\, E_{2s} \, - \, E_{1s} \, - \, \hbar\omega_{k'} \, - \, \hbar\omega_k \,) \qquad (921)
$$

The emitted photons have continuous spectra. In the expression for the matrix elements M, the last two terms differ in the time-order that the two photons are emitted. On inserting the expressions for the interactions into M, one can pull out the common factors leaving a dimensionless matrix element M'. This leads to the expression

$$
M \; = \; \left(\frac{q^2}{2 \, m \, c^2} \right) \left(\frac{2 \, \pi \, \hbar \, c^2}{V} \right) \frac{1}{\sqrt{\omega_k \, \omega_{k'}}} \; M' \qquad (922)
$$

where M' is the dimensionless factor given by

$$
\begin{aligned}
M' \;\; = \;\; & \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \; < \; 1s \mid \, \exp\left[\, - \, i \, (\, \underline{k} \, + \, \underline{k}' \,) \, . \, \underline{r} \, \right] \mid 2s \; > \\
& + \; \frac{2}{m} \sum_{n''l''m''} \frac{< \, 1s \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\underline{p}} \, \exp[\, - \, i \, \underline{k} \, . \, \underline{r} \,] \mid n''l''m'' \; > \; < \, n''l''m'' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \, . \, \hat{\underline{p}} \, \exp[\, - \, i \, \underline{k}' \, . \, \underline{r} \,] \mid 2s \; >}{E_{2s} \; - \; E_{n''l''m''} \; - \; \hbar\omega_{k'}} \\
& + \; \frac{2}{m} \sum_{n''l''m''} \frac{< \, 1s \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \, . \, \hat{\underline{p}} \, \exp[\, - \, i \, \underline{k}' \, . \, \underline{r} \,] \mid n''l''m'' \; > \; < \, n''l''m'' \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\underline{p}} \, \exp[\, - \, i \, \underline{k} \, . \, \underline{r} \,] \mid 2s \; >}{E_{2s} \; - \; E_{n''l''m''} \; - \; \hbar\omega_k}
\end{aligned}
$$
$$(923)$$

The first term is negligible, since $1 \gg \underline{k} \, . \, \underline{r}$ and the electronic eigenstates are orthogonal. The order of magnitude of the second term is given by the electronic kinetic energy divided by the excitation energy. Hence, the reduced matrix elements have a magnitude of the order of unity. The transition rate is given by

$$
\frac{1}{\tau} \; = \; \left(\frac{e^2}{2 \, m \, c^2} \right)^2 \frac{(\, 2 \, \pi \,)^3}{V^2} \sum_{\underline{k},\alpha;\underline{k}',\alpha'} \frac{\hbar \, c^2}{k \, k'} \, |M'|^2 \, \delta(\, E_{2s} \, - \, E_{1s} \, - \, \hbar\omega_k \, - \, \hbar\omega_{k'} \,) \quad (924)
$$

[46]Due to the Lamb shift, there is a $2p$ state with slightly lower energy than the $2s$ state. However, due to the small magnitude of the energy difference, the part of the decay process involving any real $2p$ transition is negligibly small.

One can assume that the dipole matrix elements of the intermediate states should be randomly oriented in space, since the initial and final electronic states are isotropic. After summing over the polarizations, the transition rate becomes isotropic. On setting

$$\sum_{\alpha,\alpha'} \mid M' \mid^2 \approx 1 \tag{925}$$

one finds

$$\frac{1}{\tau} = \left(\frac{e^2}{2\,m\,c^2} \right)^2 \frac{\hbar\,c^2}{(\,2\,\pi\,)^3} \int \frac{d^3\underline{k}}{k} \int \frac{d^3\underline{k}'}{k'} \delta(\,E_{2s}\,-\,E_{1s}\,-\,\hbar\omega_k\,-\,\hbar\omega_{k'}\,) \tag{926}$$

Since the integrand is independent of the direction of k and k', the angular integrations can be performed leaving

$$\frac{1}{\tau} = \left(\frac{e^2}{m\,c^2} \right)^2 \frac{\hbar\,c^2}{2\,\pi} \int_0^\infty dk\,k \int_0^\infty dk'\,k'\,\delta(\,E_{2s}\,-\,E_{1s}\,-\,\hbar\omega_k\,-\,\hbar\omega_{k'}\,) \tag{927}$$

On integrating over k', one obtains

$$\frac{1}{\tau} = \left(\frac{e^2}{m\,c^2} \right)^2 \frac{c}{2\,\pi} \int_0^{\frac{\omega}{c}} dk\,k\,(\,\frac{\omega_{21}}{c}\,-\,k\,) \tag{928}$$

where ω_{12} is related to the energy difference of the $1s$ and $2s$ states. An elementary integration yields

$$\frac{1}{\tau} = \left(\frac{e^2}{m\,c^2} \right)^2 \frac{c}{12\,\pi} \left(\frac{\omega_{12}}{c} \right)^3 \tag{929}$$

The first factor has dimensions of length squared and can be recognized as the square of the classical radius of the electron. However, since

$$\frac{\omega_{12}}{c} = \frac{3}{8} \frac{e^2}{\hbar\,c\,a} \tag{930}$$

and

$$a = \frac{\hbar^2}{m\,e^2} \tag{931}$$

or

$$\frac{e^2}{m\,c^2} = \frac{a\,e^4}{\hbar^2\,c^2} \tag{932}$$

one finds the decay rate is approximated by

$$\frac{1}{\tau} = \frac{1}{12\,\pi} \left(\frac{3}{8} \right)^3 \frac{c}{a} \left(\frac{e^2}{\hbar\,c} \right)^7 \tag{933}$$

Thus, the estimated decay rate is 8.75 sec^{-1}. The exact value calculated by Shapiro and Breit[47] is 8.266 sec^{-1}.

---

[47] J. Shapiro and G. Breit, Phys. Rev. **113**, 179 (1959).

11.1.10 The Absorption of Radiation

If a process occurs in which only a photon with quantum numbers $(\underline{k}, \alpha)$ is absorbed, then the numbers of quanta in the initial and final state of the electromagnetic field are given by

$$
\begin{aligned}
n'_{\underline{k},\alpha} &= n_{\underline{k},\alpha} - 1 \\
n'_{\underline{k}',\beta} &= n_{\underline{k}',\beta}
\end{aligned}
\tag{934}
$$

The matrix elements of the paramagnetic interaction are given by

$$
< n'l'm' \, \{n'_{\underline{k}',\beta}\} \mid \hat{H}_{para} \mid nlm \, \{n_{\underline{k}',\beta}\} >
$$

$$
= - \left(\frac{q}{m \, c} \right) \sum_{\underline{k},\alpha} \sqrt{n}_{\underline{k},\alpha} \sqrt{\frac{2 \, \pi \, \hbar \, c^2}{V \, \omega_k}} < n'l'm' \mid \hat{\underline{p}} \cdot \hat{\epsilon}_\alpha(\underline{k}) \, \exp\left[+ i \, \underline{k} \cdot \underline{r} \right] \mid nlm >
\tag{935}
$$

The photon absorption rate is found from the Fermi-Golden rule expression

$$
\frac{1}{\tau} = \frac{2 \, \pi}{\hbar} \left(\frac{q}{m \, c} \right)^2 \left(\frac{2 \, \pi \, \hbar \, c^2}{V \, \omega_k} \right) n_{\underline{k},\alpha} \sum_{n'l'm'} \delta(\, E_{nlm} + \hbar\omega_k - E_{n'l'm'} \,)
$$

$$
\times \mid < n'l'm' \mid \hat{\underline{p}} \cdot \hat{\epsilon}_\alpha(\underline{k}) \, \exp\left[+ i \, \underline{k} \cdot \underline{r} \right] \mid nlm > \mid^2
\tag{936}
$$

This is related to the lifetime due to stimulated emission, if the initial and final states are interchanged.

The scattering cross-section for photon absorption $\sigma_{\text{absorb}}(\omega)$ is found by relating the number of photons absorbed (per second) to the product of the incident flux and the cross-section. The photon flux is given by the photon density times the velocity of light

$$
\underline{j} = \frac{n_{\underline{k},\alpha}}{V} \, c \, \hat{k}
\tag{937}
$$

Hence, the cross-section can be written as

$$
\sigma_{\text{absorb}}(\omega_k) = \left(\frac{V}{n_{\underline{k},\alpha} \, c} \right) \frac{2 \, \pi}{\hbar} \left(\frac{q}{m \, c} \right)^2 \left(\frac{2 \, \pi \, \hbar \, c^2}{V \, \omega_k} \right) n_{\underline{k},\alpha} \sum_{n'l'm'} \delta(\, E_{nlm} + \hbar\omega_k - E_{n'l'm'} \,)
$$

$$
\times \mid < n'l'm' \mid \hat{\underline{p}} \cdot \hat{\epsilon}_\alpha(\underline{k}) \, \exp\left[+ i \, \underline{k} \cdot \underline{r} \right] \mid nlm > \mid^2
\tag{938}
$$

which simplifies to

$$
\sigma_{\text{absorb}}(\omega_k) = \left(\frac{4 \, \pi^2 \, e^2}{m^2 \, \omega_k \, c} \right) \sum_{n'l'm'} \delta(\, E_{nlm} + \hbar\omega_k - E_{n'l'm'} \,)
$$

$$
\times \mid < n'l'm' \mid \hat{\underline{p}} \cdot \hat{\epsilon}_\alpha(\underline{k}) \, \exp\left[+ i \, \underline{k} \cdot \underline{r} \right] \mid nlm > \mid^2
\tag{939}
$$

The absorption cross-section is independent of the volume of the electromagnetic cavity and the number of photons in the incident beam. As a function of frequency, the Born approximation for the cross-section for photon absorption contains delta function lines corresponding to the atomic excitation energies. Measured absorption lines do have natural widths $\Delta\omega_{nl,n'l'}$ and the absorbtion spectra can be approximated by the sums of Lorentzian functions. The widths
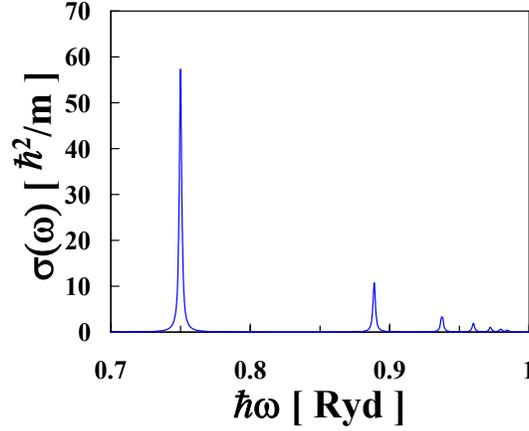
Figure 27: A sketch of the photon absorption cross-section $\sigma(\omega)$ (in units of $\frac{\hbar^2}{m}$) as a function of photon energy $\hbar\omega$ (in units of Rydbergs). The plot over-emphasizes the role of the photon lifetimes, since the ratio of the line-width to the photon frequency is of the order of $(\frac{e^2}{\hbar c})^3$.

of the lines are governed by half the sum of the decay rates of the initial and final electronic levels.

$$\Delta\omega_{nl,n'l'} \;=\; \frac{1}{2} \left(\frac{1}{\tau_{nl}} \;+\; \frac{1}{\tau_{n'l'}} \right) \tag{940}$$

This formula implies that rapidly decaying levels will yield broad lines, but does not imply the converse[48]. The spectral widths can be described by the inclusion of the effects of interaction to higher orders[49]. The higher-order processes produce small shifts of the atomic energy levels and also give the energies small imaginary parts, resulting in a Lorentzian line shape. Since a typical atomic transition rate is of the order of 10^8 sec^{-1} and a typical photon frequency is of the order of 10^{15} sec^{-1}, the widths of the lines can usually be neglected.

The absorption cross-section can be evaluated in the dipole approximation

$$\sigma_{\text{absorb}}(\omega_k) \;=\; \left(\frac{4 \, \pi^2 \, e^2}{m^2 \, \omega_k \, c} \right) \sum_{n'l'm'} \delta(\, E_{nlm} \;+\; \hbar\omega_k \;-\; E_{n'l'm'} \,)$$

---

[48]Lines in the absorption spectra with weak intensities can be broad if the final states are rapidly decaying.

[49]V. F. Weisskopf and E. Wigner, Z. Physik, **63**, 54 (1930), Zeit. für Physik, **65**, 18 (1930).

$$\times \ | \ < \ n'l'm' \ | \ \hat{\underline{p}} \cdot \hat{\epsilon}_\alpha(\underline{k}) \ | \ nlm \ > \ |^2 \tag{941}$$

which can be re-written as

$$\sigma_{\mathrm{absorb}}(\omega_k) \ = \ \left(\frac{4 \ \pi^2 \ e^2 \ \omega_k}{c} \right) \sum_{n'l'm'} \delta(\ E_{nlm} \ + \ \hbar\omega_k \ - \ E_{n'l'm'} \ )$$

$$\times \ | \ < \ n'l'm' \ | \ \underline{r} \cdot \hat{\epsilon}_\alpha(\underline{k}) \ | \ nlm \ > \ |^2 \tag{942}$$

For an isotropic medium, the electronic states are degenerate with respect to the z-components of the orbital angular momentum, so the initial state (n, l, m) should be averaged over the different values of m

$$\frac{1}{(2l+1)} \sum_{m=-l}^{l} \tag{943}$$

and the values of m' for the final states are summed over all possible values. This averaging process results in an isotropic absorption rate, and is equivalent to averaging the polarization vector over all directions in space. Therefore, in the dipole approximation, the absorption cross-section for an isotropic medium is given by the expression

$$\sigma_{\mathrm{absorb}}(\omega) \ = \ \frac{4 \ \pi^2}{3} \left(\frac{e^2}{\hbar \ c} \right) \sum_{n'l'm'} \omega_{n'l',nl} \ | \ < \ n'l'm' \ | \ r \ | \ nlm \ > \ |^2 \, \delta(\omega_{n'l',nl} - \omega \ ) \tag{944}$$

The strength of each absorption line can be found by integrating the cross-section over a narrow frequency range centered on the frequency of the absorption line. (More specifically, the width of the interval of integration must be greater than the natural line-width.) The integrated intensity of the transition $(nlm) \rightarrow (n'l'm')$ is given by

$$\int_{\omega_{nl,n'l'}-\epsilon}^{\omega_{nl,n'l'}+\epsilon} d\omega \ \sigma_{\mathrm{absorb}}(\omega) \ = \ \frac{4 \ \pi^2}{3} \left(\frac{e^2}{\hbar \ c} \right) \omega_{n'l',nl} \ | \ < \ n'l'm' \ | \ \underline{r} \ | \ nlm \ > \ |^2 \tag{945}$$

The intensity of each line is proportional to the "oscillator strength" $f_{nl \rightarrow n'l'}$ defined as

$$f_{nl \rightarrow n'l'} \ = \ \frac{2 \ m \ \omega_{n'l',nl}}{\hbar} \ | \ < \ n'l'm' \ | \ \underline{r} \ | \ nlm \ > \ |^2 \tag{946}$$

The intensities and the frequencies of all the transitions are related via sum rules[50]. These sum rules involve quantities of the form

$$\sum_{n'l'm'} \omega_{n'l'm',nlm}^p \ | \ < \ n'l'm' \ | \ \underline{r} \ | \ nlm \ > \ |^2 \tag{947}$$

Table 5: Sum Rules for Dipole Transitions

p	$\sum_{n'l'm'} \omega^p_{n'l'm',nlm} \mid < n'l'm' \mid \underline{r} \mid nlm > \mid^2$
0	$< nlm \mid r^2 \mid nlm >$
1	$\frac{3}{2} \frac{\hbar}{m}$
2	$\frac{2}{m} (E_{nlm} - < nlm \mid V \mid nlm >)$
3	$\frac{\hbar}{2\,m} < nlm \mid \nabla^2 V \mid nlm >$

and have values given in the Table(5). The sum rules can be used to provide checks of experimental data.

Sum Rules for Dipole Radiation

There exists a systematic way of deriving sum rules for the weighted intensities of the dipole allowed transitions. The sum rules are of the form

$$\sum_{n'l'm'} \omega^p_{nl,n'l'} \mid < nlm \mid \hat{A} \mid n'l'm' > \mid^2 \tag{948}$$

where

$$\hbar\, \omega_{nl,n'l'} = E_{nl} - E_{n'l'} \tag{949}$$

and p is a positive integer.

Consider the expression

$$F(t) = < nlm \mid \hat{A}(t)\, \hat{A}^\dagger(0) \mid nlm > \tag{950}$$

where the operator $\hat{A}(t)$ is given in the Heisenberg representation

$$\hat{A}(t) = \exp\left[+ \frac{i}{\hbar} \hat{H}_0\, t \right] \hat{A} \exp\left[- \frac{i}{\hbar} \hat{H}_0\, t \right] \tag{951}$$

---

[50]W. Thomas, Naturwiss. **11**, 527 (1925).
F. Reiche and W. Thomas, Zeit. für Physik, **34**, 510 (1925).
W. Kuhn, Zeit. für Physik, **33**, 408 (1925).

Then, on taking successive derivatives of $F(t)$ with respect to t, one finds

$$\left(\frac{\partial F}{\partial t} \right) = \frac{i}{\hbar} < nlm \mid [\, \hat{H}_0 \, , \, \hat{A}(t) \,] \, \hat{A}^\dagger(0) \mid nlm > \qquad (952)$$

and

$$\left(\frac{\partial^2 F}{\partial t^2} \right) = \left(\frac{i}{\hbar} \right)^2 < nlm \mid [\, \hat{H}_0 \, , \, [\, \hat{H}_0 \, , \, \hat{A}(t) \,] \,] \, \hat{A}^\dagger(0) \mid nlm > \qquad (953)$$

etc. This process shows that the p-th derivative is expressed as p nested commutators

$$\left(\frac{\partial^p F}{\partial t^p} \right) = \left(\frac{i}{\hbar} \right)^p < nlm \mid [\, \hat{H}_0 \, , \, [\, \ldots \, [\, \hat{H}_0 \, , \, [\, \hat{H}_0 \, , \, \hat{A}(t) \,] \,] \ldots \,] \,] \, \hat{A}^\dagger(0) \mid nlm >$$
$$(954)$$

Alternatively, one can insert a complete set of states in the definition for $F(t)$ yielding

$$F(t) = \sum_{n'l'm'} < nlm \mid \hat{A}(t) \mid n'l'm' > < n'l'm' \mid \hat{A}^\dagger(0) \mid nlm > \qquad (955)$$

but since the states $\mid nlm >$ are eigenstates of $\hat{H}_0$, one has

$$F(t) = \sum_{n'l'm'} \exp\left[i \, \omega_{nl,n'l'} \, t \right] \, \Big| < nlm \mid \hat{A} \mid n'l'm' > \Big|^2 \qquad (956)$$

On taking the p-th derivative of this form of $F(t)$, one finds

$$\left(\frac{\partial^p F}{\partial t^p} \right) = \sum_{n'l'm'} i^p \, \omega_{nl,n'l'}^p \, \exp\left[i \, \omega_{nl,n'l'} \, t \right] \, \Big| < nlm \mid \hat{A} \mid n'l'm' > \Big|^2 \qquad (957)$$

The sum rules are found by equating the two forms of the p-th time-derivative and then setting $t = 0$

$$\sum_{n'l'm'} \left(E_{nl} - E_{n'l'} \right)^p \, \Big| < nlm \mid \hat{A} \mid n'l'm' > \Big|^2$$
$$= < nlm \mid [\, \hat{H}_0 \, , \, [\, \ldots \, , \, [\, \hat{H}_0 \, , \, [\, \hat{H}_0 \, , \, \hat{A} \,] \,] \,] \ldots \,] \,] \, \hat{A}^\dagger \mid nlm > (958)$$

Hence, the p-th moment of the matrix elements of $\hat{A}$ is related to the expectation value of the product of the p-th nested commutator of $\hat{H}_0$ and $\hat{A}$ multiplied by $\hat{A}^\dagger$.

The expectation value of p nested commutators of $\hat{A}$ can be expressed as the expectation value of the product of $p-q$ nested commutators of $\hat{A}$ with q nested commutators of $\hat{A}^\dagger$. This can be demonstrated by noting that the expectation

value is homogeneous in time. The q-th nested commutator of the operator $\hat{A}$ can be defined by

$$\hat{B}_q = [\, \hat{H}_0 \,, \, \hat{B}_{q-1} \,] \tag{959}$$

where

$$\hat{B}_0 = \hat{A} \tag{960}$$

Likewise, $\hat{C}_q$ can be defined as the q-th nested commutator of $\hat{A}^\dagger$. However, for any pair of operators $\hat{B}_{p-q-1}$ and $\hat{C}_q$, one has

$$
\begin{aligned}
&< \, nlm \mid \hat{B}_{p-q-1}(t) \; \hat{C}_q(0) \mid nlm \, > \\
&= \sum_{n'l'm'} \exp\left[\, i \, \omega_{nl,n'l'} \, t \,\right] \; < \, nlm \mid \hat{B}_{p-q-1} \mid n'l'm' \, > < \, n'l'm' \mid \hat{C}_q \mid nlm \, > \\
&= \; < \, nlm \mid \hat{B}_{p-q-1}(0) \; \hat{C}_q(-t) \mid nlm \, >
\end{aligned}
\tag{961}
$$

This is an expression of the homogeneity of time. Hence, on taking a derivative with respect to t and then setting $t = 0$, one finds

$$< \, nlm \mid [\, \hat{H}_0 \,, \, \hat{B}_{p-q-1} \,] \, \hat{C}_q \mid nlm \, > \; = \; (\, -\, 1\,) \; < \, nlm \mid \hat{B}_{p-q-1} \, [\, \hat{H}_0 \,, \hat{C}_q \,] \mid nlm \, > \tag{962}$$

On using the definition of the operators $\hat{B}_p$ and C_q, the above equation reduces to

$$< \, nlm \mid \hat{B}_{p-q} \, \hat{C}_q \mid nlm \, > \; = \; (\, -\, 1\,) \; < \, nlm \mid \hat{B}_{p-q-1} \, \hat{C}_{q+1} \mid nlm \, > \tag{963}$$

By induction, this shows that the nested commutators can be distributed between the two sides of the expression.

$$< \, nlm \mid \hat{B}_p \, \hat{C}_0 \mid nlm \, > \; = \; (\, -\, 1\,)^q \; < \, nlm \mid \hat{B}_{p-q} \, \hat{C}_q \mid nlm \, > \tag{964}$$

which was to be shown.

11.1.11 The Photoelectric Effect

The differential scattering cross-section for the absorption of a photon by a hydrogen atom in the ground state accompanied by the emission of an electron shall be derived. For emitted electrons with sufficiently high energies, the wave function for the photo-emitted electron can be approximated by a plane-wave. The transition rate is given by the Fermi-Golden rule expression involving the paramagnetic interaction

$$\frac{1}{\tau} = \frac{2\,\pi}{\hbar} \left(\frac{e}{m\,c}\right)^2 \left(\frac{2\,\pi\,\hbar\,c^2}{V\,\omega_k}\right) n_{\underline{k},\alpha} \sum_{\underline{k}'} \mid < \, \underline{k}' \mid \underline{p} \,.\, \hat{\epsilon}_\alpha(\underline{k}) \, \exp\left[\, i\,\underline{k}\,.\,\underline{r} \,\right] \mid 1s \, > \mid^2 \delta(\, E_{1s} + \hbar\,\omega_k - \frac{\hbar^2\,k'^2}{2\,m} \,) \tag{965}$$

The cross-section is given by

$$
\sigma = \sum_{\underline{k}'} \left(\frac{4 \pi^2 e^2}{m^2 \omega_k c} \right) | < \underline{k}' | \underline{p} \cdot \hat{\epsilon}_\alpha(\underline{k}) \exp \left[i \underline{k} \cdot \underline{r} \right] | 1s > |^2 \delta(E_{1s} + \hbar \omega_k - \frac{\hbar^2 k'^2}{2 m})
$$

(966)

where the initial wave function is given by

$$
\psi_{1s}(r) = \frac{1}{\sqrt{\pi a^3}} \exp \left[- \frac{r}{a} \right]
$$

(967)

As long as the emitted electron is not close to threshold, the final state wave function can be approximated by a plane-wave

$$
\psi_{\underline{k}'}(\underline{r}) = \frac{1}{\sqrt{V}} \exp \left[i \underline{k}' \cdot \underline{r} \right]
$$

(968)

The sum over final states of the electron can be replaced by an integral over the magnitude of its momentum and its direction

$$
\sum_{\underline{k}'} \rightarrow \frac{V}{(2 \pi)^3} \int_0^\infty dk' \, k'^2 \int d\Omega_{k'}
$$

(969)

It is seen that the factor of the volume in the density of final states cancels with the factors from the normalization of the electron's final state. The differential cross-section corresponds to the part of the cross-section where the outgoing electron is emitted into the solid angle $d\Omega_{k'}$. Hence,

$$
\frac{d\sigma}{d\Omega'} = \frac{V}{2 \pi} \left(\frac{e^2}{m^2 \omega_k c} \right) \int_0^\infty dk' \, k'^2 | < \underline{k}' | \underline{p} \cdot \hat{\epsilon}_\alpha(\underline{k}) \exp \left[i \underline{k} \cdot \underline{r} \right] | 1s > |^2 \delta(E_{1s} + \hbar \omega_k - \frac{\hbar^2 k'^2}{2 m})
$$

(970)

The integration over the magnitude of electron's final momentum k' can be performed by using the properties of the energy conserving delta function. The magnitude of electron's final momentum is denoted by k_f

$$
k_f^2 = \frac{2 m}{\hbar^2} (\hbar \omega_k + E_{1s})
$$

(971)

The result of the integration over k' is

$$
\frac{d\sigma}{d\Omega'} = \frac{V}{2 \pi \hbar^2} \left(\frac{e^2}{m \omega_k c} \right) k_f | < k_f \, d\Omega' | \underline{p} \cdot \hat{\epsilon}_\alpha(\underline{k}) \exp \left[i \underline{k} \cdot \underline{r} \right] | 1s > |^2
$$

(972)

It is assumed that the initial photon is propagating along the x-axis and is polarized along the z-direction. The matrix elements involving the momentum operator only yield a finite result when $\hat{p}$ acts on $\psi_{1s}(\underline{r})$, since $\underline{k} \cdot \hat{\epsilon}_\alpha(\underline{k}) = 0$. However,

$$
\hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{p} \, \psi_{1s}(r) = i \frac{\hbar \, \cos\theta}{a} \psi_{1s}(r)
$$

(973)

164
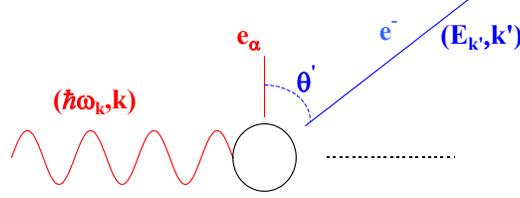
Figure 28: The geometry for the photo-emission of an electron from an atom. An electromagnetic wave, with polarization along the z-axis, is incident along the x-axis. The photo-emitted electron propagates along the direction $\underline{k}'$.

which results in the replacement $\hat{p}_z \rightarrow i \frac{\hbar \, \cos\theta}{a}$. Thus, one finds

$$\frac{d\sigma}{d\Omega'} = \frac{V}{2\,\pi} \left(\frac{e^2}{m\,\omega_k\,c\,a^2} \right) k_f \mid \, < \, k_f \, d\Omega' \mid \cos\theta \, \exp\left[i \, \underline{k} \cdot \underline{r} \right] \mid 1s \, > \mid^2 \quad (974)$$

where (θ, φ) are the polar coordinates of the vector $\underline{r}$. The matrix elements are evaluated using the dipole approximation for the photon wave function and set

$$\exp\left[i \, \underline{k} \cdot \underline{r} \right] \approx 1 + i \, \underline{k} \cdot \underline{r} + \dots \quad (975)$$

and only keep the first term of the expansion. The factor $\cos\theta$ can be expressed as a spherical harmonic through

$$\cos\theta = \sqrt{\frac{4\,\pi}{3}} \, Y_0^1(\theta, \varphi) \quad (976)$$

and the final state electronic wave function can be expressed in terms of the Rayleigh expansion

$$\exp\left[i \, k_f \, \hat{k}' \cdot \underline{r} \right] = 4\,\pi \sum_{l,m} i^l \, j_l(\, k_f \, r \,) \, Y_m^l(\theta, \varphi) \, Y_m^{l\,*}(\theta', \varphi') \quad (977)$$

where (θ', φ') are the polar coordinates of the electron's final momentum. The angular integration over the polar coordinates (θ, φ) can be performed by using the orthogonality relations for the spherical harmonics. The end result is

$$< \, k_f \, d\Omega' \mid \cos\theta \mid 1s \, > = - \, 4\,\pi\,i \, \cos\theta' \, \frac{1}{\sqrt{\pi \, a^3 \, V}} \int_0^\infty dr \, r^2 \, j_1(k_f r) \, \exp\left[-\frac{r}{a} \right] \quad (978)$$

where the $\cos\theta'$ dependence refers to the direction of the emitted electron's angular momentum. The radial integral is evaluated to yield

$$< \, k_f \, d\Omega' \mid \cos\theta \mid 1s \, > = - \, 4\,\pi\,i \, \cos\theta' \, \sqrt{\frac{a^3}{\pi \, V}} \, \frac{2 \, k_f \, a}{(\, 1 + k_f^2 \, a^2 \,)^2} \quad (979)$$

165

Therefore, the differential cross-section is given by

$$\left(\frac{d\sigma}{d\Omega'} \right) = 8 \left(\frac{k_f}{k} \right) \left(\frac{e^2 \, a}{m \, c^2} \right) \cos^2 \theta' \left(\frac{2 \, k_f \, a}{(\, 1 \, + \, k_f^2 \, a^2 \,)^2} \right)^2 \qquad (980)$$

Using

$$a = \frac{\hbar^2}{m \, e^2} \qquad (981)$$

the photo-emission cross-section can be re-written as

$$\left(\frac{d\sigma}{d\Omega'} \right) = 8 \left(\frac{k_f}{k} \right) a^2 \left(\frac{e^2}{\hbar \, c} \right)^2 \cos^2 \theta' \left(\frac{2 \, k_f \, a}{(\, 1 \, + \, k_f^2 \, a^2 \,)^2} \right)^2 \qquad (982)$$

Thus, although the photon is propagating along the x-direction, the electron is preferentially emitted along the direction of the polarization ($\theta' \approx 0$). This can be understood as being due to the effect of c being large, so that the photon's momentum is negligible compared with the energy, therefore, (in the dipole approximation) only the direction of the polarization determines the angular distribution of the emitted electron. It should be noted that in the relativistic case, where the momentum of the photon is important, the electrons are predominantly ejected in the direction of the photon[51]. This formula also breaks down for emitted electrons with low energies. In this case, the correct electronic wave function for the continuous spectrum of $\hat{H}_0$ should be used[52]. The inclusion of the Coulomb attraction of the ion in the final state has the effect of reducing the cross-section near the threshold.

11.1.12 Impossibility of absorption of photons by free-electrons.

Free electrons are described by the non-interacting Hamiltonian $\hat{H}_0$ where

$$\hat{H}_0 = \frac{\hat{p}^2}{2 \, m} \qquad (983)$$

which plane-waves as energy eigenstates

$$\phi_{k'}(\underline{r}) = \frac{1}{\sqrt{V}} \exp\left[- \, i \, \underline{k}' . \, \underline{r} \right] \qquad (984)$$

corresponding to the energy eigenvalues

$$E_{k'} = \frac{\hbar^2 \, k'^2}{2 \, m} \qquad (985)$$

The matrix elements for electromagnetic transitions in which a photon $(\underline{k}, \alpha)$ is absorbed is given by

$$- \left(\frac{q}{m \, c} \right) \sqrt{\frac{2 \, \pi \, \hbar \, c}{V \, \omega_k}} \ < \ \underline{k}'' \, | \, \hat{\underline{p}} . \, \hat{\epsilon}_\alpha(\underline{k}) \ \exp\left[+ \, i \, \underline{k} . \, \underline{r} \right] | \, \underline{k}' > \ \sqrt{n_{\underline{k},\alpha}} \quad (986)$$

---

[51]F. Sauter, Ann. Phys. **9**, 217 (1931), Ann. Phys. **11**, 454 (1931).
[52]M. Stobbe, Ann. Phys. **7**, 661 (1930).

which is evaluated as

$$
- \left(\frac{q}{m \, c} \right) \, \sqrt{\frac{2 \, \pi \, \hbar \, c}{V \, \omega_k}} \, \underline{\hat{p}} \, . \, \hat{\epsilon}_\alpha(\underline{k}) \, \delta_{\underline{k}+\underline{k}'-\underline{k}''} \, \sqrt{n}_{\underline{k},\alpha} \tag{987}
$$

This shows that momentum is conserved. Furthermore, for the transition rate
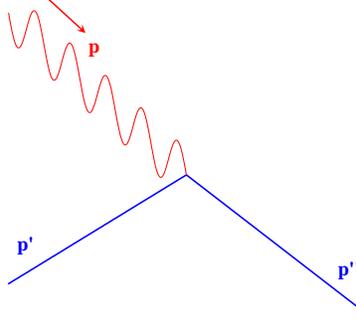
Figure 29: The absorption of a photon via the paramagnetic interaction.

to represent a real process, it is necessary that energy is conserved between the initial and final states

$$
\hbar \, \omega_k \; + \; \frac{\hbar^2 \, k'^2}{2 \, m} \; = \; \frac{\hbar^2 k''^2}{2 \, m} \tag{988}
$$

It is impossible for this process to satisfy the conditions for conservation of energy and momentum. This can be seen by appealing to the relativistic formulation where the four-vector momentum is conserved

$$
p_\mu \; + \; p'_\mu \; = \; p''_\mu \tag{989}
$$

Hence,

$$
(\, p^\mu \; + \; p^{\mu \prime} \,) \, (\, p_\mu \; + \; p'_\mu \,) \; = \; p^{\mu \prime\prime} \, p''_\mu \tag{990}
$$

but the electron's momenta form a Lorentz scalar which is related to the rest mass

$$
p^{\mu \prime} \, p'_\mu \; = \; p^{\mu \prime\prime} \, p''_\mu \; = \; m^2 \, c^2 \tag{991}
$$

and the photon has zero mass

$$
p^\mu \, p_\mu \; = \; 0 \tag{992}
$$

Therefore, one finds that the cross-terms vanish

$$
p^\mu \, p'_\mu \; = \; 0 \tag{993}
$$

In the rest frame of the electron one has $p_\mu{}' \; = \; (\, m \, c \, , \, 0 \,)$, so the energy of the photon is identically zero. Therefore, there is no photon and the absorption process is impossible.

167

11.2 Scattering of Light

Kramers and Heisenberg evaluated the scattering cross-section for light incident on atomic electrons[53]. The incident photon is denoted by $(\underline{k}, \alpha)$ and the scattered photon by $(\underline{k}', \alpha')$. The scattering cross-section involves the paramagnetic interaction to second-order and the diamagnetic interaction to first-order. The matrix elements of the diamagnetic interaction are given by

$$
< n'l'm'\underline{k}'\alpha' \mid \hat{H}_{dia} \mid nlm\underline{k}\alpha > \;=\; \left(\frac{e^2}{2\, m\, c^2} \right) \; < n'l'm'\underline{k}'\alpha' \mid \underline{\hat{A}} \,.\, \underline{\hat{A}} \mid nlm\underline{k}\alpha >
$$

$$
= \left(\frac{e^2}{2\, m\, c^2} \right) \; < n'l'm'\underline{k}'\alpha' \mid (a_{\underline{k},\alpha} a^\dagger_{\underline{k}',\alpha'} \;+\; a^\dagger_{\underline{k}',\alpha'} a_{\underline{k},\alpha}) \; \exp\left[\, i\, (\, \underline{k}\, -\, \underline{k}'\,)\,.\, \underline{r} \, \right] \mid nlm\underline{k}\alpha >
$$

$$
\times \left(\frac{2\,\pi\,\hbar\, c^2}{\sqrt{\omega_k\,\omega_{k'}}\, V} \right) \hat{\epsilon}_\alpha(\underline{k}) \,.\, \hat{\epsilon}_{\alpha'}(\underline{k}') \tag{994}
$$

where it has been assumed that only the initial and final photon are present. On making use of the long-wavelength approximation $\lambda \gg a$, the matrix elements simplify to

$$
< n'l'm'\underline{k}'\alpha' \mid \hat{H}_{dia} \mid nlm\underline{k}\alpha > \;\approx\; \left(\frac{e^2}{2\, m\, c^2} \right) \; < n'l'm' \mid nlm >
$$

$$
\times \left(\frac{2\,\pi\,\hbar\, c^2}{\sqrt{\omega_k\,\omega_{k'}}\, V} \right) \hat{\epsilon}_\alpha(\underline{k}) \,.\, \hat{\epsilon}_{\alpha'}(\underline{k}') \tag{995}
$$

The scattering cross-section will be expressed in terms of a transition rate and the transition rate will be calculated using a similar procedure to that which was used in describing two-photon decay. An arbitrary state $\mid \psi_n >$ can be expressed in terms of a complete set of non-interacting states $\mid n >$

$$
\mid \psi_n > = \sum_{n'} C_{n'}(t) \; \exp\left[-\, \frac{i}{\hbar}\, E_{n'}\, t \right] \mid n' > \tag{996}
$$

where $C_n(t)$ are time-dependent coefficients. Initially, the system is assumed to be in an energy eigenstate $\mid n >$ of the unperturbed Hamiltonian, and due to the interaction makes a transition to a state $\mid n' >$. The probability of finding the system in the state $\mid n' >$ at time t is then given by $|C_{n'}(t)|^2$. It shall be assumed that the interaction is turned off when $t \rightarrow -\infty$. The interaction can be turned off at large negative times by introducing a multiplicative factor of $\exp[\, +\, \eta\, t\,]$ in the interaction, where η is an infinitesimally small positive constant. To first-order in the diamagnetic interaction, one finds

$$
C^{(1)}_{n'}(t) \;=\; \left(\frac{-\,i}{\hbar} \right) \delta_{n',n} \left(\frac{e^2}{2\, m\, c^2} \right) \left(\frac{2\,\pi\,\hbar\, c^2}{\sqrt{\omega_k\,\omega_{k'}}\, V} \right) 2\, \hat{\epsilon}_\alpha(\underline{k}) \,.\, \hat{\epsilon}_{\alpha'}(\underline{k}')
$$

$$
\int_{-\infty}^{t} dt' \; \exp\left[\frac{i}{\hbar}\, (\, \hbar\, \omega'\, +\, E_{n'}\, -\, \hbar\, \omega\, -\, E_n\,)\, t' \right] \tag{997}
$$

[53]H. A. Kramers and W. Heisenberg, Z. Physik, **31**, 681 (1925).

where $\omega = c\,k$ and $\omega' = c\,k'$ and the long-wavelength approximation has been used. The small quantity η has been absorbed as a small imaginary part to the initial state energy

$$E_n \rightarrow E_n + i\,\eta\,\hbar \qquad (998)$$

The paramagnetic interaction is of order of e and the diamagnetic interaction is of order e^2. Thus, to second-order in e, one must include the diamagnetic interaction and the paramagnetic interaction to second-order. There are two
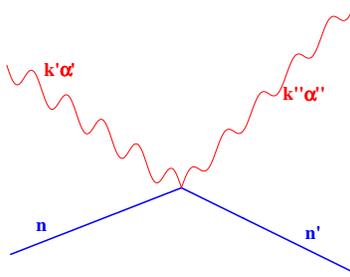
Figure 30: Photon scattering processes due to the diamagnetic interaction to first-order.

terms which are second-order in the paramagnetic interactions that represent:

(a) absorption of the photon $(\underline{k}, \alpha)$ followed by the emission of a photon $(\underline{k}', \alpha')$.

(b) emission of a photon $(\underline{k}', \alpha')$ followed by the absorption of the photon $(\underline{k}, \alpha)$.
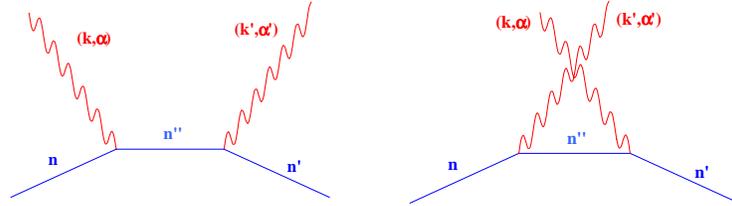
Figure 31: Photon scattering processes due to the paramagnetic interaction to second-order.

The second-order contribution to the transition amplitude is given by

$$C_{n'}^{(2)}(t) = \left(\frac{-i}{\hbar} \right)^2 \left(\frac{e}{m\,c} \right)^2 \left(\frac{2\,\pi\,\hbar\,c^2}{\sqrt{\omega_k\,\omega_{k'}}\,V} \right) \int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt''$$

$$\left[\sum_{n''} \exp[\,\frac{i}{\hbar}\,(\hbar\omega' + E_{n'} - E_{n''})\,t'\,]\,\exp[\,-\,\frac{i}{\hbar}\,(E_n - E_{n''} + \hbar\omega)\,t''\,] \right.$$

169

$$\times \; < \; n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid n''l''m'' \; > \; < \; n''l''m'' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \underline{\hat{p}} \mid nlm \; >$$

$$+ \; \sum_{n''} \; \exp[\; \frac{i}{\hbar} \; (E_{n'} - E_{n''} - \hbar\omega) \; t' \;] \; \exp[\; -\frac{i}{\hbar} \; (E_n - E_{n''} - \hbar\omega') \; t'' \;]$$

$$\times \; < \; n'l'm' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \underline{\hat{p}} \mid n''l''m'' \; > \; < \; n''l''m'' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid nlm \; > \Bigg]$$

$$(999)$$

The earliest time integration can be evaluated leading to

$$C_{n'}^{(2)}(t) \;\; = \;\; \left(\frac{-i}{\hbar} \right) \left(\frac{e}{m \; c} \right)^2 \left(\frac{2 \; \pi \; \hbar \; c^2}{\sqrt{\omega_k \; \omega_{k'}} \; V} \right) \int_{-\infty}^{t} dt' \; \sum_{n''} \; \exp[\; \frac{i}{\hbar} \; (\hbar\omega' + \hbar\omega + E_{n'} - E_n) \; t' \;]$$

$$\Bigg[\frac{< \; n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid n''l''m'' \; > \; < \; n''l''m'' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \underline{\hat{p}} \mid nlm \; >}{(\; E_n \; - \; E_{n''} \; + \; \hbar \; \omega \;)}$$

$$+ \; \frac{< \; n'l'm' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \underline{\hat{p}} \mid n''l''m'' \; > \; < \; n''l''m'' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid nlm \; >}{(\; E_n \; - \; E_{n''} \; - \; \hbar \; \omega' \;)} \Bigg] \qquad (1000)$$

as long as the denominators are non-vanishing.

The coefficients $C_{n'}^{(1)}(t)$ and $C_{n'}^{(2)}(t)$ have the same type of time-dependence. The remaining integration over time yields

$$\left(\frac{-i}{\hbar} \right) \int_{-\infty}^{t} dt' \; \exp\left[\; \frac{i}{\hbar} \; (\hbar\omega' + E_{n'} - E_n - \hbar\omega - i\hbar\eta) \; t' \; \right]$$

$$= \; - \; \frac{\exp[\; \frac{i}{\hbar} \; (\hbar\omega' + E_{n'} - E_n - \hbar\omega - i\hbar\eta) \; t \;]}{(\hbar\omega' + E_{n'} - E_n - \hbar\omega - i\hbar\eta)} \qquad (1001)$$

The transition rate is given by

$$\frac{1}{\tau} \; = \; \frac{\partial}{\partial t} \left(\; \left| \; C_{n'}^{(1)}(t) \; + \; C_{n'}^{(2)}(t) \; \right|^2 \right) \qquad (1002)$$

but the time-dependence of the squared modulus is contained in the common factor

$$\left| \frac{\exp[\; \frac{i}{\hbar} \; (\hbar\omega' + E_{n'} - E_n - \hbar\omega - i\hbar\eta) \; t \;]}{(\hbar\omega' + E_{n'} - E_n - \hbar\omega - i\hbar\eta)} \right|^2 \; = \; \frac{\exp\left[\; 2 \; \eta \; t \; \right]}{(\hbar\omega' + E_{n'} - E_n - \hbar\omega)^2 \; + \; \hbar^2\eta^2}$$

$$(1003)$$

Therefore, one finds the transition rate is given by the expression

$$\frac{1}{\tau} \; = \; \frac{2 \; \eta \; \exp\left[\; 2 \; \eta \; t \; \right]}{(\hbar\omega' + E_{n'} - E_n - \hbar\omega)^2 \; + \; \hbar^2\eta^2} \left(\frac{e^2}{m \; c^2} \right)^2 \left(\frac{2 \; \pi \; \hbar \; c^2}{\sqrt{\omega_k \; \omega_{k'}} \; V} \right)^2 M^2 \quad (1004)$$

where the matrix elements are given by

$$
\begin{aligned}
M \;=\; \Big[& \;\hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') \;<\; n'l'm' \mid nlm \;> \\
& + \frac{1}{m} \sum_{n''} \frac{<\; n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid n''l''m'' \;><\; n''l''m'' \mid \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{\hat{p}} \mid nlm \;>}{(\; E_n \;-\; E_{n''} \;+\; \hbar\,\omega \;)} \\
& + \frac{1}{m} \sum_{n''} \frac{<\; n'l'm' \mid \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{\hat{p}} \mid n''l''m'' \;><\; n''l''m'' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid nlm \;>}{(\; E_n \;-\; E_{n''} \;-\; \hbar\,\omega' \;)} \Big]
\end{aligned}
$$
(1005)

On taking the limit $\eta \;\rightarrow\; 0$, the first factor in the decay rate reduces to an energy conserving delta function. Therefore, one obtains the Fermi-Golden rule expression

$$
\frac{1}{\tau} \;=\; \frac{2\,\pi}{\hbar} \left(\frac{e^2}{m\,c^2} \right)^2 \left(\frac{2\,\pi\,\hbar\,c^2}{\sqrt{\omega_k\,\omega_{k'}}\,V} \right)^2 M^2\, \delta(\hbar\omega_{k'} + E_{n'} - E_n - \hbar\omega_k)
\tag{1006}
$$

The magnitudes of the final state photon quantum numbers $(\underline{k}')$ must be integrated over, since these are not measured. This integration imparts a physical meaning to the expression for the rate which contains the Dirac delta function. We shall assume that the direction of the scattered photon is to be measured and that the photon is absorbed by a detector which subtends a solid angle $d\Omega'$ to the atom. Therefore, the scattering rate is given by

$$
\frac{1}{\tau_{d\Omega'}} \;=\; \frac{2\,\pi}{\hbar} \left(\frac{e^2}{m\,c^2} \right)^2 \frac{V}{(\,2\,\pi\,)^3}\, d\Omega' \int_0^\infty dk'\, k'^2 \left(\frac{2\,\pi\,\hbar\,c^2}{\sqrt{\omega_k\,\omega_{k'}}\,V} \right)^2 \mid M \mid^2 \delta(\hbar\omega_{k'} + E_{n'} - E_n - \hbar\omega_k)
\tag{1007}
$$

Since $\hbar\,\omega_{k'} \;=\; \hbar\,c\,k'$, the integration over the delta function can be performed, yielding

$$
\frac{1}{\tau_{d\Omega'}} \;=\; \frac{2\,\pi}{\hbar} \left(\frac{e^2}{m\,c^2} \right)^2 \frac{V\,d\Omega'}{(\,2\,\pi\,)^3} \frac{\omega'^2}{\hbar\,c^3} \left(\frac{2\,\pi\,\hbar\,c^2}{\sqrt{\omega\,\omega'}\,V} \right)^2 \mid M \mid^2
\tag{1008}
$$

The scattering cross-section is defined as the transition rate divided by the photon flux. The photon flux is found by noting that it has been assumed that there is one photon per volume V so the photon density is $\frac{1}{V}$ and the speed of light is c. Hence, the photon flux is given by $\frac{c}{V}$. Therefore, the cross-section is determined by the Kramers-Heisenberg formula

$$
\left(\frac{d\sigma}{d\Omega'} \right) \;=\; \left(\frac{e^2}{m\,c^2} \right)^2 \left(\frac{\omega'}{\omega} \right) \mid M \mid^2
\tag{1009}
$$

The magnitude of the scattering rate is determined by the quantity r_e which has the dimensions of length

$$
r_e \;=\; \left(\frac{e^2}{m\,c^2} \right)
\tag{1010}
$$

171

This quantity is often called the classical radius of the electron. The quantity r_e can be expressed as

$$r_e = \left(\frac{e^2}{m\,c^2} \right) = \left(\frac{e^2}{\hbar\,c} \right) \left(\frac{\hbar}{m\,c} \right) \approx 2.82 \times 10^{-15} \text{ m} \qquad (1011)$$

11.2.1 Rayleigh Scattering

Rayleigh scattering corresponds to the limit in which the light is elastically scattered. Hence, one has

$$\omega = \omega' \qquad (1012)$$

In the case of elastic scattering, all the terms in the Kramers-Heisenberg formula are equally important. That all terms have a similar magnitude can be seen by re-writing the first term $\hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\epsilon}_{\alpha'}(\underline{k}')$ in a way which is similar to the second. The scalar product of the polarization vectors can be expressed as

$$\hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') = \sum_{i,j} \hat{\epsilon}_\alpha(\underline{k})_i \, \delta_{i,j} \, \hat{\epsilon}_{\alpha'}(\underline{k}')_j \qquad (1013)$$

but one can re-write the Kronecker delta function in terms of the commutation relation

$$[\, x_i \, , \, \hat{p}_j \,] = i\,\hbar\,\delta_{i,j} \qquad (1014)$$

Thus, one can express the scalar product as a commutator

$$\hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') = \frac{1}{i\,\hbar} \sum_{i,j} \hat{\epsilon}_\alpha(\underline{k})_i \, [\, x_i \, , \, \hat{p}_j \,] \, \hat{\epsilon}_{\alpha'}(\underline{k}')_j$$

$$= \frac{1}{i\,\hbar} \, [\, \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{r} \, , \, \hat{\underline{p}} \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') \,] \qquad (1015)$$

Since, in the dipole approximation, the diamagnetic contribution to the matrix elements M is proportional to the overlap integral

$$< n'l'm' \mid nlm > \qquad (1016)$$

the initial and final states must be identical if this is non-zero. Hence, the result is equivalent to the expectation value in the state $\mid nlm >$. On replacing the matrix elements by the expectation value and then insert a complete set of electronic states, one finds

$$< n'l'm' \mid nlm > \hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\epsilon}_{\alpha'}(\underline{k}')$$

$$= \frac{1}{i\,\hbar} \sum_{n''l''m''} \Bigg[\, < n'l'm' \mid \underline{r} \cdot \hat{\epsilon}_\alpha(\underline{k}) \mid n''l''m'' > < n''l''m'' \mid \underline{\epsilon}_{\alpha'}(\underline{k}') \cdot \hat{\underline{p}} \mid nlm >$$

$$- < n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \hat{\underline{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{r} \cdot \hat{\epsilon}_\alpha(\underline{k}) \mid nlm > \Bigg] \qquad (1017)$$

The matrix elements of $\underline{r}$ can be expressed in terms of the matrix elements of $\hat{\underline{p}}$ via

$$
\begin{aligned}
< n'l'm' \mid \hat{\underline{p}} \mid n''l''m'' > \; &= \; \frac{1}{2\,i\,\hbar} \; < n'l'm' \mid [\, \underline{r}\, ,\, \hat{p}^2\,] \mid n''l''m'' > \\
&= \; \frac{m}{i\,\hbar} \; < n'l'm' \mid [\, \underline{r}\, ,\, \hat{H}_0\,] \mid n''l''m'' > \\
&= \; \frac{m}{i\,\hbar} \; (\, E_{n''l''m''} \, - \, E_{n'l'm'}\,) \; < n'l'm' \mid \underline{r} \mid n''l''m'' >
\end{aligned}
$$

$$(1018)$$

Therefore, one finds

$$
< n'l'm' \mid \underline{r} \mid n''l''m'' > \; = \; \frac{i}{m\,\omega_{n'',n'}} \; < n'l'm' \mid \hat{\underline{p}} \mid n''l''m'' > \tag{1019}
$$

where

$$
E_{n''l''m''} \, - \, E_{n'l'm'} \; = \; \hbar\,\omega_{n''n'} \tag{1020}
$$

Thus, the elastic scattering term in the Kramers-Heisenberg formula is given by

$$
\begin{aligned}
&\delta_{nlm,n'l'm'} \; \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \\
&= \; \frac{1}{m} \sum_{n''l''m''} \frac{< n'l'm' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_\alpha(\underline{k}) \mid n''l''m'' > < n''l''m'' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \mid nlm >}{\hbar\,\omega_{n''n'}} \\
&- \; \frac{1}{m} \sum_{n''l''m''} \frac{< n'l'm' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \mid n''l''m'' > < n''l''m'' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_\alpha(\underline{k}) \mid nlm >}{\hbar\,\omega_{nn''}}
\end{aligned}
$$

$$(1021)$$

but since for elastic scattering $E_{nlm} \; = \; E_{n'l'm'}$, one has

$$
\begin{aligned}
&\delta_{nlm,n'l'm'} \; \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \\
&= \; \frac{1}{m} \sum_{n''l''m''} \frac{< n'l'm' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_\alpha(\underline{k}) \mid n''l''m'' > < n''l''m'' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \mid nlm >}{E_{n''} \, - \, E_n} \\
&+ \; \frac{1}{m} \sum_{n''l''m''} \frac{< n'l'm' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \mid n''l''m'' > < n''l''m'' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_\alpha(\underline{k}) \mid nlm >}{E_{n''} \, - \, E_n}
\end{aligned}
$$

$$(1022)$$

On substituting this back into the expression for the matrix elements M, one obtains

$$
\begin{aligned}
M \; = \; &\Bigg[\frac{1}{m} \sum_{n''l''m''} \frac{< n'l'm' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_\alpha(\underline{k}) \mid n''l''m'' > < n''l''m'' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \mid nlm >}{E_{n''} \, - \, E_n} \\
&+ \; \frac{1}{m} \sum_{n''l''m''} \frac{< n'l'm' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \mid n''l''m'' > < n''l''m'' \mid \hat{\underline{p}} \, . \, \hat{\epsilon}_\alpha(\underline{k}) \mid nlm >}{E_{n''} \, - \, E_n}
\end{aligned}
$$

173

$$
+ \frac{1}{m} \sum_{n'l''m''} \frac{< n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{\hat{p}} \cdot \hat{\epsilon}_{\alpha}(\underline{k}) \mid nlm >}{(E_n - E_{n''} + \hbar \omega)}
$$

$$
+ \frac{1}{m} \sum_{n''l''m''} \frac{< n'l'm' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \underline{\hat{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{\hat{p}} \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') \mid nlm >}{(E_n - E_{n''} - \hbar \omega)} \Bigg]
$$

$$(1023)$$

which simplifies to

$$
M \;\; = \;\; \frac{\omega}{m \, \hbar} \sum_{n''l''m''} \Bigg[\frac{< n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{\hat{p}} \cdot \hat{\epsilon}_{\alpha}(\underline{k}) \mid nlm >}{\omega_{n''n} \, (\, \omega_{nn''} + \omega \,)}
$$

$$
- \; \frac{< n'l'm' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \underline{\hat{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{\hat{p}} \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') \mid nlm >}{\omega_{n''n} \, (\, \omega_{nn''} - \omega \,)} \Bigg]
$$

$$(1024)$$

In the limit of small photon frequencies compared with the electronic energies, one can expand the denominators of the matrix element as

$$
\frac{1}{\omega_{nn''} \, (\, \omega_{nn''} \pm \omega \,)} \;\; = \;\; \frac{1}{\omega_{nn''}^2} \; \mp \; \frac{\omega}{\omega_{nn''}^3} \; + \; \ldots \qquad (1025)
$$

When this low-frequency expansion is substituted into the matrix elements, the leading term vanishes. This can be seen since the leading term is proportional to

$$
\sum_{n''} \frac{1}{\omega_{nn''}^2} \Bigg[\; < n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{\hat{p}} \cdot \hat{\epsilon}_{\alpha}(\underline{k}) \mid nlm >
$$

$$
- \; < n'l'm' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \underline{\hat{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{\hat{p}} \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') \mid nlm > \Bigg]
$$

$$(1026)$$

which can be expressed as

$$
m^2 \sum_{n''} \Bigg[\; < n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{r} \mid n''l''m'' > < n''l''m'' \mid \underline{r} \cdot \hat{\epsilon}_{\alpha}(\underline{k}) \mid nlm >
$$

$$
- \; < n'l'm' \mid \hat{\epsilon}_{\alpha}(\underline{k}) \cdot \underline{r} \mid n''l''m'' > < n''l''m'' \mid \underline{r} \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') \mid nlm > \Bigg]
$$

$$(1027)$$

or, on using the completeness relation, one finds the expectation value of the commutator is given by

$$
m^2 \; < n'l'm' \mid [\, \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{r} \, , \; \underline{r} \cdot \hat{\epsilon}_{\alpha}(\underline{k}) \,] \mid nlm > = 0 \qquad (1028)
$$

Thus, the leading term of the low-frequency expansion vanishes. Therefore, the scattering rate is expressed as

$$
\left(\frac{d\sigma}{d\Omega'} \right) \;=\; \left(\frac{r_e}{m\,\hbar} \right)^2 \omega^4 \left| \sum_{n''l''m''} \left(\frac{1}{\omega_{nn''}} \right)^3 \right.
$$
$$
\times \left[\; < n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{\hat{p}} \cdot \hat{\epsilon}_\alpha(\underline{k}) \mid nlm > \right.
$$
$$
\left. + \; < n'l'm' \mid \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{\hat{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{\hat{p}} \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') \mid nlm > \left. \right] \right|^2
$$

(1029)

Finally, the scattering rate can be expressed in terms of the dipole matrix elements as

$$
\left(\frac{d\sigma}{d\Omega'} \right) \;=\; \left(\frac{r_e\,m}{\hbar} \right)^2 \omega^4 \left| \sum_{n''l''m''} \left(\frac{1}{\omega_{nn''}} \right)^3 \right.
$$
$$
\times \left[\; < n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{r} \mid n''l''m'' > < n''l''m'' \mid \underline{r} \cdot \hat{\epsilon}_\alpha(\underline{k}) \mid nlm > \right.
$$
$$
\left. + \; < n'l'm' \mid \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{r} \mid n''l''m'' > < n''l''m'' \mid \underline{r} \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') \mid nlm > \left. \right] \right|^2
$$

(1030)

Hence, at long-wavelengths, the scattering cross-section varies as ω^4 as expected from Rayleigh's law. Since the typical electronic frequency $\omega_{nn''}$ is in the ultra-violet spectrum, then

$$
\omega_{nn'} \;\gg\; \omega
$$

(1031)

for all frequencies in the visible optical spectrum. This leads to the phenomena of blue skies in the day and red sunsets at dusk.

11.2.2 Thomson Scattering

Thomson scattering occurs for photons with sufficiently high energies

$$
\omega \;\gg\; \omega_{nn''}
$$

(1032)

so that the photon energy is greater than the atomic binding-energy. In this case, the second and third terms in the Kramers-Heisenberg formula can be neglected. This is because

$$
\omega \sim \omega' \;\gg\; \frac{1}{m} \; < n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \cdot \underline{\hat{p}} \mid n''l''m'' > < n''l''m'' \mid \underline{\hat{p}} \cdot \hat{\epsilon}_\alpha(\underline{k}) \mid nlm >
$$

(1033)

Therefore, the scattering predominantly occurs elastically and the scattering cross-section is given by

$$\left(\frac{d\sigma}{d\Omega'} \right) = r_e^2 \left| \hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\epsilon}_{\alpha'}(\underline{k}') \right|^2 \qquad (1034)$$

which is independent of ω. The above result is dependent on the scattering angle via the polarization vectors.

In the investigation of the angular dependence of Thomson scattering, it is convenient to introduce a coordinate system which is defined by the polarization vectors and direction of propagation of the incident photon and its polarization $\hat{\epsilon}_1(\underline{k})$. The coordinate system is composed of the three orthogonal unit vectors $(\hat{\epsilon}_1(\underline{k}), \hat{\epsilon}_2(\underline{k}), \hat{k})$. Thus the direction of the polarization vector $\hat{\epsilon}_1(\underline{k})$ defines the x-direction. In this coordinate system, the scattered photon $(\underline{k}', \alpha')$ is in the direction $\underline{k}'$ with polar coordinates $(\theta_{k'}, \varphi_{k'})$. The polarization of the final photons $\hat{\epsilon}_\alpha(\underline{k}')$ must be transverse to $\underline{k}'$. Two polarization vectors are defined according to

$$\hat{\epsilon}_1(\underline{k}') = (\cos\theta_{k'} \cos\varphi_{k'}, \cos\theta_{k'} \sin\varphi_{k'}, -\sin\theta_{k'}) \qquad (1035)$$

which lies in the plane of $\underline{k}$ and $\underline{k}'$ and

$$\hat{\epsilon}_2(\underline{k}') = (-\sin\varphi_{k'}, \cos\varphi_{k'}, 0) \qquad (1036)$$

which lies in the plane perpendicular to $\underline{k}$. In terms of the chosen polarization
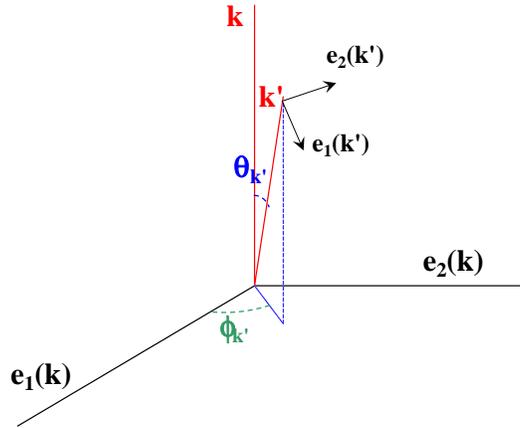
Figure 32: The coordinate system and polarization vectors used to describe Thomson scattering.

vectors, the scattering cross-section for incident radiation that is polarized along

176

the x-direction takes on the form

$$\left(\frac{d\sigma}{d\Omega'} \right)_{x-\text{pol}} = r_e^2 \left\{ \begin{array}{ll} \cos^2 \theta_{k'} \cos^2 \varphi_{k'} & \text{for } \alpha' = 1 \\ \sin^2 \varphi_{k'} & \text{for } \alpha' = 2 \end{array} \right. \tag{1037}$$

if the polarizations of the final photon are measured.

If the incident beam has its polarization along the x-direction, and the detector is not sensitive to the polarization, then the final polarization must be summed over. In this case of a polarized beam and a polarization insensitive detector, the cross-section is given by

$$\left(\frac{d\sigma}{d\Omega'} \right)_{x-\text{pol}} = r_e^2 \left(\cos^2 \theta_{k'} \cos^2 \varphi_{k'} + \sin^2 \varphi_{k'} \right) \tag{1038}$$

where the polarizations of the final state photon have been summed over.

If the incident beam of photons is unpolarized, then φ is undefined since the azimuthal direction of the scattered photon is defined with respect to the assumed polarization $\hat{\epsilon}_1(\underline{k})$. In the case of an unpolarized incident beam the expression should be integrated over φ and divided by 2π. The scattering rate is given by

$$\left(\frac{d\sigma}{d\Omega'} \right)_{\text{unpol}} = \frac{r_e^2}{2} \left\{ \begin{array}{ll} \cos^2 \theta_{k'} & \text{for } \alpha' = 1 \\ 1 & \text{for } \alpha' = 2 \end{array} \right. \tag{1039}$$

if the polarizations of the final state photons are measured. This result is identical to that obtained by assuming that the initial beam is composed of one half of the number photons polarized along the x-direction and the other half of the number of photons polarized along the y-direction. That is

$$\left(\frac{d\sigma}{d\Omega'} \right)_{\text{unpol}} = r_e^2 \left\{ \begin{array}{ll} \frac{1}{2} \cos^2 \theta_{k'} \left(\cos^2 \varphi_{k'} + \sin^2 \varphi_{k'} \right) & \text{for } \alpha' = 1 \\ \frac{1}{2} \left(\sin^2 \varphi_{k'} + \cos^2 \varphi_{k'} \right) & \text{for } \alpha' = 2 \end{array} \right. \tag{1040}$$

The cross-section for unpolarized photons with a polarization insensitive detector is given by

$$\left(\frac{d\sigma}{d\Omega'} \right)_{\text{unpol}} = \frac{r_e^2}{2} \left(1 + \cos^2 \theta_{k'} \right) \tag{1041}$$

where the final polarizations have been summed over.

The total cross-section σ is obtained by integrating over all directions. The total Thomson scattering cross-section is independent of whether the initial beam was polarized or unpolarized. The final result is

$$\sigma = \frac{8\pi}{3} r_e^2 \tag{1042}$$

which has a magnitude of $6.65 \times 10^{-29} \ m^2$. More massive charged particles, such as protons, can also produce Thomson scattering but the cross-sections for

these processes are smaller by factors of $(\frac{m}{M})^2$. The derivation of the Thomson scattering cross-section breaks down for photons which have energies of the order of the electron's rest energy

$$\hbar \, \omega \, \sim \, m_e \, c^2 \tag{1043}$$

For photons with these high-energies, one must describe the scattering process relativistically. In this energy region, Compton scattering dominates.

Classical Interpretation

The classical counter-parts of Rayleigh and Thomson scattering can be described by a two-step process. In the first step, the incident classical electromagnetic field causes an electron to undergo forced oscillations. In the second step, the oscillating electrons emit electromagnetic radiation.

In the first process, an electron bound harmonically to the atom which responds to an electromagnetic field $\underline{E}_0 \exp[\, i \, \omega \, t \,]$ can be described by the equation of motion

$$\underline{\ddot{r}} \, + \, \omega_0^2 \, \underline{r} \, = \, \frac{q}{m} \, \underline{E}_0 \, \Re e \, \exp\left[\, i \, \omega \, t \,\right] \tag{1044}$$

where ω_0 is the frequency of the electron's natural motion. In the steady state, one finds

$$\underline{r} \, = \, \frac{\frac{q}{m} \, \underline{E}_0}{\omega_0^2 \, - \, \omega^2} \, \Re e \, \exp\left[\, i \, \omega \, t \,\right] \tag{1045}$$

The acceleration of the charged particle can be described by

$$\underline{\ddot{r}} \, = \, - \, \frac{\frac{q}{m} \, \omega^2 \, \underline{E}_0}{\omega_0^2 \, - \, \omega^2} \, \Re e \, \exp\left[\, i \, \omega \, t \,\right] \tag{1046}$$

The accelerating charged particle radiates electromagnetic energy. The emitted power is given by the Larmor formula

$$\begin{aligned} P(\omega) \, &= \, \frac{2 \, q^2 \, r^2 \, \omega^4}{3 \, c^3} \\ &= \, \frac{2 \, q^4 \, \underline{E}_0^2}{3 \, m^2 \, c^3} \, \frac{\omega^4}{(\, \omega_0^2 \, - \, \omega^2 \,)^2} \end{aligned} \tag{1047}$$

while the incident energy flux is given by

$$\frac{c}{4 \, \pi} \, \underline{E}_0^2 \tag{1048}$$

Hence, the scattering cross-section is described by

$$\sigma \, = \, \frac{8 \, \pi}{3} \, r_e^2 \, \frac{\omega^4}{(\, \omega_0^2 \, - \, \omega^2 \,)^2} \tag{1049}$$

178

This formula has the correct frequency dependence in the limit $\omega \ll \omega_0$ in which case the classical cross-section varies as ω^4, as expected for Rayleigh scattering. On the other hand, in the limit $\omega \gg \omega_0$ the cross-section becomes frequency independent, as is expected for Thomson scattering.

11.2.3 Raman Scattering

For inelastic scattering, one has $\hbar \omega \neq \hbar \omega'$, therefore, the condition of conservation of energy requires that

$$E_{nlm} + \hbar \omega = E_{n'l'm'} + \hbar \omega' \tag{1050}$$

Since it is most probable that the initial electron is in the ground state, one has

$$E_{n'l'm'} > E_{nlm} \tag{1051}$$

which leads to the inequality

$$\hbar \omega > \hbar \omega' \tag{1052}$$

Hence, the final photon has less energy than the initial photon. That is, the
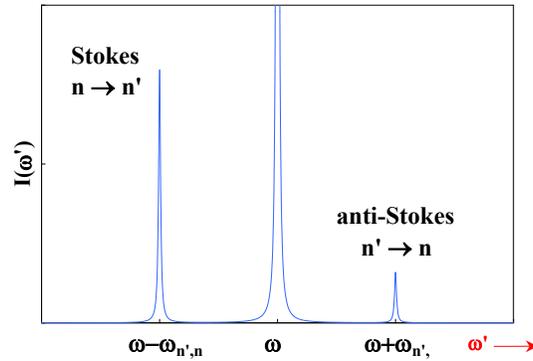
Figure 33: The schematic frequency dependence of the observed intensity expected in a Raman scattering experiment. The ratio of intensities of the Stokes and anti-Stokes lines provides a relative measure of the initial occupation of the low-energy state n and the higher-energy excited state n'.

electromagnetic field has lost energy and left the electron in an excited state. This inelastic process describes the Stoke's line. On the other hand, if the electron is initially in an excited state, then it is possible that the electron loses energy and makes a transition to the ground state. In this case,

$$E_{n'l'm'} < E_{nlm} \tag{1053}$$

179

so the final photon is more energetic

$$\hbar \, \omega \; < \; \hbar \, \omega' \tag{1054}$$

This process results in the anti-Stokes line.

11.2.4 Radiation Damping and Resonance Fluorescence

In the analysis of photon scattering, it has been assumed that the energy de-nominators ($E_n \; - \; E_{n''} \; + \; \hbar \, \omega$) do not vanish. If the energy denominator vanishes, the Kramers-Heisenberg formula becomes singular, however, the phys-ically observed scattering cross-section may become large but does not diverge. This is the phenomenon of resonance-fluorescence. Using the classical model, one can describe the scattering cross-section, if damping is introduced to rep-resent the lifetime of the electronic states. That is, the dynamics of the bound electron is modelled by a damped harmonic oscillator

$$\underline{\ddot{r}} \; + \; \gamma \, \underline{\dot{r}} \; + \; \omega_0^2 \, \underline{r} \; = \; \frac{q}{m} \, \underline{E}_0 \, \Re e \, \exp \left[\, i \, \omega \, t \, \right] \tag{1055}$$

which has the solution

$$\underline{r} \; = \; \frac{q}{m} \, \underline{E}_0 \, \Re e \, \frac{1}{\omega_0^2 \; + \; i \, \gamma \, \omega \; - \; \omega^2} \, \exp \left[\, i \, \omega \, t \, \right] \tag{1056}$$

since γ is related to the decay rate and is of the order of 10^8 sec^{-1}, it is usually negligible compared with the frequency of light which is estimated as $\omega \; \sim \; 10^{15}$ sec^{-1}. Following our previous arguments, one finds that the scattering cross-section is given by

$$\sigma(\omega) \; = \; \frac{8 \, \pi}{3} \, \left(\frac{q^2}{m \, c^2} \right)^2 \, \frac{\omega^4}{(\, \omega^2 \; - \; \omega_0^2 \,)^2 \; + \; \gamma^2 \, \omega^2} \tag{1057}$$

which no longer diverges when the resonance condition is satisfied, because of the damping of the electronic states.

The lifetime of a quantum mechanical state which at $t \, = \, 0$ is represented by $\mid \psi_n(0) >$ calculated to second-order in the interaction $\hat{H}_I$ is given by the Fermi-Golden rule expression. The rate can be expresses as the limit $\eta \to 0$ by

$$\frac{1}{\tau_n} \; = \; - \, \frac{2}{\hbar} \, \Im m \, \sum_{n'} \, \frac{\mid < \psi_{n'} \mid \hat{H}_I \mid \psi_n > \mid^2}{E_n \; - \; E_{n'} \; + \; i \, \eta} \tag{1058}$$

whereas the energy-shift found in second-order (Rayleigh-Schrödinger) pertur-bation theory is also given by the limit $\eta \to 0$ of

$$\Delta E_n \; = \; \Re e \, \sum_{n} \, \frac{\mid < \psi_{n'} \mid \hat{H}_I \mid \psi_n > \mid^2}{E_n \; - \; E_{n'} \; + \; i \, \eta} \tag{1059}$$

so

$$E_n = E_n^{(0)} + \Delta E_n \tag{1060}$$

Hence, due to the form of the expressions for the shift and the lifetime as the real and imaginary parts of a complex function, it is possible to consider an unstable state as having a complex energy[54] given by

$$E_n - i\,\Gamma_n \approx E_n^{(0)} + \Delta E_n - i\,\frac{\hbar}{2\,\tau_n} \tag{1061}$$

That is, the lifetime can be considered as giving the state an energy-width Γ_n. This is the natural width of the electronic state. The factor of two in the width can be understood by considering the time-dependence of the state $\mid \psi_n(t) >$ which is given by

$$\mid \psi_n(t) > = \exp\left[-\frac{i}{\hbar}\,(\,E_n - i\,\Gamma_n\,)\,t \right]\mid \psi_n(0) > \tag{1062}$$

Hence, the probability $P_n(t)$ that the state has not decayed at time t is given by

$$
\begin{aligned}
P_n(t) &= \mid < \psi_n(0) \mid \psi_n(t) > \mid^2 \\
&= \mid < \psi_n(0) \mid \exp\left[-\frac{i}{\hbar}\,(\,E_n - i\,\Gamma_n\,)\,t \right]\mid \psi_n(0) > \mid^2 \\
&= \exp\left[-\frac{2}{\hbar}\,\Gamma_n\,t \right] \tag{1063}
\end{aligned}
$$

due to the normalization of the initial state. This time-dependence of $P_n(t)$ is interpreted in terms of the exponential decay of the probability for finding the initial state

$$P_n(t) = P_n(0)\,\exp\left[-\frac{t}{\tau_n} \right] \tag{1064}$$

This leads to the identification of the relation between the energy-width and the lifetime

$$\Gamma_n = \frac{\hbar}{2\,\tau_n} \tag{1065}$$

Hence, the lifetime τ_n of an unstable or metastable state can be incorporated by introducing an imaginary part Γ_n to the energy.

Therefore, for the case of resonant scattering, one should replace the energies by complex numbers such that the real part represents the state's energy and the imaginary part describes half the state's decay rate. In the case of resonant scattering, the Kramers-Heisenberg formula is modified[55] to

$$\left(\frac{d\sigma}{d\Omega'} \right) = \left(\frac{e^2}{m\,c^2} \right)^2 \left(\frac{\omega'}{\omega} \right)\mid M \mid^2 \tag{1066}$$

---

[54]That is, the perturbation produces a complex shift of the energy-shift which related to the self-energy $\Sigma_n(E)$ which is to be discussed later

[55]P. A. M. Dirac, Proc. Roy. Soc. **A 114**, 710 (1927).

where the matrix elements are given by

$$
\begin{aligned}
M \;=\; & \Bigg[\; \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\epsilon}_{\alpha'}(\underline{k}') \; < \; n'l'm' \mid nlm \; > \\
& + \frac{1}{m} \sum_{n''l''m''} \frac{< \; n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \, . \, \hat{\underline{p}} \mid n''l''m'' \; > \; < \; n''l''m'' \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\underline{p}} \mid nlm \; >}{(\; E_n \; - \; E_{n''} \; - \; i \, \Gamma_{n''} \; + \; \hbar \, \omega \;)} \\
& + \frac{1}{m} \sum_{n''l''m''} \frac{< \; n'l'm' \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\underline{p}} \mid n''l''m'' \; > \; < \; n''l''m'' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \, . \, \hat{\underline{p}} \mid nlm \; >}{(\; E_n \; - \; E_{n''} \; - \; i \, \Gamma_{n''} \; - \; \hbar \, \omega' \;)} \; \Bigg]
\end{aligned}
$$

$$(1067)$$

Since close to resonance, the resonant denominator is given by $\Gamma \; \sim \; \frac{\hbar \, c}{a} \, (\frac{e^2}{\hbar \, c})^4$ whereas the numerator is of the order of $\frac{e^2}{a}$. Hence, on-resonance, the matrix elements can be of the order $(\frac{e^2}{\hbar \, c})^{-3}$ larger than the non-resonant matrix elements. Therefore, on resonance, the non-resonant terms may be neglected. In the following, it shall be assumed that the resonant state is non-degenerate

$$
\left(\frac{d\sigma}{d\Omega'} \right) \;=\; \left(\frac{e^2}{m^2 \, c^2} \right)^2 \left(\frac{\omega'}{\omega} \right) \frac{\mid \; < \; n'l'm' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \, . \, \hat{\underline{p}} \mid n''l''m'' \; > \; < \; n''l''m'' \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\underline{p}} \mid nlm \; > \; \mid^2}{(\; E_n \; - \; E_{n''} \; + \; \hbar \, \omega \;)^2 \; + \; \Gamma^2_{n''}}
$$

$$(1068)$$

This expression can be re-expressed in terms of the product of two factors

$$
\begin{aligned}
\left(\frac{d\sigma}{d\Omega'} \right) \;=\; & \left(\frac{e^2 \, 2 \, \pi \, \hbar}{m^2 \, \omega \, V} \right) \frac{\mid \; < \; nlm \mid \hat{\epsilon}_\alpha(\underline{k}) \, . \, \hat{\underline{p}} \mid n''l''m'' \; > \; \mid^2}{(\; E_n \; - \; E_{n''} \; + \; \hbar \, \omega \;)^2 \; + \; \Gamma^2_{n''}} \frac{V}{c} \\
& \times \frac{2 \, \pi}{\hbar} \left(\frac{e^2 \, 2 \, \pi \, \hbar}{m^2 \, \omega' \, V} \right) \mid \; < \; n''l''m'' \mid \hat{\epsilon}_{\alpha'}(\underline{k}') \, . \, \hat{\underline{p}} \mid n'l'm' \; > \; \mid^2 \frac{V \, \omega'^2}{(\; 2 \, \pi \;)^3 \, \hbar \, c^3}
\end{aligned}
$$

$$(1069)$$

which is the probability for absorption from the ground state to the resonant state $\mid n''l''m'' >$ (divided by the incident flux) times the probability for its decay via emission. On resonance, it appears that the process corresponds to two sequential processes, first absorption and secondly emission.

For energies slightly off-resonance, the resonant scattering is expected to interfere with the non-resonant scattering process. Likewise, if the resonant state is degenerate, the sum over the degeneracy must be performed before the matrix elements are squared leading to constructive interference.

The difference between a resonant process and two step process, is determined by the lifetime of the intermediate state $\mid n'l'm' >$ compared with the frequency width of the photon beam. The frequency width of the photon beam may be limited by the monochromator, or by the time-scale of the experiment if it involves a pulsed light source. If the lifetime of the intermediate state is sufficiently long compared with the time scale of experiment, it may be possible to observe the decay long after the incident light has been switched off. In

this case, the resonance can be considered to be composed of two independent processes[56]. Furthermore, it may be possible to perform further experiments on the surviving intermediate state. In the opposite case, where the lifetime of the intermediate state is shorter than the time-scale of the experiment, the intermediate state will have decayed before the experiment has terminated.

11.2.5   Natural Line-Widths

The interaction representation will be used to calculate the natural width for the absorption of light $(\underline{k}, \alpha)$, by introducing the lifetimes of the initial and final state. Strictly speaking, one should not take the exponential decay of a probability $P_n(t)$ of finding an electron in state ψ_n too literally. If one considers the approximate exponential decay as being rigorous, this implies that the Hamiltonian should be non-Hermitean which is strictly forbidden. One should think of the decaying wave function as a wave packet or linear superposition of exact energy eigenstates (with energies denoted by E). The Fourier transform of the time-dependent wave function should provide the energy-distribution $\rho_n(E)$ of the exact energy eigenstates in the wave packet $\mid \psi_n(t) >$

$$\rho_n(E) \;\; = \;\; \frac{1}{2 \, \pi \, \hbar} \, \int_{-\infty}^{\infty} \, dt \, \exp\left[\, + \, \frac{i}{\hbar} \, E \, t \, \right] \;\; < \, \psi_n(0) \mid \psi_n(t) \, > \tag{1070}$$

On assuming the approximate form of a decaying wave packet

$$< \, \psi_n(0) \mid \psi_n(t) \, > \; = \; \exp\left[\, - \, \frac{i}{\hbar} \, E_n \, t \, - \, \frac{\mid t \mid}{2 \, \tau_n} \, \right] \tag{1071}$$

where the decay includes transitions to all possible final states, one finds

$$\begin{aligned}
\rho_n(E) \;\; &= \;\; \frac{1}{2 \, \pi \, i} \, \left(\, \frac{1}{E \, - \, E_n \, - \, i \, \frac{\hbar}{2 \, \tau_n}} \, - \, \frac{1}{E \, - \, E_n \, + \, i \, \frac{\hbar}{2 \, \tau_n}} \, \right) \\[2mm]
&= \;\; \frac{1}{\pi} \, \frac{\frac{\hbar}{2 \, \tau_n}}{(\, E \, - \, E_n \,)^2 \, + \, (\frac{\hbar}{2 \, \tau_n})^2} \tag{1072}
\end{aligned}$$

This can only be an approximate form of the energy-distribution since the energy must be bounded from below. The existence of a lower-bound to energy distribution implies that the width of the electronic energy level has to be energy-dependent $\frac{\hbar}{2 \, \tau_n} \; = \; \Gamma_n$ as this must become zero below a threshold energy. However, it should be noted that the width of the energy-distribution will determine the approximate exponential decay. Since the perturbations introduce an energy-dependent width to the wave packet, causality requires that the energy-shift ΔE_n should also be energy-dependent. Hence, the effects of the perturbation (such as the energy-shift and lifetime) should be described in

---

[56]V. Weisskopf, Ann. der Physik, **9**, 23 (1931).

terms of a self-energy $\Sigma_n(E)$

$$
\begin{aligned}
\Sigma_n(E) \;&=\; \Re e\;\Sigma_n(E) \;+\; i\;\Im m\;\Sigma_n(E) \\
&\approx\; \Delta E_n \;-\; i\;\Gamma_n
\end{aligned}
\tag{1073}
$$

The energy-dependent self-energy appears most naturally if one uses Brillouin-Wigner perturbation theory to calculate the correction to an approximate energy E_n. From second-order Brillouin-Wigner perturbation theory, one finds that the energy-dependent self-energy, when evaluated just above the real E axis, is given by

$$
\Sigma_n(E + i\eta) \;=\; \sum_{n'} \frac{|\;<\;n'\;|\;\hat{H}_I\;|\;n\;>\;|^2}{E\;+\;i\,\eta\;-\;E_{n'}}
\tag{1074}
$$

This complex self-energy has a real and imaginary part. The imaginary part can be thought of as occurring via amplification of the infinitesimal imaginary term $i\,\eta$ in the denominator, and can be seen to be non-zero when the energy E of the component in the wave packet falls in the region when the spectral density of the approximate $E_{n'}$ is finite. Hence, since the $E_{n'}$ are bounded from below, then so is the energy-distribution $\rho_n(E)$ since

$$
\rho_n(E) \;=\; -\,\frac{1}{\pi}\,\frac{\Im m\;\Sigma_n(E + i\eta)}{(\;E\;-\;E_n\;-\;\Re e\;\Sigma_n(E)\;)^2\;+\;(\;\Im m\;\Sigma_n(E + i\eta)\;)^2}
\tag{1075}
$$

The real part of the self-energy must also be energy-dependent, since it is related to the imaginary part via the Kramer's-Kronig relations

$$
\begin{aligned}
\Re e\;\Sigma_n(E) \;&=\; -\,\frac{1}{\pi}\,\int_{-\infty}^{\infty} dz\,\frac{\Im m\;\Sigma_n(z + i\eta)}{E\;-\;z} \\
\Im m\;\Sigma_n(E + i\eta) \;&=\; +\,\frac{\mathcal{P}r}{\pi}\,\int_{-\infty}^{\infty} dz\,\frac{\Re e\;\Sigma_n(z)}{E\;-\;z}
\end{aligned}
\tag{1076}
$$

where the Principal Part of an integral with a simple pole is defined as

$$
\mathcal{P}r\,\int_{-\infty}^{\infty} dz\,\frac{f(z)}{z} \;=\; \lim_{\epsilon \to 0}\,\left(\;\int_{+\epsilon}^{\infty} dz\,\frac{f(z)}{z}\;+\;\int_{-\infty}^{-\epsilon} dz\,\frac{f(z)}{z}\;\right)
\tag{1077}
$$

Hence, the real part of the self-energy is also energy-dependent. The Kramers-Kronig relation is an expression of causality.

Since the electronic states in the expression for the Fermi-Golden rule decay rate

$$
\frac{1}{\tau_{nl \to n'l'}} \;=\; \frac{2\,\pi}{\hbar}\,|\;<\;n'l'm'\;|\;\hat{H}_I\;|\;nlm\;>\;|^2\,\delta(\;E_{n'l'}\;-\;E_{nl}\;-\;\hbar\,\omega\;)
\tag{1078}
$$

are to be interpreted as wave packets with a distribution of energies, the factor expressing conservation of energy should be expressed in terms of the energy

conservation for the components of the wave packets. Hence, the decay rate should be written as the convolution

$$
\begin{aligned}
\frac{1}{\tau_{nl \to n'l'}} &= \frac{2\pi}{\hbar} \mid < n'l'm' \mid \hat{H}_I \mid nlm > \mid^2 \int_{-\infty}^{\infty} dE' \, \rho_{n'l'}(E') \int_{-\infty}^{\infty} dE \, \rho_{nl}(E) \, \delta(\, E' - E - \hbar\omega\,) \\
&= \frac{2\pi}{\hbar} \mid < n'l'm' \mid \hat{H}_I \mid nlm > \mid^2 \int_{-\infty}^{\infty} dE \, \rho_{n'l'}(\, E + \hbar\omega\,) \, \rho_{nl}(E)
\end{aligned}
\tag{1079}
$$

We shall use the approximation for the energy distributions suggested by eqn(1072). In this case, the convolution is evaluated by contour integration as

$$
\frac{1}{\tau_{nl \to n'l'}} = \frac{2}{\hbar} \mid < n'l'm' \mid \hat{H}_I \mid nlm > \mid^2 \frac{\frac{\hbar}{2\,\tau_n} + \frac{\hbar}{2\,\tau_{n'l}}}{(\,\hbar\omega + E_{nl} - E_{n'l'}\,)^2 + (\frac{\hbar}{2\,\tau_{nl}} + \frac{\hbar}{2\,\tau_{n'l'}})^2}
\tag{1080}
$$

since only the terms with poles on the opposite sides of the real-axis yield non-zero contributions. From this, one can show that the optical absorption cross-section is given by

$$
\sigma_{absorb}(\omega) = \frac{4\pi}{3} \left(\frac{e^2}{\hbar c} \right) \sum_{n'l'm'} \mid < n'l'm' \mid r \mid nlm > \mid^2 \frac{\omega_{n'l',nl} \left(\frac{1}{2\,\tau_{n'l'}} + \frac{1}{2\,\tau_{nl}} \right)}{(\,\omega_{n'l',nl} - \omega\,)^2 + (\,\frac{1}{2\,\tau_{n'l'}} + \frac{1}{2\,\tau_{nl}}\,)^2}
\tag{1081}
$$

which was first derived by Weisskopf and Wigner[57]. Hence, the natural width is given by the average of the decay rates for the initial and final electronic states. This leads to the conclusion that even weak lines can be broad, if the final electronic state has a short lifetime.

11.3 Renormalization and Regularization

Quantum Electrodynamics treats the interactions between charged particles and the electromagnetic field, and often contains infinities. The zero-point energy of the electromagnetic field is one such infinity. In most cases, these infinities can be ignored since they are not measurable, since the infinities occur as modifications caused by the introduction of interactions between the charged particles of a hypothetical system with an electromagnetic field. That is, the infinities occur in the form of a renormalization of the quantities of the non-interacting theory. These infinite renormalizations do not lead to the rejection of the theory of Quantum Electrodynamics since the quantities of the non-interacting system are not measurable. To be sure, the infinities occur in relations between hypothetical quantities and physically measurable quantities, and so these infinities can be ignored since the hypothetical quantities are undefined. However, it is possible to use the theory to eliminate the unmeasurable quantities, thereby yielding relations between physically measurable quantities to other physically

---

[57]V. F. Weisskopf and E. Wigner, Z. Physik, **63**, 54 (1930).

measurable quantities. In Quantum Electrodynamics, the infinities cancel in equations which only contain physical measurable quantities. This fortunate circumstance makes the theory of Quantum Electrodynamics renormalizable.

First, it shall be shown how the infinite zero-point energy of the electromagnetic field can lead to a (finite) physically measurable force between its containing walls.

11.3.1 The Casimir Effect

The zero-point energy of the electromagnetic field can lead to measurable effects. In general relativity, the total energy including the zero-point energy of the electromagnetic radiation is the source for the gravitational field. The Casimir effect[58] shows that the zero-point energy of the electromagnetic radiation produces a force on the walls of the cavity. We shall consider a cubic volume $V = L^3$ which is enclosed by conducting walls that acts as a cavity for the electromagnetic radiation. This volume is divided into two by a metallic partition, which is located at a distance d from one side of the cavity. We shall
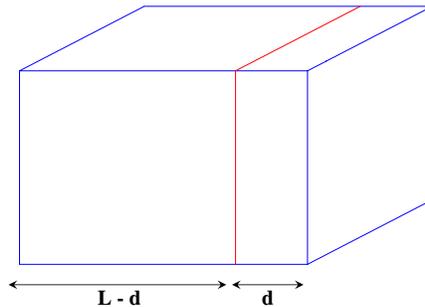
Figure 34: The geometry of the partitioned electromagnetic cavity used to consider the Casimir effect.

evaluate the total energy of this configuration and then deduce the form of the interaction between the partition and the walls of the cavity.

We shall consider the total energy due to the zero-point fluctuations in the container. Since the zero-point energy is divergent due to the presence of arbitrarily large frequencies, we shall introduce a convergence factor. The introduction of a convergence factor to remove infinities is the process of regularization. The introduction of the convergence factor can be motivated by the observation that, in matter, electromagnetic radiation becomes exponentially damped at

---

[58]H. B. G. Casimir, Physica **19**, 846 (1953).

large frequencies. Hence, one can write

$$E \; = \; \frac{1}{2} \sum_{\underline{k},\alpha} \hbar \, \omega_{\underline{k},\alpha} \, \exp \left[\, - \, \lambda \, \frac{\omega_{\underline{k},\alpha}}{c} \, \right] \tag{1082}$$

and then take the limit $\lambda \to 0$.

The presence of the conducting walls introduces boundary conditions such that the EM field is zero at every boundary. The boundary conditions restrict the allowed values of $\underline{k}$ so that the components satisfy

$$k_i \, L \; = \; \pi \, n_i \tag{1083}$$

for $i = x, y$ and n_i are positive integers. The boundary condition for the remaining two boundaries leads to the restriction

$$k_z \, d \; = \; \pi \, n_z \tag{1084}$$

The energy of the radiation in one part of the partition can be expressed as

$$E_d \; = \; \hbar \, c \, \frac{L^2}{\pi^2} \int_0^\infty dk_x \int_0^\infty dk_y \sum_{n_z=1}^\infty \sqrt{ k_x^2 \, + \, k_y^2 \, + \, \left(\frac{n_z \, \pi}{d} \right)^2 } \, \exp \left[- \lambda \sqrt{ k_x^2 \, + \, k_y^2 \, + \, \left(\frac{n_z \, \pi}{d} \right)^2 } \, \right] \tag{1085}$$

where the two polarizations have been summed over. The integration has cylindrical symmetry but only extends over the quadrant with positive k_x and k_y, therefore, it shall be re-written as

$$E_d \; = \; \hbar \, c \, \frac{2 \, L^2}{4 \, \pi} \int_0^\infty dk \, k \sum_{n_z=1}^\infty \sqrt{ k^2 \, + \, \left(\frac{n_z \, \pi}{d} \right)^2 } \, \exp \left[- \lambda \sqrt{ k^2 \, + \, \left(\frac{n_z \, \pi}{d} \right)^2 } \, \right] \tag{1086}$$

or, on changing variable to the dimensionless $\kappa \; = \; k^2 \, (\frac{d}{n_z \, \pi})^2$

$$E_d \; = \; \hbar \, c \, \frac{L^2}{4 \, \pi} \sum_{n_z=1}^\infty \left(\frac{n_z \, \pi}{d} \right)^3 \int_0^\infty d\kappa \, \sqrt{\kappa \, + \, 1} \, \exp \left[\, - \, \frac{n_z \, \pi \, \lambda}{d} \, \sqrt{\kappa \, + \, 1} \, \right] \tag{1087}$$

The factor of n_z^3 can be expressed as a third-order derivative of the exponential factor w.r.t. λ

$$E_d \; = \; - \, \hbar \, c \, \frac{L^2}{4 \, \pi} \sum_{n_z=1}^\infty \int_0^\infty d\kappa \, \frac{1}{\kappa \, + \, 1} \, \frac{\partial^3}{\partial \lambda^3} \left(\, \exp \left[\, - \, \frac{n_z \, \pi \, \lambda}{d} \, \sqrt{\kappa \, + \, 1} \, \right] \right) \tag{1088}$$

The summation over n_z can be performed, leading to

$$\begin{aligned} E_d \; &= \; - \, \hbar \, c \, \frac{L^2}{4 \, \pi} \int_0^\infty d\kappa \, \frac{1}{\kappa \, + \, 1} \, \frac{\partial^3}{\partial \lambda^3} \left(\frac{\exp[\, - \, \frac{\pi \, \lambda}{d} \, \sqrt{\kappa \, + \, 1} \,]}{1 \, - \, \exp[\, - \, \frac{\pi \, \lambda}{d} \, \sqrt{\kappa \, + \, 1} \,]} \right) \\ &= \; - \, \hbar \, c \, \frac{L^2}{4 \, \pi} \int_0^\infty d\kappa \, \frac{1}{\kappa \, + \, 1} \, \frac{\partial^3}{\partial \lambda^3} \left(\frac{1}{\exp[\, \frac{\pi \, \lambda}{d} \, \sqrt{\kappa \, + \, 1} \,] \, - \, 1} \right) \end{aligned} \tag{1089}$$

187

Let $t = \sqrt{\kappa + 1}$ so

$$E_d \;=\; -\,\hbar\,c\,\frac{2\,L^2}{4\,\pi}\,\int_1^\infty \frac{dt}{t}\,\frac{\partial^3}{\partial\lambda^3}\left(\frac{1}{\exp[\,\frac{\pi\,\lambda}{d}\,t\,]\;-\;1}\right) \tag{1090}$$

The factor of t^{-1} can be eliminated by performing one of the differentials with respect to λ.

$$E_d \;=\; \frac{\hbar\,c\,L^2}{2\,d}\,\int_1^\infty dt\,\frac{\partial^2}{\partial\lambda^2}\left(\frac{\exp[\,\frac{\pi\,\lambda}{d}\,t\,]}{(\,\exp[\,\frac{\pi\,\lambda}{d}\,t\,]\;-\;1\,)^2}\right) \tag{1091}$$

We shall set

$$s \;=\; \exp[\,\frac{\pi\,\lambda\,t}{d}\,]\;-\;1 \tag{1092}$$

therefore

$$E_d \;=\; \frac{\hbar\,c\,L^2}{2\,d}\,\frac{\partial^2}{\partial\lambda^2}\left(\frac{d}{\pi\,\lambda}\right)\int_{s_0}^\infty \frac{ds}{s^2} \tag{1093}$$

where the lower limit of integration depends on λ and is given by

$$s_0 \;=\; \exp[\,\frac{\pi\,\lambda}{d}\,]\;-\;1 \tag{1094}$$

The integration can be performed trivially, yielding

$$\begin{aligned}
E_d \;&=\; \frac{\hbar\,c\,L^2}{2\,d}\,\frac{\partial^2}{\partial\lambda^2}\left(\frac{d}{\pi\,\lambda\,s_0}\right)\\[6pt]
&=\; \frac{\hbar\,c\,L^2}{2\,d}\,\frac{\partial^2}{\partial\lambda^2}\left(\frac{\frac{d}{\pi\,\lambda}}{\exp[\,\frac{\pi\,\lambda}{d}\,]\;-\;1}\right)\\[6pt]
&=\; \frac{\hbar\,c\,L^2}{2\,d}\,\frac{\partial^2}{\partial\lambda^2}\left[\left(\frac{d}{\pi\,\lambda}\right)^2\left(\frac{\frac{\pi\,\lambda}{d}}{\exp[\,\frac{\pi\,\lambda}{d}\,]\;-\;1}\right)\right] \tag{1095}
\end{aligned}$$

The last factor in the above expression can be expanded as

$$\frac{x}{\exp[x]\;-\;1} \;=\; \sum_{n=0}^\infty B_n\,\frac{x^n}{n!} \tag{1096}$$

where B_n are the Bernoulli numbers, which are given by $B_0 \;=\; 1$, $B_1 \;=\; -\frac{1}{2}$, $B_2 \;=\; \frac{1}{6}$, $B_3 \;=\; 0$, $B_4 \;=\; -\frac{1}{30}$, etc. Therefore, the energy of the electromagnetic cavity at zero temperature, is finite for a finite value of the cut-off λ but diverges as λ^{-4} when $\lambda \to 0$. The zero point energy of the cavity can be expressed as

$$E_d \;=\; \frac{\hbar\,c\,L^2}{2}\left[\frac{\pi^2}{d^3}\sum_{n=0}\frac{B_n(n-2)(n-3)}{n!}\left(\frac{\pi\,\lambda}{d}\right)^{n-4}\right] \tag{1097}$$

188

where the $n = 0$ term diverges as λ^{-4} in the limit as $\lambda \to 0$ and is proportional to the volume of the cavity $d \, L^2$. The term with $n = 1$ also diverges, but diverges as λ^{-3} and has the form of a surface energy since it is proportional to L^2. The terms with $n = 2$ and $n = 3$ are identically equal to zero. The term with $n = 4$ remains finite in the limit $\lambda \to 0$ and all the higher-order terms vanish in this limit. Explicitly, one has

$$E_d \;=\; \frac{\hbar \, c \, L^2}{2 \, d} \left[\frac{6 \, B_0 \, d^2}{\pi^2 \, \lambda^4} \;+\; \frac{2 \, B_1 \, d}{\pi \, \lambda^3} \;+\; \frac{2 \, B_4 \, \pi^2}{4! \, d^2} \;+\; O(\lambda) \right] \qquad (1098)$$

The first term in the energy is proportional to $L^2 d$, which is the volume of the cavity and the second term is proportional to L^2 the surface area of the walls. The third term is independent of the cut-off and the higher order terms vanish in the limit $\lambda \to 0$.

The Casimir force is the force between two planes, which originates from the energy of the field[59]. This energy can be separated out into a volume part and parts due to the creation of the surfaces and an interaction energy between the surfaces. In order to eliminate both the volume dependence of the energy and the surface energies, we are considering two configurations of the partitions in the cavity. In one configuration the plane divides the volume into two unequal volumes $d \, L^2$ and $(L - d) \, L^2$, and the other configuration is a reference configuration where the cavity is partitioned into two equal volumes $\frac{L^3}{2}$. The difference of energies for these configurations is given by

$$\Delta E \;=\; E_d \;+\; E_{L-d} \;-\; 2 \, E_{\frac{L}{2}} \qquad (1099)$$

In the limit $L \to \infty$ this is expected to reduce to the energy of interaction between the planes separated by distance d. Since the volume and surface areas of the two partitions are identical, one finds that the difference in energy of the two configurations is finite and is given by

$$\lim_{L \gg d, \lambda \to 0} \Delta E \;\to\; -\frac{\pi^2}{720} \frac{\hbar \, c \, L^2}{d^3} \qquad (1100)$$

The d-dependence of the energy difference leads to an attractive force between the two plates separated by a distance d, which is the Casimir force

$$F \;=\; -\frac{\pi^2}{240} \, \hbar \, c \, \frac{L^2}{d^4} \qquad (1101)$$

The force is proportional to L^2 which is the area of the wall of the cavity. The predicted force was measured by Sparnaay[60]. A more recent experiment involving a similar force between a planar surface and a sphere has achieved greater

---

[59] Our considerations only include the part of Fock space that corresponds to having zero numbers of excited quanta. Hence, the Casimir force is due to the properties of the field, and is not due to the transmission of real particles (photons) between the planes.

[60] M. J. Sparnaay, Physica **24**, 751 (1958).
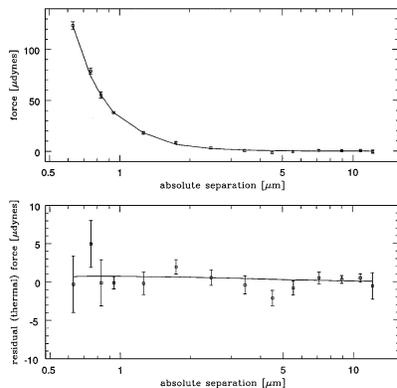
Figure 35: The separation-dependent force between two closely spaced metallic surfaces due to the modification of the zero-point energy. The lower panel shows the difference between the experimental results and the theoretical prediction for the Casimir Force. [After S. K. Lamoreaux, Phys. Rev. Lett. **78**, 5, (1997).]

accuracy[61].

To summarize, the physical quantity is the force or difference in energies when one wall is moved. When the change in energy is calculated, the difference between the two divergent energies is finite and independent of the choice of cut-off[62].

Cut-Off Independence

It is the boundary condition and not the cut-off that plays an important role in the Casimir effect. For simplicity, one can choose zero boundary conditions. The zero-point energy of a cylindrical electromagnetic cavity of radius R and length d can be expressed as the sum

$$
E_d \;=\; 2\,\frac{\hbar\,c}{2}\,\frac{\pi\,R^2}{(\,2\,\pi\,)}\,\sum_{n_z=1}^{\infty}\int dk_\rho\,k_\rho\,\sqrt{k_\rho^2\,+\,\left(\frac{\pi\,n_z}{d}\right)^2}\;F\!\left(\sqrt{k_\rho^2\,+\,\left(\frac{\pi\,n_z}{d}\right)^2}\right)
$$
(1102)

where $F(z)$ is an arbitrary cut-off function (which may depend on an arbitrary parameter λ which is ultimately going to be set to zero). The cut-off must not effect the low energy-modes so one can choose $F(0) = 1$ and all the derivatives of $F(z)$ to be zero for finite values of z. These assumptions are all in accord with

---

[61]S. K. Lamoreaux, Phys. Rev. Lett. **78**, 5 (1997).
[62]The independence of any cut-off procedure can be shown by evaluating the divergent sums by using the Euler-Maclaurin summation formula.

190
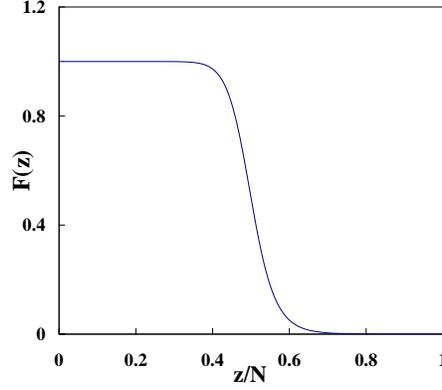
Figure 36: The schematic form of the cut-off function $F(z)$.

the ideal case of no cut-off function or $F(z) = 1$. The energy can be written as

$$E_d = \frac{\hbar \, c \, R^2}{2} \sum_{n_z=1}^{\infty} f(n_z)$$

(1103)

where

$$f(n_z) = \int_0^{\infty} dk_\rho \, k_\rho \, \sqrt{k_\rho^2 + \left(\frac{\pi \, n_z}{d}\right)^2} \, F\left(\sqrt{k_\rho^2 + \left(\frac{\pi \, n_z}{d}\right)^2}\right)$$

(1104)

The summation can be performed by changing it into an integral, however the corrections due to smoothing will be kept. This is accomplished by the Euler-Maclaurin formula. The integral between 0 and N of a function can be roughly expressed as a summation

$$\int_0^N dx \, f(x) \approx \frac{1}{2} \, f(0) + \sum_{n=1}^{N-1} f(n) + \frac{1}{2} \, f(N)$$

(1105)

by choosing to approximate the integral by the area under a histogram where the x variable is binned into intervals of width unity centered around $x = n$. The corrections at $n = 0$ and $n = N$ are needed to account for the fact that the range of integration excludes half the width of the rectangular blocks centered on $n = 0$ and $n = N$. The Euler-Maclaurin formulae is equivalent to finding a good smooth polynomial fit to the integrand, and then integrating the polynomial. It generates corrections which are given by the derivatives at the end points

$$\int_0^N dx \, f(x) = \frac{1}{2} \, f(0) + \sum_{n=1}^{N-1} f(n) + \frac{1}{2} \, f(N)$$
$$+ \frac{B_2}{2!} \left(f^{(1)}(0) - f^{(1)}(N) \right) + \frac{B_4}{4!} \left(f^{(3)}(0) - f^{(3)}(N) \right) + \ldots$$

(1106)

191

We shall assume that $f(n)$ and all its derivatives vanishes in the limit of large n, $\lim_{N \to \infty} f(N) \to 0$, due to the behavior of the cut-off function. The corrections in the Euler-Maclaurin summation formulae can be evaluated by noting that the first derivative of $f(n)$ with respect to n is given by

$$f^{(1)}(n) = \frac{\pi^2 \, n}{d^2} \int_0^\infty dk_\rho \, \frac{k_\rho}{\sqrt{k_\rho^2 + (\frac{\pi \, n}{d})^2}} \, F\left(\sqrt{k_\rho^2 + \left(\frac{\pi \, n}{d} \right)^2} \right) \quad (1107)$$

since the derivatives of $F(z)$ all vanish for finite z. The integration over the variable k_ρ is re-expressed in terms of an integration over the variable z, defined by

$$z = \sqrt{k_\rho^2 + \left(\frac{\pi \, n}{d} \right)^2} \quad (1108)$$

so

$$dz = dk_\rho \, \frac{k_\rho}{\sqrt{k_\rho^2 + \left(\frac{\pi \, n}{d} \right)^2}} \quad (1109)$$

The integration is evaluated through integration by parts

$$
\begin{aligned}
f^{(1)}(n) &= \frac{\pi^2 \, n}{d^2} \int_{\frac{\pi n}{d}}^\infty dz \, F(z) \\
&= \frac{\pi^2 \, n}{d^2} \int_{\frac{\pi n}{d}}^\infty dz \left(\frac{\partial z}{\partial z} \right) F(z) \\
&= \frac{\pi^2 \, n}{d^2} \left. z \, F(z) \right|_{\frac{\pi n}{d}}^\infty \\
&= -\frac{\pi^3 \, n^2}{d^3} \quad (1110)
\end{aligned}
$$

In deriving the above expression, the condition that the first-order derivative of $F(z)$ vanishes for finite z has been used. It immediately follows that

$$f^{(2)}(n) = -\frac{2 \, \pi^3 \, n}{d^3} \quad (1111)$$

and

$$f^{(3)}(n) = -\frac{2 \, \pi^3}{d^3} \quad (1112)$$

and all higher order derivatives vanish. Hence, one finds that at $z = 0$ all the m-th order derivatives $f^{(m)}(0)$ vanish, except for $m = 3$ which is given by

$$f^{(3)}(0) = -\frac{2 \, \pi^3}{d^3} \quad (1113)$$

192

Hence, on evaluating the energy of the cylindrical cavity (and using the zero boundary conditions), one finds that the energy is composed of the sum of an integral and a finite number of other terms. The integral part of the expression only depends on the volume of the cavity and is proportional to a divergent integral, and hence drops out when the energy differences are taken. The only terms that yield non-zero contributions to the energy difference originate with $f^{(3)}(0)$ and depend on d. It is these terms that give rise to the Casimir force. This approach also showed that any particular choice made for the cut-off is irrelevant.

Mathematical Interlude:
The Euler-Maclaurin Summation Formula.

The Euler-Maclaurin formula allows one to accurately evaluate the difference of finite summations and their approximate evaluations in the form of integrals.

The Euler-Maclaurin Formula

If N is an integer and $f(x)$ is a smooth differentiable function defined for all real values of x between 0 and N, then the summation

$$ S = \sum_{n=1}^{N-1} f(n) \tag{1114} $$

can be approximated by an integral

$$ I = \int_0^N dx \, f(x) \tag{1115} $$

In particular, by utilizing the "trapedoizal rule", one expects that

$$ I \sim S + \frac{1}{2} \left(f(0) + f(N) \right) \tag{1116} $$

The Euler-Maclaurin formula provides expressions for the difference between the sum and the integral in terms of the higher-derivatives $f(n)$ at the end points of the interval 0 and N. For any integer p, one has

$$ S + \frac{1}{2} \left(f(0) + f(N) \right) - I = \sum_{n=1}^{p} \frac{B_{2n}}{(2n)!} \left(f^{2n-1}(N) - f^{2n-1}(0) \right) + R \tag{1117} $$

where B_1 = -1/2, B_2 = 1/6, B_3 = 0, B_4 = -1/30, B_5 = 0, B_6 = 1/42, B_7 = 0, B_8 = -1/30, ... are the Bernoulli numbers, and R is an error term which is normally small if the series on the right is truncated at a suitable value of p.

The Remainder Term

The remainder R when the series is truncated after p terms is given by

$$R = (-1)^p \int_0^N dx \ f^{(p+1)}(x) \frac{P_{p+1}(x)}{(p+1)!} \tag{1118}$$

where $P_n(x) = B_n(x - [x])$ are the periodic Bernoulli polynomials. The remainder term can be estimated as

$$| R | \leq \frac{2}{(2\pi)^p} \int_0^N dx \mid f^{2p-1}(x) \mid \tag{1119}$$

Derivation by Induction

First we shall examine the properties of the Bernoulli polynomials and the Bernoulli numbers. Then we shall indicate how the Euler-Maclaurin formula can be obtained by induction.

The Bernoulli polynomials $B_n(x)$, for n = 0, 1, 2, ... are defined by the generating function expansion

$$G(z,x) = \frac{z \ e^{zx}}{e^z - 1} = \sum_{n=0}^{\infty} B_n(x) \frac{z^n}{n!} \tag{1120}$$

Furthermore, when $x = 0$, one has

$$G(z,0) = \frac{z}{e^z - 1} = \sum_{n=0}^{\infty} B_n \frac{z^n}{n!} \tag{1121}$$

where B_n are the Bernoulli constants. Hence, the Bernoulli constants are the Bernoulli polynomials evaluated at $x = 0$, i.e. $B_n(0) = B_n$. Furthermore, on differentiating the generating function w.r.t. x, one finds

$$\frac{\partial G(z,x)}{\partial x} = z \ G(z,x) \tag{1122}$$

which implies that

$$\sum_{n=0}^{\infty} \frac{\partial B_n(x)}{\partial x} \frac{z^n}{n!} = z \sum_{n=0}^{\infty} B_n(x) \frac{z^n}{n!} \tag{1123}$$

On equating the coefficients of z^n in the above equation, one obtains the important relation

$$\frac{\partial B_n(x)}{\partial x} = n \ B_{n-1}(x) \tag{1124}$$

Therefore, by integration it easy to show that $B_n(x)$ are polynomials of degree n. The first few Bernoulli polynomials can be explicitly constructed from the

194

generating function expansion. The first few polynomials are given by

$$
\begin{aligned}
B_0(x) &= 1 \\
B_1(x) &= x - \frac{1}{2} \\
B_2(x) &= x^2 - x + \frac{1}{6} \\
B_3(x) &= x^3 - \frac{3}{2}x^2 + \frac{1}{2}x \\
B_4(x) &= x^4 - 2x^3 + x^2 - \frac{1}{30} \\
B_5(x) &= x^5 - \frac{5}{2}x^4 + \frac{5}{3}x^3 - \frac{1}{6}x \\
&\cdots
\end{aligned}
\tag{1125}
$$

From the generating function expansion, one can show that the Bernoulli polynomials are either even or odd functions of $x - \frac{1}{2}$. The generating function can be expressed

$$
G(z,x) = e^{z(x-\frac{1}{2})} \left(\frac{z}{e^{\frac{z}{2}} - e^{-\frac{z}{2}}} \right) = \sum_{n=0}^{\infty} B_n(x) \frac{z^n}{n!}
\tag{1126}
$$

where the second factor is an even function of z, thus, the generating function is invariant under the combined transformation $z \to -z$ and $(x - \frac{1}{2}) \to -(x - \frac{1}{2})$. Therefore, one has

$$
\sum_{n=0}^{\infty} B_n\left(\frac{1}{2} + x - \frac{1}{2}\right) \frac{z^n}{n!} = \sum_{n=0}^{\infty} B_n\left(\frac{1}{2} + \frac{1}{2} - x\right) (-1)^n \frac{z^n}{n!}
\tag{1127}
$$

so the polynomials satisfy

$$
B_n(x) = (-1)^n B_n(1-x)
\tag{1128}
$$

In particular for $x = 1$, one has

$$
B_n(1) = (-1)^n B_n(0)
\tag{1129}
$$

The generating function with $x = 0$ can be re-written as the sum of its even and odd parts

$$
G(z,0) = \left(\frac{\frac{z}{2}}{\tanh \frac{z}{2}} \right) - \frac{z}{2} = \sum_{n=0}^{\infty} B_n(0) \frac{z^n}{n!}
\tag{1130}
$$

The even part has only even terms in its Taylor expansion, and there is only one term in the odd part. Hence, the odd Bernoulli numbers vanish for $n > 1$, i.e. $B_{2n+1}(0) = 0$ for $n > 0$. Therefore, for $n \geq 2$, one has $B_n(0) = B_n(1)$. This equality can be used to evaluate the integrals of the Bernoulli polynomial over

the range from 0 to 1. On expressing the integral of $B_n(x)$ in terms of $B_{n+1}(x)$, one has

$$
\begin{aligned}
\int_0^1 dx\, B_n(x) &= \frac{1}{(n+1)} \int_0^1 dx\, \frac{\partial B_{n+1}(x)}{\partial x} \\
&= \frac{B_{n+1}(1) - B_{n+1}(0)}{(n+1)} \\
&= 0 \ \text{ for } n \geq 1
\end{aligned}
\tag{1131}
$$

Hence, the Bernoulli polynomials may be defined recursively via the relation

$$
\frac{\partial B_n(x)}{\partial x} = n\, B_{n-1}(x)
\tag{1132}
$$

if the constant of integration is fixed by

$$
\int_0^1 dx\, B_n(x) = 0 \ \text{ for } n \geq 1
\tag{1133}
$$

The periodic Bernoulli functions $P_n(x)$ can be defined by

$$
P_n(x) = B_n(x - [x])
\tag{1134}
$$

where $[x]$ is the integral part of x. This definition of $P_n(x)$ reproduces to the Bernoulli polynomials on the interval $(0,1)$ since $[x] = 0$ in this interval. The functions $P_n(x)$ are periodic over an extended range of x with period 1.

The Euler-Maclaurin formula can be obtained by mathematical induction. Consider the integral

$$
\int_n^{n+1} dx\, f(x) = \int_n^{n+1} dx\, u\, \frac{\partial v}{\partial x}
\tag{1135}
$$

with the identification of

$$
u = f(x)
\tag{1136}
$$

and

$$
\frac{\partial v}{\partial x} = 1 = P_0(x)
\tag{1137}
$$

since $P_0(x) = 1$. Therefore, on using the recursion relation involving the derivative of the Bernoulli polynomials, one finds that

$$
v = P_1(x)
\tag{1138}
$$

Integrating by parts, one obtains

$$
\int_n^{n+1} dx\, f(x) = [\, f(x)\, P_1(x)\,]_n^{n+1} - \int_n^{n+1} dx\, \frac{\partial f(x)}{\partial x}\, P_1(x)
\tag{1139}
$$

196

but since the periodic Bernoulli polynomial $P_1(x)$ is given by

$$P_1(x) = (x - [x]) - \frac{1}{2} \tag{1140}$$

it has the value of $1/2$ at the limits of integration. Hence, the integration reduces to

$$\int_n^{n+1} dx \, f(x) = \left(\frac{f(n+1) + f(n)}{2} \right) - \int_n^{n+1} dx \, \frac{\partial f(x)}{\partial x} \, P_1(x) \tag{1141}$$

Summing the above expression from $n = 1$ to $n = N - 1$, yields

$$\int_1^N dx \, f(x) = \left(\frac{f(1) + f(N)}{2} \right) + \sum_{n=2}^{N-1} f(n) - \int_1^N dx \, \frac{\partial f(x)}{\partial x} \, P_1(x) \tag{1142}$$

Adding $\left(\frac{f(1) + f(N)}{2} \right)$ to both sides of the equation and rearranging, one finds

$$\sum_{n=1}^N f(n) = \int_1^N dx \, f(x) + \left(\frac{f(1) + f(N)}{2} \right) + \int_1^N dx \, \frac{\partial f(x)}{\partial x} \, P_1(x) \tag{1143}$$

The last two terms, therefore, give the error when the sum is approximated by an integral. The first correction is simply the end point corrections from the "trapezoidal rule", and the second correction has to be evaluated to yield the Euler-Maclaurin formula. The last correction is of the form of an integral which can be expressed in terms of the sum of the integrals

$$\int_n^{n+1} dx \, f'(x) \, P_1(x) \tag{1144}$$

where the prime refers to the derivative of $f(x)$ w.r.t. x. The above expression can be evaluated by integrating by parts. The integrand is re-written as

$$\int_n^{n+1} dx \, f'(x) \, P_1(x) = \int_n^{n+1} dx \, u \, \frac{\partial v}{\partial x} \tag{1145}$$

where one identifies the two factors as

$$u = f'(x)$$
$$\frac{\partial v}{\partial x} = P_1(x) \tag{1146}$$

Since the indefinite integral is evaluated as

$$\int^x dx' \, P_1(x') = \frac{1}{2} \, P_2(x) \tag{1147}$$

197

the integration by parts yields

$$\int_n^{n+1} dx \, P_1(x) \, f'(x) \; = \; \left[\frac{P_2(x) \, f'(x)}{2} \right]_n^{n+1} - \frac{1}{2} \int_n^{n+1} dx \, f''(x) \, P_2(x) \quad (1148)$$

However, one has $P_2(0) = P_2(1) = B_2$, therefore the above expression simplifies to

$$\int_n^{n+1} dx \, P_1(x) \, f'(x) \; = \; B_2 \left(\frac{f'(n+1) - f'(n)}{2} \right) - \frac{1}{2} \int_n^{n+1} dx \, f''(x) \, P_2(x)$$
$$(1149)$$

Then, on summing the above expression from $n = 1$ to $n = N - 1$, one finds

$$\int_1^N dx \, P_1(x) \, f'(x) \; = \; B_2 \left(\frac{f'(N) - f'(1)}{2} \right) - \frac{1}{2} \int_1^N dx \, f''(x) \, P_2(x) \quad (1150)$$

This yields the first term in the series of end point corrections in the Euler-Maclaurin formula, where the correction is the sum of the first derivatives at the end points multiplied by $B_2/2!$. The above process can be iterated yielding a complete proof of the Euler-Maclaurin summation formula.

In order to get bounds on the size of the error when the sum is approximated by the integral, we note that the Bernoulli polynomials on the interval $[0, 1]$ attain their maximum absolute values at the endpoints and the value $B_n(1)$ is the n-th Bernoulli number.

References

T. M. Apostol, *"An Elementary View of Euler's Summation Formula"*, American Mathematical Monthly, **106**, 409-418 (1999).

D. H. Lehmer, *"On the Maxima and Minima of Bernoulli Polynomials"*, American Mathematical Monthly, **47**, 533-538 (1940).

11.3.2 The Lamb Shift

The Lamb shift is a shift between the energy levels of the $2\,^2S_{\frac{1}{2}}$ and the $2\,^2P_{\frac{1}{2}}$ levels of hydrogen from the predictions of the Dirac equation as they have the same n and j values ($j = l \pm s$). The Dirac equation predicts that these two levels should be degenerate. However, these levels were measured by Lamb and Retherford[63] who found that the $2\,^2S_{\frac{1}{2}}$ level is higher than the $2\,^2P_{\frac{1}{2}}$ level by 1058 MHz or 0.033 cm^{-1}. Bethe explained this in terms of the interaction between the bound electron and the quantized electromagnetic field[64]. Similar shifts should also occur between the $n\,^2S_{\frac{1}{2}}$ and the $n\,^2P_{\frac{1}{2}}$ levels, but the

---

[63]W. E. Lamb Jr. and R. C. Retherford, Phys. Rev.**72**, 241 (1947).
[64]H. A. Bethe, Phys. Rev. **72**, 339 (1947).

magnitude of the shifts should be much smaller, as the magnitude varies as n^{-3}.

Qualitatively, the electron interacts with the fluctuating electromagnetic field and with the potential due to the nucleus. The zero-point fluctuations cause the electron to deviate from its quantum orbit by an amount given by Δr and, therefore, experiences a potential given by

$$V(r + \Delta r) \ = \ V(r) \ + \ \Delta \underline{r} \ . \ \underline{\nabla} V(r) \ + \ \frac{1}{2!} \ (\ \Delta \underline{r} \ . \ \underline{\nabla} \ )^2 \ V(r) \ + \ldots \quad (1151)$$

so one expects an energy-shift given by

$$\Delta E \ = \ \frac{1}{3 \ . \ 2!} \ < \ \Delta r^2 \ > < \ nlm \mid \nabla^2 V(r) \mid nlm \ > \quad (1152)$$

Due to the form of the Coulomb potential

$$V(r) \ = \ - \ \frac{e^2}{r} \quad (1153)$$

the Laplacian is related to a point charge density at the nucleus

$$\nabla^2 V(r) \ = \ 4 \ \pi \ e^2 \ \delta(r) \quad (1154)$$

Hence, the shift due to the fluctuations in the electron's potential energy occurs primarily at the origin. The effect of the electromagnetic fluctuations on the kinetic energy are not state specific, and can be considered as a uniform shift of all the energy levels, like the electron's rest mass energy $m \ c^2$. Thus, the relative energy shift of the levels is solely determined by the potential at the origin. Therefore, the states with non-zero angular momenta do not experience the relative energy-shift since the electronic wave functions vanish at the origin. Thus, only the 2s state experiences a shift but the 2p state is unshifted.

The magnitude of the Lamb shift can be ascertained by expressing Δr in terms of the zero-point fluctuations in the electromagnetic field[65]. If it is assumed that the electron is bound to the atom harmonically, $\Delta \underline{r}$ is determined from the equation of motion

$$\Delta \underline{\ddot{r}} \ + \omega_0^2 \ \Delta \underline{r} \ = \ \frac{q}{m} \ \underline{E} \quad (1155)$$

where the electric field $\underline{E}$ has components that are fluctuating with wave vector $\underline{k}$ or equivalently with frequency ω. This has the result that the position fluctuates at the frequency ω with an amplitude given by

$$\Delta \underline{r_\omega} \ = \ \frac{q}{m} \ E_\omega \ \frac{1}{\omega_0^2 \ - \ \omega^2} \quad (1156)$$

[65]T. A. Welton, Phys. Rev. **74**, 1557 (1948).

$\Delta\underline{\mathbf{r}}(t)$
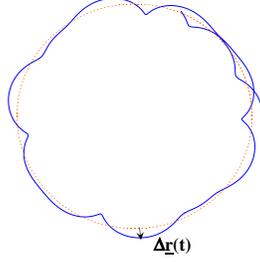
Figure 37: A cartoon depicting the modification of the classical orbit of an electron due to the zero-point fluctuations of the electromagnetic field.

where E_ω is the Fourier component of the fluctuating electric field. Hence, the ω-component of the mean squared fluctuation[66] in the particle's position is given by

$$< \mid \Delta\underline{r}^2_\omega \mid > \; = \; \left(\frac{q}{m} \right)^2 \; < \mid \frac{E^2_\omega}{(\, \omega^2_0 \, - \, \omega^2 \,)^2} \mid > \qquad (1157)$$

On approximating the electromagnetic energy associated with the fluctuating electromagnetic field $< \mid E^2_\omega \mid >$ by the half the sum of the zero-point energies of the photon modes, one has

$$\frac{V}{8 \, \pi} \; < \mid E^2_\omega \mid > \; = \; 2 \, \frac{1}{4} \, \hbar \, \omega \qquad (1158)$$

where the factor 2 represents the two types of polarization of the normal modes. Therefore, on summing over the normal modes, one finds that the mean squared deviation of the electron's trajectory from the classical orbit is proportional to

$$\frac{V}{(\, 2 \, \pi \, c \,)^3} \; \int d\Omega \; \int_0^\infty \; d\omega \; \omega^2 \; < \mid \frac{E^2_\omega}{(\, \omega^2_0 \, - \, \omega^2 \,)^2} \mid >$$
$$= \; \frac{4 \, \pi \, \hbar}{V} \; \frac{V}{(\, 2 \, \pi \, c \,)^3} \; \int d\Omega \; \int_0^\infty \; d\omega \; \frac{\omega^3}{(\, \omega^2_0 \, - \, \omega^2 \,)^2} \qquad (1159)$$

The integration over ω can be approximated as

$$\int_{\omega_0}^{\frac{mc^2}{\hbar}} \; \frac{d\omega}{\omega} \; = \; \ln \frac{m \, c^2}{\hbar \, \omega_0} \qquad (1160)$$

where an upper and lower cut-off have been introduced to prevent the integral from diverging[67]. The expectation value of the second derivative of the potential

---

[66]The average squared fluctuation of the electromagnetic field should, in principle, be calculated as an average over a volume in time and space which encompasses the electron's trajectory.

[67]The upper limit can be considered as being determined by the spatial dimension of the volume in which the electromagnetic fluctuations are being averaged over. The divergence at the lower limit of integration is unphysical and is caused by the neglect of the lifetimes of the electronic states. The inclusion of the lifetimes result in the integrand being finite at the resonance frequency ω_0

for the 2s state is given by

$$< \mid \nabla^2 V \mid > = 4 \pi e^2 \frac{1}{\pi a^3} \tag{1161}$$

where the second factor represents the 2s electron density at the origin. The corresponding factor for an ns level is expected to vary proportionally to n^{-3}. Combining the above expressions, one finds that the 2s level is shifted by an energy given by

$$\Delta E_{2s} = \frac{4}{2 \pi} \left(\frac{e^2}{\hbar c} \right)^3 \left(\frac{m e^4}{\hbar^2} \right) \ln \frac{m c^2}{\hbar \omega_0} \tag{1162}$$

where the frequency of the electron's orbit ω_0 has been chosen as a lower cut-off on the frequency of the electromagnetic fluctuations. Since the logarithmic factor is approximately given by $\sim -2 \ln \frac{e^2}{\hbar c}$, one can see that the above estimate is consistent with the Lamb shift having a magnitude of approximately 4.372 $\times 10^{-6}$ eV.

11.3.3 The Self-Energy of a Free Electron

The corrections to the energy of a free electron due to its coupling to the electromagnetic field are to be considered[68]. It shall be assumed that the electromagnetic field is in the ground state $\mid \{0\} >$, and the energy of an electron in a state with momentum $\underline{q}$ will be evaluated via perturbation theory.

The lowest-order correction to the electron's energy comes from the diamagnetic interaction. From first-order perturbation theory, one finds the correction

$$\Delta E_q^{(1)} = < \underline{q} \{0\} \mid \hat{H}_{dia} \mid \underline{q} \{0\} > \tag{1163}$$

On using a plane-wave to represent the electronic wave function

$$\psi_{\underline{q}}(\underline{r}) = \frac{1}{\sqrt{V}} \exp \left[i \underline{q} \cdot \underline{r} \right] \tag{1164}$$

then the first-order change in the electron's energy due to the coupling to the field is given by

$$\Delta E_q^{(1)} = \left(\frac{e^2}{2 m c^2} \right) \sum_{\underline{k},\alpha,\underline{k}',\alpha'} \left(\frac{2 \pi \hbar c^2}{V} \right) \frac{\hat{\epsilon}_\alpha(\underline{k}) \cdot \hat{\epsilon}_{\alpha'}(\underline{k}')}{\sqrt{\omega_k \omega_{k'}}} < \{0\} \mid a_{\underline{k}',\alpha'} a^\dagger_{\underline{k},\alpha} \mid \{0\} >$$

$$\times \frac{1}{V} \int d^3\underline{r} \exp \left[- i \underline{q} \cdot \underline{r} \right] \exp \left[i (\underline{k}' - \underline{k}) \cdot \underline{r} \right] \exp \left[i \underline{q} \cdot \underline{r} \right]$$

[68]W. Heisenberg and W. Pauli, Z. Physik, **56**, 1 (1929),
W. Heisenberg and W. Pauli, Z. Physik, **59**, 168 (1930).
I. Waller, Z. Physik, **59**, 168 (1930).
I. Waller, Z. Physik, **61**, 721 & 837 (1930).
I. Waller, Z. Physik, **62**, 673 (1930).
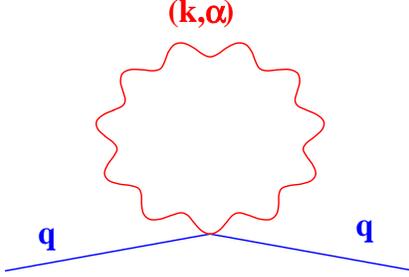
Figure 38: The first-order correction to the rest mass of the electron due to the diamagnetic interaction.

$$= \left(\frac{e^2}{2\ m\ c^2} \right) \sum_{\underline{k},\alpha,\underline{k},\alpha'} \left(\frac{2\ \pi\ \hbar\ c^2}{V} \right) \frac{\hat{\epsilon}_\alpha(\underline{k})\ .\ \hat{\epsilon}_{\alpha'}(\underline{k}')}{\sqrt{\omega_k\ \omega_{k'}}}\ \delta_{\underline{k},\underline{k}'}\ \delta_{\alpha,\alpha'} \tag{1165}$$

since the electronic matrix elements give rise to the condition of conservation of momentum. Hence, the correction to the energy is found as

$$\begin{aligned}
\Delta E_q^{(1)} &= \left(\frac{e^2}{2\ m\ c^2} \right) \frac{V}{(\ 2\ \pi\ )^3} \left(\frac{2\ \pi\ \hbar\ c}{V} \right) 2 \int d^3\underline{k}\ \frac{1}{k} \\
&= \left(\frac{e^2}{2\ m\ c^2} \right) \frac{V}{(\ 2\ \pi\ )^3} \left(\frac{2\ \pi\ \hbar\ c}{V} \right) 8\ \pi \int_0^\infty dk\ k \\
&= \left(\frac{e^2\ \hbar}{\pi\ m\ c} \right) \int_0^\infty dk\ k
\end{aligned} \tag{1166}$$

which diverges. This contribution is independent of the electron's momentum $\underline{q}$, and since $\underline{k} = \underline{k}'$ it can be seen that the contribution of the diamagnetic interaction to first-order is independent of the quantum state of the electron. This contribution to the electron's energy can be lumped together with the electron's rest-energy $m\ c^2$. However, since the corrections are being evaluated for non-relativistic electrons for which the rest energy is not observable, it is customary to ignore the rest-energy and, therefore, this correction shall no longer be considered.

The paramagnetic interaction, when taken to second-order, also yields a correction to the electron's self-energy. This correction can be considered to be due to a virtual process in which the electron emits a photon and then re-absorbs it. The second-order correction to the energy is evaluated from

$$\Delta E_q^{(2)} = \sum_{\underline{q}',\underline{k},\alpha} \frac{<\ \underline{q}\ \{0\}\ |\ \hat{H}_{para}\ |\ \underline{q}'\ 1_{\underline{k},\alpha}\ >\ <\ \underline{q}'\ 1_{\underline{k},\alpha}\ |\ \hat{H}_{para}\ |\ \underline{q}\ \{0\}\ >}{E_{\underline{q}'}\ +\ \hbar\ \omega_k\ -\ E_{\underline{q}}} \tag{1167}$$
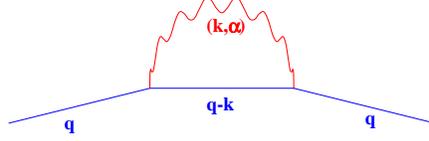
Figure 39: The second-order self-energy correction of a free electron due to the paramagnetic interaction. The electron with momentum $\underline{q}$ emits a virtual photon with momentum $\underline{k}$ and then reabsorbs it.

where $\mid \underline{q}' \, 1_{\underline{k},\alpha} \, >$ is a one-photon intermediate state of the electron-photon system. We assume that the process does not conserve energy, so that the denominator is finite. The matrix elements are evaluated as

$$
< \underline{q}' \, 1_{\underline{k},\alpha} \mid \hat{H}_{para} \mid \underline{q} \, \{0\} \, > \;\; = \;\; \sqrt{\frac{2 \, \pi \, \hbar \, c^2}{V \, \omega_k}} \; \hbar \; \hat{\epsilon}_\alpha(\underline{k}) \, . \, \underline{q}
$$

$$
\times \;\; \frac{1}{V} \int d^3\underline{r} \; \exp\left[\, - \, i \, \underline{q}' \, . \, \underline{r} \, \right] \exp\left[\, - \, i \, \underline{k} \, . \, \underline{r} \, \right] \exp\left[\, i \, \underline{q} \, . \, \underline{r} \, \right]
$$

$$
= \;\; \sqrt{\frac{2 \, \pi \, \hbar \, c^2}{V \, \omega_k}} \; \hbar \; \hat{\epsilon}_\alpha(\underline{k}) \, . \, \underline{q} \;\; \delta_{\underline{q}'+\underline{k}-\underline{q}} \tag{1168}
$$

which leads to momentum conservation. The second-order correction to the electron's energy takes the form

$$
\Delta E_q^{(2)} \;\; = \;\; \left(\frac{e^2}{m^2 \, c^2} \right) \sum_{\underline{k},\alpha} \left(\frac{2 \, \pi \, \hbar \, c^2}{V \, \omega_k} \right) \frac{\mid \hbar \, \underline{q} \, . \, \hat{\epsilon}_\alpha(\underline{k}) \mid^2}{\frac{\hbar^2 \, q^2}{2 \, m} - \frac{\hbar^2 \, (\underline{q}-\underline{k})^2}{2 \, m} - \hbar \, \omega_k} \tag{1169}
$$

On summing over the polarizations by using the diadic completeness relation[69]

$$
\sum_\alpha \hat{\epsilon}_\alpha(\underline{k}) \, \hat{\epsilon}_\alpha(\underline{k}) \;\; = \;\; \hat{I} \, - \, \hat{k} \, \hat{k} \tag{1170}
$$

---

[69]The completeness relation merely expresses the fact that any vector in a three-dimensional space can be expressed in terms of the components along three orthogonal directions $\hat{e}_i$

$$
\underline{A} \;\; = \;\; \sum_{i=1}^3 A_i \, \hat{e}_i
$$

where the components are given by the scalar product

$$
A_i \;\; = \;\; \underline{A} \, . \, \hat{e}_i
$$

Hence, the completeness relation follows as

$$
\underline{I} \;\; = \;\; \sum_i \hat{e}_i \, \hat{e}_i
$$

.

one finds that the numerator is given by

$$\sum_\alpha \hbar^2 \mid \underline{q} \cdot \hat{\epsilon}_\alpha(\underline{k}) \mid^2 = \hbar^2 q^2 (1 - \cos^2\theta)$$

(1171)

where θ is the angle between $\underline{q}$ and $\underline{k}$

$$\underline{q} \cdot \underline{k} = q k \cos\theta$$

(1172)

Hence, one has

$$
\begin{aligned}
\Delta E_q^{(2)} &= \left(\frac{e^2}{m^2 c^2} \right) \frac{V}{(2\pi)^3} \int d^3\underline{k} \left(\frac{2\pi \hbar c^2}{V \omega_k} \right) \frac{\hbar^2 q^2 (1 - \cos^2\theta)}{\frac{\hbar^2 q^2}{2 m} - \frac{\hbar^2 (q-k)^2}{2 m} - \hbar \omega_k} \\
&= \left(\frac{e^2 \hbar}{2\pi m^2 c} \right) \int_0^\infty dk \, k \int_0^\pi d\theta \, \sin\theta \, \frac{\hbar^2 q^2 (1 - \cos^2\theta)}{\frac{\hbar^2 q k}{m} \cos\theta - \frac{\hbar^2 k^2}{2 m} - \hbar c k}
\end{aligned}
$$

(1173)

This contribution can be written as being explicitly proportional to the kinetic energy of the electron, and a factor of k can be cancelled from the numerator and the denominator

$$\Delta E_q^{(2)} = \frac{\hbar^2 q^2}{2 m} \left(\frac{e^2}{\hbar c} \right) \frac{2}{\pi} \int_0^\infty dk \int_0^\pi d\theta \, \sin\theta \, \frac{(1 - \cos^2\theta)}{2 q \cos\theta - k - \frac{2 m c}{\hbar}}$$

(1174)

It should be evident that the integral diverges logarithmically at large k. The divergent part of the integral can be written as

$$
\begin{aligned}
\Delta E_q^{(2)} &\sim - \frac{\hbar^2 q^2}{2 m} \left(\frac{e^2}{\hbar c} \right) \frac{2}{\pi} \int_0^\pi d\theta \, \sin\theta \, (1 - \cos^2\theta) \int_{\frac{2mc}{\hbar}}^\infty \frac{dk}{k} \\
&= - \frac{\hbar^2 q^2}{2 m} \left(\frac{e^2}{\hbar c} \right) \frac{8}{3\pi} \int_{\frac{2mc}{\hbar}}^\infty \frac{dk}{k}
\end{aligned}
$$

(1175)

If an upper cut-off λ_+^{-1} is introduced, then the correction to the electron's kinetic energy can be estimated as

$$\Delta E_q^{(2)} = - \frac{\hbar^2 q^2}{2 m} \frac{8}{3\pi} \left(\frac{e^2}{\hbar c} \right) \ln \left(\frac{\hbar}{2 m c \lambda_+} \right)$$

(1176)

This shift can be interpreted as a (second-order) renormalization of the electron's mass from the un-renormalized mass to the physical mass m^*

$$\frac{1}{m^*} = \frac{1}{m} \left[1 - \frac{8}{3\pi} \left(\frac{e^2}{\hbar c} \right) \ln \left(\frac{\hbar}{2 m c \lambda_+} \right) + \dots \right]$$

(1177)

It is the renormalized mass m^* which would be determined by an experiment.

204

11.3.4 The Self-Energy of a Bound Electron

The Lamb shift (a quantum electrodynamic shift of the $2s$ level of Hydrogen upwards by 1058 MHz) is caused the self-energy of a bound electron. The self-energy of the state nlm can be estimated from second-order perturbation theory using the dipole approximation, as is appropriate for a completely non-relativistic calculation. The second-order shift is given by

$$\Delta E_{nlm}^{(2)} = \left(\frac{e^2}{m^2 c^2} \right) \sum_{\underline{k}, \alpha, n'l'm'} \left(\frac{2 \pi \hbar c^2}{V \omega_k} \right) \frac{| < n'l'm' | \hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{\hat{p}} | nlm > |^2}{E_{nlm} - E_{n'l'm'} - \hbar \omega_k}$$

(1178)

On summing over the polarizations using the completeness relation, one obtains

$$\Delta E_{nlm}^{(2)} = \left(\frac{e^2}{m^2 c^2} \right) \frac{\hbar c}{(2 \pi)} \int_0^\infty dk \, k \int_0^\pi d\theta_k \, \sin \theta_k \sum_{n'l'm'} \frac{| < n'l'm' | \underline{\hat{p}} | nlm > |^2 (1 - \cos^2 \theta_k)}{E_{nlm} - E_{n'l'm'} - \hbar \omega_k}$$

(1179)

where θ_k is the angle subtended between $\underline{k}$ and the matrix elements of $\underline{p}$. The angular integration can be performed, yielding

$$\Delta E_{nlm}^{(2)} = \left(\frac{e^2}{m^2 c^2} \right) \frac{2 \hbar c}{3 \pi} \int_0^\infty dk \, k \sum_{n'l'm'} \frac{| < n'l'm' | \underline{\hat{p}} | nlm > |^2}{E_{nlm} - E_{n'l'm'} - \hbar \omega_k}$$

(1180)

In the completely non-relativistic limit, the integration over k can be shown to be linearly divergent at the upper limit of integration.

Hans Bethe argued[70] that, within the same dipole approximation, the correction to the kinetic energy of the electron in the state $| nlm >$ is given by an expression analogous to that of an electron in a continuum state n

$$\Delta T_n^{(2)} = \frac{2}{3 \pi} \left(\frac{e^2}{\hbar c} \right) \left(\frac{\hbar}{m c} \right)^2 \int_0^\infty d\omega \, \omega \sum_{n'} \frac{| < n' | \underline{\hat{p}} | n > |^2}{E_n - E_{n'} - \hbar \omega}$$

(1181)

Since momentum is conserved for continuum states (on average), only the state where $n = n'$ contribute so the denominator simplifies[71]. The expression for the mass renormalization is divergent and is given by

$$\begin{aligned}
\Delta T_n^{(2)} &= - \frac{2}{3 \pi} \left(\frac{e^2}{\hbar c} \right) \left(\frac{\hbar}{m c} \right)^2 \int_0^\infty d\omega \, \omega \sum_{n'} \frac{| < n' | \underline{\hat{p}} | n > |^2}{\hbar \omega} \\
&= - \frac{4}{3 \pi} \left(\frac{e^2}{\hbar c} \right) \left(\frac{\hbar}{m c^2} \right) \int_0^\infty \frac{d\omega \, \omega}{\omega} \frac{< n | \hat{p}^2 | n >}{2 m}
\end{aligned}$$

(1182)

where the completeness relation has been used. This expression is valid if n labels either a continuum or a discrete state, since only the mass of the electron

---

[70]H. A. Bethe, Phys. Rev. **72**, 339 (1947).

[71]Since we are now using the dipole approximation, the recoil of the free electron which was taken into account in our previous analysis is now being ignored. [See the denominator of the first line of eqn(1173).]

is being altered and the expectation value of $\hat{p}$ is unaltered. Thus, Bethe argued, the kinetic energy of an electron in a bound state which has the physical mass m^* should be approximated as

$$
\begin{aligned}
< \ nlm \ | \ \frac{\hat{p}^2}{2 \ m^*} \ | \ nlm \ > \quad = \quad & < \ nlm \ | \ \frac{\hat{p}^2}{2 \ m} \ | \ nlm \ > \\
& - \frac{4}{3 \ \pi} \left(\frac{e^2}{\hbar \ c} \right) \left(\frac{\hbar}{m \ c^2} \right) \int_0^\infty \frac{d\omega \ \omega}{\omega} \ \frac{< \ nlm \ | \ \hat{p}^2 \ | \ nlm \ >}{2 \ m}
\end{aligned}
$$

$$(1183)$$

Now the bare Hamiltonian for an electron bound to a nucleus is given by

$$
\hat{H}_0 \ = \ \frac{\hat{p}^2}{2 \ m} \ + \ V(r) \tag{1184}
$$

and the unperturbed energy of the state $| \ nlm \ >$ of a hypothetical electron with mass m is calculated in the non-relativistic Schrödinger theory as

$$
E_{nlm}^{(0)} \ = \ < \ nlm \ | \ \frac{\hat{p}^2}{2 \ m} \ | \ nlm \ > \ + \ < \ nlm \ | \ V(r) \ | \ nlm \ > \tag{1185}
$$

However, in order to obtain a sensible numerical value for the approximate energy, $E_{nlm}^{(0)}$ has to be expressed in terms of the observed physical mass m^*. Therefore, the bare Hamiltonian has to be expressed in terms of the physical mass and compensating radiative corrections to the mass

$$
\begin{aligned}
E_{nlm}^{(0)} \quad = \quad & < \ nlm \ | \ \frac{\hat{p}^2}{2 \ m^*} \ | \ nlm \ > \ + \ < \ nlm \ | \ V(r) \ | \ nlm \ > \\
& + \frac{4}{3 \ \pi} \left(\frac{e^2}{\hbar \ c} \right) \left(\frac{\hbar}{m \ c^2} \right) \int_0^\infty \frac{d\omega \ \omega}{\omega} \ \frac{< \ nlm \ | \ \hat{p}^2 \ | \ nlm \ >}{2 \ m} \\
= \quad & < \ nlm \ | \ \frac{\hat{p}^2}{2 \ m^*} \ | \ nlm \ > \ + \ < \ nlm \ | \ V(r) \ | \ nlm \ > \\
& + \frac{4}{3 \ \pi} \left(\frac{e^2}{\hbar \ c} \right) \left(\frac{\hbar}{m \ c^2} \right) \int_0^\infty \frac{d\omega \ \omega}{\omega} \ \sum_{n'l'm'} \frac{| \ < \ n'l'm' \ | \ \hat{p} \ | \ nlm \ > \ |^2}{2 \ m}
\end{aligned}
$$

$$(1186)$$

The completeness relation was used in obtaining the last line. The second term in the above expression for the unperturbed energy is the correction due to the mass renormalization[72] which should be combined with the second-order radiative correction. The total energy (to second-order) is given by

$$
E_{nlm} \quad = \quad E_{n,l,m}^{(0)} \ + \ \Delta E_{n,l,m}^{(2)}
$$

---

[72]Renormalization is an idea which Bethe attributed to H. A. Kramers. Kramers had proposed that physical quantities should be expressed in terms of observable quantities, with all mention of bare quantities removed. Kramers was advocating a classical treatment from which Bethe created a non-relativistic quantum treatment.

$$= \quad < nlm \mid \frac{\hat{p}^2}{2\, m^*} \mid nlm > \; + \; < nlm \mid V(r) \mid nlm >$$

$$+ \frac{2}{3\,\pi} \left(\frac{e^2}{\hbar\, c} \right) \left(\frac{\hbar}{m\, c} \right)^2 \int_0^\infty d\omega\, \omega \sum_{n'l'm'} \frac{\mid < n'l'm' \mid \hat{p} \mid nlm > \mid^2}{\hbar\, \omega}$$

$$+ \frac{2}{3\,\pi} \left(\frac{e^2}{\hbar\, c} \right) \left(\frac{\hbar}{m\, c} \right)^2 \int_0^\infty d\omega\, \omega \sum_{n'l'm'} \frac{\mid < n'l'm' \mid \underline{\hat{p}} \mid nlm > \mid^2}{E_{nlm}\; -\; E_{n'l'm'}\; -\; \hbar\, \omega}$$

$$(1187)$$

The overall (second-order) shift from the Schrödinger estimate of the energy for the state $\mid nlm >$ (as calculated with the physical mass) is given by the sum of the last two terms, which is expressed as

$$\Delta E_{nlm}^{\text{shift}} \quad = \quad \frac{2}{3\,\pi} \left(\frac{e^2}{\hbar\, c} \right) \left(\frac{\hbar}{m\, c} \right)^2 \int_0^\infty d\omega\, \omega \sum_{n'l'm'} \frac{\mid < n'l'm' \mid \underline{\hat{p}} \mid nlm > \mid^2 \; (E_{nlm}\; -\; E_{n'l'm'})}{(E_{nlm}\; -\; E_{n'l'm'}\; -\; \hbar\, \omega)\, \hbar\, \omega}$$

$$(1188)$$

The integration over ω is logarithmically divergent, and can be made to converge by introducing an upper cut-off $\omega_+ \; = \; c\, \lambda^{-1}$. Therefore, the difference of the linearly divergent self-energy of the bound electron and the linearly divergent self-energy of the free electron is only logarithmically divergent. After introducing the cut-off, one finds the result

$$\Delta E_{nlm}^{\text{shift}} \quad = \quad - \frac{2}{3\,\pi} \left(\frac{e^2}{\hbar\, c} \right) \sum_{n'l'm'} \frac{\mid < n'l'm' \mid \underline{\hat{p}} \mid nlm > \mid^2}{m^2\, c^2}$$

$$\times \; (E_{nlm}\; -\; E_{n'l'm'})\; \ln \left| \frac{\hbar c \lambda^{-1} + E_{n'l'm'} - E_{nlm}}{E_{n'l'm'} - E_{nlm}} \right|$$

$$(1189)$$

The square of the mass m can safely be replaced by the square of the physical mass m^* in the expression for the energy shift, since we are only working to first order in $\frac{e^2}{\hbar\, c}$. All other quantities have been expressed in terms of the physical mass. If the rest energy of the electron is used as the upper cut-off energy $m\, c^2 \sim 0.5 \times 10^6$ eV, and assuming that the averaged logarithm of the electron excitation energy corresponds to an energy of the order of 17.8 Ryd, then the logarithm has a value of about 7.63 and is not sensitive to the precise value of $E_{nlm} \; - \; E_{n'l'm'}$ and, therefore, can be taken outside the summation

$$\Delta E_{nlm}^{\text{shift}} \quad = \quad - \frac{2}{3\,\pi} \left(\frac{e^2}{\hbar\, c} \right) \ln \left| \frac{2\, \hbar^2\, c^2}{Z^2\, e^4} \right|$$

$$\times \; \sum_{n'l'm'} \frac{\mid < n'l'm' \mid \underline{\hat{p}} \mid nlm > \mid^2}{m^2\, c^2} \; (E_{nlm}\; -\; E_{n'l'm'})$$

$$(1190)$$

As later shown by Dyson[73], that divergences found in any order in $\frac{e^2}{\hbar\, c}$ can

[73]F. J. Dyson, Phys. Rev. **75**, 1736 (1949).

be removed by consistently using the ideas of mass and charge renormalization[74]. Hence, a completely consistent relativistic theory does yield a finite shift, without the need to invoke any cut-off[75]. The weighted sum over the matrix elements can be evaluated by expressing it in terms of an expectation value involving commutators of $\hat{H}_0$ with $\hat{\underline{p}}$. That is

$$\sum_{n'l'm'} \, | \, < \, n'l'm' \, | \, \hat{\underline{p}} \, | \, nlm \, > \, |^2 \, (\, E_{nlm} \, - \, E_{n'l'm'} \,)$$

$$= \, \sum_{n'l'm'} \, < \, nlm \, | \, \hat{\underline{p}} \, | \, n'l'm' \, > \, < \, n'l'm' \, | \, [\, \hat{\underline{p}} \, , \, \hat{H}_0 \,] \, | \, nlm \, > \quad (1191)$$

and using the completeness relation, one obtains

$$= \, < \, nlm \, | \, \hat{\underline{p}} \, [\, \hat{\underline{p}} \, , \, \hat{H}_0 \,] \, | \, nlm \, > \qquad (1192)$$

Likewise,

$$\sum_{n'l'm'} \, | \, < \, n'l'm' \, | \, \hat{\underline{p}} \, | \, nlm \, > \, |^2 \, (\, E_{nlm} \, - \, E_{n'l'm'} \,)$$

$$= \, - \, \sum_{n'l'm'} \, < \, nlm \, | \, [\, \hat{\underline{p}} \, , \, \hat{H}_0 \,] \, | \, n'l'm' \, > \, < \, n'l'm' \, | \, \hat{\underline{p}} \, | \, nlm \, >$$

$$(1193)$$

which results in

$$= \, - \, < \, nlm \, | \, [\, \hat{\underline{p}} \, , \, \hat{H}_0 \,] \, \hat{\underline{p}} \, | \, nlm \, > \qquad (1194)$$

Thus, on adding the above two equations and dividing by two, one finds

$$\sum_{n'l'm'} \, | \, < \, n'l'm' \, | \, \hat{\underline{p}} \, | \, nlm \, > \, |^2 \, (\, E_{nlm} \, - \, E_{n'l'm'} \,)$$

$$= \, \frac{1}{2} \, < \, nlm \, | \, [\, \hat{\underline{p}} \, , \, [\, \hat{\underline{p}} \, , \, \hat{H}_0 \,] \,] \, | \, nlm \, > \qquad (1195)$$

On substituting

$$\hat{\underline{p}} \, = \, - \, i \, \hbar \, \underline{\nabla} \qquad (1196)$$

and

$$\hat{H}_0 \, = \, \frac{\hat{p}^2}{2 \, m} \, + \, V(r) \qquad (1197)$$

---

[74] This statement does not imply that a properly renormalized perturbation theory is convergent. In fact, one may argue that if the coupling constant changed sign then systems containing electrons would be unstable to BCS pairing. Since the radius of convergence of any expansion is limited by the closest singularity, perturbation theory may only have a zero radius of convergence. In this case, the theory may be expected to contain non-analytic terms of the form $\exp[\, - \, \hbar \, c / \, e^2]$.

[75] F. J. Dyson, Phys. Rev. **173**, 617 (1948).

into the expression for the matrix elements, one obtains

$$
\sum_{n'l'm'} | < n'l'm' \,|\, \underline{\hat{p}} \,|\, nlm > |^2 \,(\, E_{nlm} \, - \, E_{n'l'm'} \,)
$$

$$
= \; - \; \frac{\hbar^2}{2} \; \int \, d^3\underline{r} \; \psi_{nlm}(\underline{r}) \; \nabla^2 V(r) \; \psi_{nlm}(\underline{r}) \tag{1198}
$$

Substituting the expressions for the matrix elements into the expression for the Lamb-shift yields

$$
\Delta E_{nlm}^{\text{shift}} \;\; = \;\; \frac{2}{3\,\pi} \, \left(\frac{e^2}{\hbar\, c} \right) \left(\frac{\hbar^2}{m^2\, c^2} \right) \, \ln \left| \frac{2\,\hbar^2\, c^2}{Z^2\, e^4} \right| \; < nlm \,|\, \nabla^2 V(r) \,|\, nlm >
$$

$$
\tag{1199}
$$

Thus, the energy-shift only occurs for bound electrons as the expectation value of the Laplacian of the potential will vanish for extended states. For a hydrogenic-like atom

$$
\nabla^2 V(r) \;=\; 4\,\pi\, Z\, e^2\, \delta^3(\underline{r}) \tag{1200}
$$

so

$$
\Delta E_{nlm}^{\text{shift}} \;\; = \;\; \frac{4\, Z\, e^2}{3} \, \left(\frac{e^2}{\hbar\, c} \right) \left(\frac{\hbar^2}{m^2\, c^2} \right) \, | \, \psi_{nlm}(0) \, |^2 \; \ln \left| \frac{2\,\hbar^2\, c^2}{Z^2\, e^4} \right| \tag{1201}
$$

Therefore, the Lamb shift only occurs for electrons with $l = 0$, since electronic wave functions with $l \neq 0$ vanish at the origin. The atomic wave function at the position of the nucleus is given by

$$
| \, \psi_{n00}(0) \, |^2 \;=\; \frac{1}{\pi} \, \left(\frac{Z}{n\, a} \right)^3 \tag{1202}
$$

This yields Bethe's estimate for the Lamb shift as

$$
\Delta E_{n00}^{\text{shift}} \;\; = \;\; \frac{4}{3\,\pi\, n^3} \, \left(\frac{e^2}{\hbar\, c} \right)^3 \left(\frac{Z^4\, e^4\, m}{\hbar^2} \right) \, \ln \left| \frac{2\,\hbar^2\, c^2}{Z^2\, e^4} \right| \tag{1203}
$$

The above formulae leads to the estimate of 1040 MHz which is in good agreement with the experimentally determined value[76]. The exact relativistic calculation[77] yields the result

$$
\Delta E_{n00}^{\text{shift}} \;\; = \;\; \frac{4}{3\,\pi\, n^3} \, Z^4 \, \left(\frac{e^2}{\hbar\, c} \right)^5 m\, c^2 \, \left(\, \ln \left| \frac{m\, c^2}{2\,\hbar\, \omega_{n,n'}} \right| \, + \, \frac{31}{120} \, \right) \tag{1204}
$$

where the mc^2 in the logarithm comes from the Dirac theory without invoking any cut-off. The most recent experimentally measured value[78] is 1057.851 MHz which is in good agreement with the theoretical value of 1057.857 MHz.

---

[76] W. E. Lamb Jr. and R. E. Retherford, Phys. Rev. **72**, 241 (1947).

[77] N. M. Kroll and W. E. Lamb Jr., Phys. Rev. **75**, 388 (1949).

[78] G. C. Bhatt and H. Grotch, Ann. Phys. **187**, 1 (1987).

11.3.5 Brehmstrahlung

Accelerating (or decelerating) charged particles radiate. We shall consider the radiation emitted by a charged particle (such as an electron) that scatters from a massive charged particle via the Coulomb interaction. It is assumed that the mass M of the massive charged particle (in most cases, this is a nucleus) is significantly greater than the electron mass, so that the recoil of the nucleus can be neglected. The (instantaneous) Coulomb interaction between the electron and the nucleus is given by

$$V(r) = -\frac{Z\,e^2}{r} \tag{1205}$$

The Hamiltonian of the unperturbed electron is simply the kinetic energy. The incident electron is assumed to have a momentum $\underline{q}$ and the scattered electron has momentum $\underline{q}'$ and the cross-section for the scattering process will be calculated via low-order perturbation theory.

Rutherford Scattering

To second-order, the scattering cross-section is expressed as Rutherford scattering which is elastic and, therefore, involves no emission of photons. The
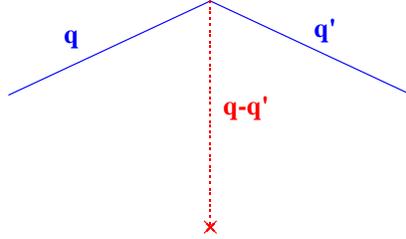
Figure 40: The Rutherford scattering process.

Rutherford scattering cross-section is found from the Fermi-Golden rule decay rate

$$\left(\frac{1}{\tau}\right)_{\text{Rutherford}} = \frac{2\,\pi}{\hbar}\,|\,<\underline{q}'\,|\,V(r)\,|\,\underline{q}\,>\,|^2\,\delta(\,E_q\,-\,E_{q'}\,) \tag{1206}$$

The matrix elements of the Coulomb potential is evaluated as

$$<\underline{q}'\,|\,V(r)\,|\,\underline{q}\,> = -\frac{4\,\pi\,Z\,e^2}{V\,|\,\underline{q}\,-\,\underline{q}'\,|^2} \tag{1207}$$

On integrating over the magnitude of the scattered electron's momentum, one obtains

$$\left(\frac{1}{\tau_{d\Omega'}}\right)_{\text{Rutherford}} = \frac{2\,\pi}{\hbar}\,\frac{V}{(\,2\,\pi\,)^3}\,d\Omega'\,\int_0^\infty dq'\,q'^2\,\left(\frac{4\,\pi\,Z\,e^2}{V\,|\,\underline{q}\,-\,\underline{q}'\,|^2}\right)^2\,\delta(\,E_q\,-\,E_{q'}\,)$$

$$= \frac{2\pi}{\hbar} \frac{V}{(2\pi)^3} \frac{m}{\hbar^2} q \left(\frac{4\pi Z e^2}{V \mid \underline{q} - \underline{q}' \mid^2} \right)^2 d\Omega'$$

$$(1208)$$

The denominator in the potential has to be evaluated on the energy shell. On introducing the scattering angle θ' and using the elastic scattering condition

$$\mid \underline{q} - \underline{q}' \mid^2 = 2 q^2 (1 - \cos\theta')$$

$$= 4 q^2 \sin^2 \frac{\theta'}{2} \qquad (1209)$$

one finds
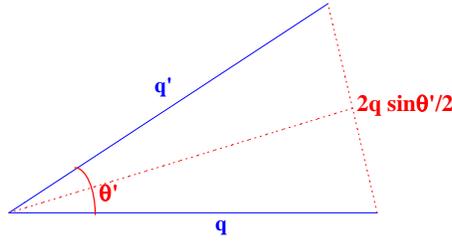
Figure 41: The geometry for Rutherford scattering. For elastic scattering, the magnitude of the initial momentum $\underline{q}$ is equal to the magnitude of the final momentum $\underline{q}'$ and the scattering angle is θ'.

$$\left(\frac{1}{\tau_{d\Omega'}} \right)_{\text{Rutherford}} = \frac{2\pi}{\hbar} \frac{V}{(2\pi)^3} \frac{m}{\hbar^2} q \left(\frac{4\pi Z e^2}{V 4 q^2 \sin^2 \frac{\theta'}{2}} \right)^2 d\Omega'$$

$$(1210)$$

On diving the scattering rate by the incident flux $\mathcal{F}$ of electrons

$$\mathcal{F} = \frac{\hbar q}{m V} \qquad (1211)$$

the elastic scattering cross-section is found to be given by

$$\left(\frac{d\sigma}{d\Omega'} \right)_{\text{Rutherford}} = \frac{2\pi}{\hbar} \frac{V}{(2\pi)^3} \frac{V m^2}{\hbar^3} \left(\frac{4\pi Z e^2}{V 4 q^2 \sin^2 \frac{\theta'}{2}} \right)^2$$

$$= \left(\frac{m Z e^2}{2 \hbar^2 q^2 \sin^2 \frac{\theta'}{2}} \right)^2 \qquad (1212)$$

which is the Rutherford scattering cross-section for electrons. The scattering cross-section diverges at $\theta' = 0$ and is always finite at $\theta' = \pi$ no matter how large q is. The scattering at $\theta' = \pi$ is known as back-scattering, and is caused
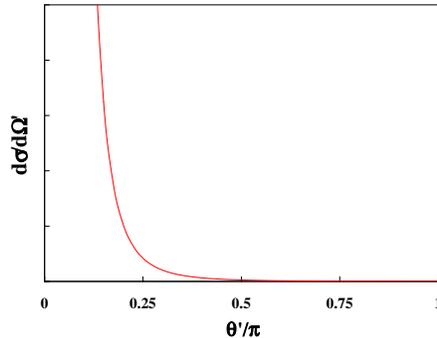
Figure 42: The scattering angle dependence of the differential scattering cross-section.

by the extremely high potential experienced by electrons with very small impact parameters. It was the large cross-section for back-scattering of charged α-particles from atoms, found by H. Geiger and E. Marsden in 1913[79], that was instrumental in verifying Rutherford's 1911 conjecture[80] that atoms have nuclei which are of very small spatial extent. The divergence in the scattering cross-section at $\theta' = 0$ is due to the long-ranged nature of the Coulomb interaction, which causes electrons to undergo scattering (no matter how slight the scattering is) at arbitrarily large distances from the nucleus.

Brehmstrahlung

Elastic scattering of electrons by the Coulomb potential is highly unlikely, since from classical electrodynamics it is known that accelerated particles radiate. Hence, it is expected that photons should be emitted in this process. This phenomenon is known as Brehmstrahlung. We shall calculate the Brehmstrahlung scattering cross-section[81] using low-order perturbation theory. The electron is scattered between the free electron eigenstates due to a perturbation which is a linear superposition of the Coulomb interaction with the nucleus and the paramagnetic interaction.

The lowest-order probability amplitude describing Brehmstrahlung is a linear superposition of two processes. These are:

(a) Scattering of an electron from the nucleus followed by the emission of a photon. The initial state of the electron is assumed to have momentum $\underline{q}$ and the final state of the electron is given by $\underline{q}'$ while the emitted photon has momentum $\underline{k}$. Therefore, from conservation of momentum, the momentum of the electron

---

[79]H. Geiger and E. Marsden, Phil. Mag. **25**, 1798 (1913).
[80]E. Rutherford, Phil. Mag. **21**, 669 (1911).
[81]H. A. Bethe and W. Heitler, Proc. Roy. Soc. **A 146**, 82 (1934).
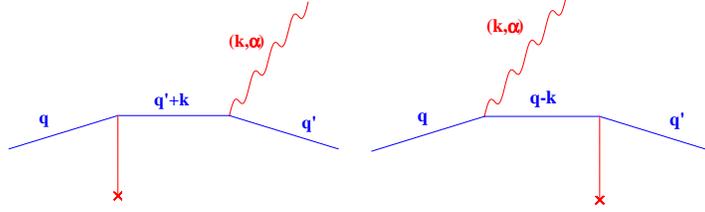
Figure 43: The two lowest-order processes contributing to Bremstrahlung.

in the intermediate state is given by $\underline{q}' + \underline{k}$.

(b) Emission of a photon followed by scattering from the nucleus. Conservation of momentum indicates that the intermediate state has momentum given by $\underline{q} - \underline{k}$.

The matrix elements for these second-order processes are given by

$$
\begin{aligned}
M_a &= \left(\frac{4 \pi Z e^2}{V \mid \underline{q} - \underline{q}' - \underline{k} \mid^2} \right) \left(\frac{e \hbar}{m c} \right) \sqrt{\frac{2 \pi \hbar c^2}{V \omega_k}} \; \hat{\epsilon}_\alpha(\underline{k}) \cdot (\underline{q}' + \underline{k}) \\
&\times \left(\frac{1}{E_{\underline{q}} - E_{\underline{q}'+\underline{k}} + i \eta} \right)
\end{aligned}
\tag{1213}
$$

and

$$
\begin{aligned}
M_b &= \left(\frac{4 \pi Z e^2}{V \mid \underline{q} - \underline{q}' - \underline{k} \mid^2} \right) \left(\frac{e \hbar}{m c} \right) \sqrt{\frac{2 \pi \hbar c^2}{V \omega_k}} \; \hat{\epsilon}_\alpha(\underline{k}) \cdot (\underline{q} - \underline{k}) \\
&\times \left(\frac{1}{E_{\underline{q}} - E_{\underline{q}-\underline{k}} - \hbar \omega_k + i \eta} \right)
\end{aligned}
\tag{1214}
$$

It should be noted that the numerators of the matrix elements simplify because the photons have transverse polarizations

$$
\epsilon_\alpha(\underline{k}) \cdot \underline{k} = 0
\tag{1215}
$$

From the energy conserving delta function in the expression for the decay rate, one finds

$$
E_{\underline{q}} = E_{\underline{q}'} + \hbar \omega_k
\tag{1216}
$$

hence the first energy-denominator can be expressed in a similar form to the second

$$
E_{\underline{q}} - E_{\underline{q}'+\underline{k}} = E_{\underline{q}'} - E_{\underline{q}'+\underline{k}} + \hbar \omega_k
\tag{1217}
$$

For small k, the energy-denominators can be expanded, yielding

$$
E_{\underline{q}'} - E_{\underline{q}'+\underline{k}} + \hbar \omega_k = \hbar \omega_k - \frac{\hbar}{m} \underline{q}' \cdot \underline{k} - \frac{\hbar^2 k^2}{2 m}
\tag{1218}
$$

213

and

$$E_{\underline{q}} - E_{\underline{q}-\underline{k}} - \hbar \, \omega_k = - \hbar \, \omega_k + \frac{\hbar}{m} \, \underline{q} \cdot \underline{k} - \frac{\hbar^2 \, k^2}{2 \, m} \tag{1219}$$

Since the energy of the photon cannot exceed the energy of the initial electron, one must have $q > k$, so the third term is smaller than the second term. Due to the large magnitude of c compared with the electron velocities $\frac{\hbar \, q}{m}$, the second and third terms can be neglected. Therefore, the photon-energy dominates both the energy-denominators. On substituting the above expressions in the sum of the matrix elements, one finds

$$\begin{aligned}
M_a + M_b &= \left(\frac{4 \, \pi \, Z \, e^2}{V \mid \underline{q} - \underline{q}' - \underline{k} \mid^2} \right) \left(\frac{e}{m \, c} \right) \sqrt{\frac{2 \, \pi \, \hbar \, c^2}{V \, \omega_k}} \\
&\quad \times \left(\frac{\hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{q}' \, \hbar}{E_{\underline{q}'} - E_{\underline{q}'+\underline{k}} + \hbar \, \omega_k + i \, \eta} + \frac{\hat{\epsilon}_\alpha(\underline{k}) \cdot \underline{q} \, \hbar}{E_{\underline{q}} - E_{\underline{q}-\underline{k}} - \hbar \, \omega_k + i \, \eta} \right) \\
&\approx \left(\frac{4 \, \pi \, Z \, e^2}{V \mid \underline{q} - \underline{q}' - \underline{k} \mid^2} \right) \left(\frac{e}{m \, c} \right) \sqrt{\frac{2 \, \pi \, \hbar \, c^2}{V \, \omega_k}} \\
&\quad \times \left(\frac{\hat{\epsilon}_\alpha(\underline{k}) \cdot (\, \underline{q}' - \underline{q} \,)}{\omega_k} \right)
\end{aligned} \tag{1220}$$

Using this approximation for the matrix elements, the transition rate is given by

$$\begin{aligned}
\frac{1}{\tau} &= \frac{2 \, \pi}{\hbar} \sum_{\underline{q}'} \sum_{\underline{k},\alpha} \left(\frac{4 \, \pi \, Z \, e^2}{V \mid \underline{q} - \underline{q}' - \underline{k} \mid^2} \right)^2 \left(\frac{e}{m \, c} \right)^2 \left(\frac{2 \, \pi \, \hbar \, c^2}{V \, \omega_k} \right) \\
&\quad \times \left| \frac{\hat{\epsilon}_\alpha(\underline{k}) \cdot (\, \underline{q}' - \underline{q} \,)}{\omega_k} \right|^2 \delta(\, E_{\underline{q}} - E_{\underline{q}'} - \hbar \, \omega_k \,)
\end{aligned} \tag{1221}$$

The terms proportional to k in the Coulomb scattering terms can be neglected, for low k values. The inelastic scattering cross-section for Brehmstrahlung is found by replacing the sums over $\underline{q}'$ and $\underline{k}'$ by integrals, and dividing by the incident flux of electrons. This procedure results in the expression

$$\begin{aligned}
\left(\frac{d^2\sigma}{d\Omega' \, d\omega_k} \right)_{\text{Brehmse}} &= \frac{q'}{q} \left(\frac{2 \, m \, Z \, e^2}{\hbar^2 \mid \underline{q} - \underline{q}' \mid^2} \right)^2 \sum_\alpha \int \frac{d\Omega_k}{4 \, \pi^2 \, \omega_k} \left(\frac{e^2}{\hbar \, c} \right) \\
&\quad \times \left| \frac{\hbar \, \hat{\epsilon}_\alpha(\underline{k}) \cdot (\, \underline{q}' - \underline{q} \,)}{m \, c} \right|^2
\end{aligned} \tag{1222}$$

If the angular distributions of the emitted photon $(d\Omega_k)$ and the scattered electron $(d\Omega')$ are both measured, the scattering cross-section can be represented as

$$\left(\frac{d^3\sigma}{d\Omega' \, d\Omega_k \, d\omega_k} \right)_{\text{Brehmse}} = \frac{q'}{q} \left(\frac{d\sigma}{d\Omega'} \right)_{\text{Rutherford}}$$

214

$$\times \frac{1}{4\,\pi^2\,\omega_k} \left(\frac{e^2}{\hbar\,c}\right) \sum_\alpha \left| \frac{\hbar\,\hat{\epsilon}_\alpha(\underline{k})\,.\,(\,\underline{q}'\,-\,\underline{q}\,)}{m\,c} \right|^2$$

$$(1223)$$

where the second factor is the probability of emitting a photon with energy $\hbar\,\omega_k$ into solid angle $d\Omega_k$. On summing over the polarization α and integrating over the directions of the emitted photon, one obtains

$$\left(\frac{d^2\sigma}{d\Omega'\,d\omega_k}\right)_{\text{Brehmse}} = \frac{q'}{q} \left(\frac{2\,m\,Z\,e^2}{\hbar^2\,|\,\underline{q}\,-\,\underline{q}'\,|^2}\right)^2$$

$$\times \frac{2}{3\,\pi\,\omega_k} \left(\frac{e^2}{\hbar\,c}\right) \left| \frac{\hbar\,(\,\underline{q}'\,-\,\underline{q}\,)}{m\,c} \right|^2 \quad (1224)$$

Hence, the scattering rate which includes the emission of a photon of energy $\hbar\,\omega_k$ is given by the product of the Rutherford scattering rate with a factor

$$\frac{q'}{q}\,\frac{2}{3\,\pi\,\omega_k} \left(\frac{e^2}{\hbar\,c}\right) \left(\frac{2\,q\,\hbar\,\sin\frac{\theta'}{2}}{m\,c}\right)^2 \qquad (1225)$$

This particular factorization of the cross-section involving the simultaneous emission of a soft photon is common to many processes involving the emission of low-energy bosons. The soft-photon theorem[82] shows that properties of the emitted low-energy photon is insensitive to anything except the global properties (such as the total charge or total magnetic moment) of the scattered particle. The cross-section involving the emission of a low-energy photon diverges as $\omega_k \rightarrow 0$, due to the factor of ω_k^{-1} in eqn(1224). This type of divergence is an infrared divergence. What this implies is that, in Brehmstrahlung, arbitrary large numbers of low-energy photons are emitted. Furthermore, similar singularities are also found in the $\omega = 0$ limit when elastic scattering corrections to the Rutherford scattering process are considered[83]. In any experiment with finite energy resolution, elastic scattering and very low-energy quasi-elastic scattering processes cannot be distinguished, so it is might be expected that the elastic scattering and quasi-elastic scattering divergences should be combined.

The divergences found in the problem of Brehmstrahlung were first considered by Bloch and Nordsieck[84] who showed that the infra-red divergences cancel. That is, the infra-red divergence does not exist[85]. The cancelation was achieved adding virtual emission processes for Rutherford scattering to the Brehmstrahlung cross-section for the emission of photons of energy less than ω_0, since these processes cannot be distinguished for sufficiently small photon

[82]F. F. Low, Phys. Rev. **96**, 1428 (1958).

[83]R. H. Dalitz, Proc. Roy. Soc. **A 206**, 509 (1950).

[84]F. Bloch and A. Nordsieck, Phys. Rev. **52**, 54 (1937).

[85]Since there are an infinite number of low-energy photons present in Brehmstrahlung, then it is expected that the classical limit of quantum theory applies so that classical electromagnetic theory should produce exact results.

frequencies ω_0. That is, on introducing an infra-red cut-off λ_-, one finds that the total inelastic scattering in which a photon with frequency less than ω_0 is emitted is given by

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{Brehmse}} = \left(\frac{d\sigma}{d\Omega}\right)_{\text{Rutherford}} \left[\frac{1}{2\pi} \left(\frac{e^2}{\hbar c}\right) \left(A \ln \frac{2\omega_0 \lambda_-}{c} + \dots \right) + \dots \right]$$
(1226)

where the factor A depends on the initial and final momentum of the electron. This result is logarithmically divergent as $\lambda_- \to 0$. On the other hand, to the same order, the elastic scattering cross-section is found as

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{Elastic}} = \left(\frac{d\sigma}{d\Omega}\right)_{\text{Rutherford}} \left[1 + \frac{1}{2\pi} \left(\frac{e^2}{\hbar c}\right) \left(A \ln \frac{\hbar}{\lambda_- m c} + \dots \right) + \dots \right]$$
(1227)

Hence, on combining the results, one finds that the quasi-elastic scattering cross-section is given by

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{Quasi-Elastic}} = \left(\frac{d\sigma}{d\Omega}\right)_{\text{Rutherford}} \left[1 + \frac{1}{2\pi} \left(\frac{e^2}{\hbar c}\right) \left(A \ln \frac{2\hbar \omega_0}{m c^2} + \dots \right) + \dots \right]$$
(1228)

so the cut-off λ_- cancels and the scattering cross-section does not diverge logarithmically. With this reasoning, Bloch and Nordsieck found that the appropriate expansion parameter is not $\frac{e^2}{\hbar c}$ but instead is given by $\frac{e^2}{\hbar c} \ln \frac{\hbar \omega_0}{m c^2}$. The higher-order perturbations may also describe processes involving larger numbers of emitted soft photons and results in a multiplicative exponential factor to the quasi-elastic scattering rate

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{Quasi-Elastic}} \approx \left(\frac{d\sigma}{d\Omega}\right)_{\text{Rutherford}} \exp \left[\frac{1}{2\pi} \left(\frac{e^2}{\hbar c}\right) B \ln \frac{2\hbar \omega_0}{m c^2} + \dots \right]$$
(1229)

Therefore, the scattering rate from soft photons vanishes in the limit $\omega_0 \to 0$. This occurs because perturbation theory causes the normalization of the starting approximate wave function to change, and hence the probabilities of the various processes are changed by including higher-order processes. In other words, since the probability of emitting an arbitrarily large number of soft-photons is finite, the probability of emitting either zero or any fixed number of soft photons must be zero. Bloch and Nordsieck's calculation was restricted to the case of emission of sufficiently low-energy photons. Pauli and Fierz[86] also considered Brehmstrahlung in a non-relativistic approximation. Pauli and Fierz showed that the infra-red divergences, discussed above, cancel. Pauli and Fierz went on to examine the remaining ultra-violet divergences, and showed that portions of the ultra-violet infinities that were found in the calculations of the scattering processes could be associated with mass renormalization. Using a relativistic theory Ito, Koba and Tomonaga[87] showed that the remaining infinities could

[86]W. Pauli and M. Fierz, Nuovo Cimento, **15**, 167 (1938).
[87]D. Ito, Z. Koba and S-I. Tomonaga, Prog. Theor. Phys. (Kyoto), **3**, 276 (1948).

be absorbed into a renormalization of the electron charge. Similar conclusions were arrived at by Lewis[88] and by Epstein[89]. Dyson[90] showed that all infinities that appear in Quantum Electrodynamics could be cured by renormalization to arbitrarily high-orders in perturbation theory.

12   The Dirac Equation

In 1928, Dirac searched for a relativistically invariant form of the one-particle Schrödinger equation for electrons

$$ i \hbar \frac{\partial}{\partial t} \psi = \hat{H} \psi \tag{1230} $$

Since this equation is only first-order in time, then the solution is uniquely specified by the initial condition for ψ. It is essential to only require an evolution equation which is first-order in time. Dirac[91] searched for a set of coupled first-order (in time) equations for a multi-component wave function ψ

$$ \psi = \begin{pmatrix} \psi^{(0)} \\ \psi^{(1)} \\ \vdots \\ \psi^{(N-1)} \end{pmatrix} \tag{1231} $$

The wave function was assumed to satisfy an equation of the form

$$ \left[i \frac{\hbar}{c} \frac{\partial}{\partial t} - \underline{\alpha} \cdot \hat{\underline{p}} \right] \psi = \beta \, m \, c \, \psi $$
$$ \left[i \frac{\hbar}{c} \frac{\partial}{\partial t} + i \hbar \, \underline{\alpha} \cdot \underline{\nabla} \right] \psi = \beta \, m \, c \, \psi \tag{1232} $$

The equations have to be of this form since, if the equation is a first-order partial differential equation in time then it must also only involve the first-order partial derivatives with respect to the spatial components for the resulting equation to be relativistically covariant. The wave function ψ is a N-component (column) wave function and the three as yet unknown components of $\underline{\alpha}$ and β are three $N \times N$ matrices. Since the Hamiltonian is the generator of time translations, then $\hat{H}$ should be equivalent to $i\hbar \frac{\partial}{\partial t}$. Hence, as the Hamiltonian operator $\hat{H}$ must be Hermitean, then the operators $\underline{\alpha}$ and β must be Hermitean matrices.

---

[88] H. W. Lewis, Phys. Rev. **73**, 173 (1948).

[89] Saul T. Epstein, Phys. Rev. **73**, 177 (1948).

[90] F. J. Dyson, Phys. Rev. **75**, 486 (1949).

[91] P. A. M. Dirac, Proc. Roy. Soc. **A 117**, 610 (1928).
P. A. M. Dirac, Proc. Roy. Soc. **A 118**, 351 (1928).

This set of equations is required to yield the dispersion relation for a relativistic particle

$$\left(\frac{E}{c} \right)^2 - \underline{p}^2 = m^2 \, c^2 \tag{1233}$$

which, following the ordinary rules of quantization, leads to the Klein-Gordon equation

$$\left[- \frac{\hbar^2}{c^2} \frac{\partial^2}{\partial t^2} + \hbar^2 \, \nabla^2 \right] \psi = m^2 \, c^2 \, \psi \tag{1234}$$

(which is a second-order partial differential equation in time). The requirement that the Dirac equation is compatible with the Klein-Gordon equation imposes conditions on the form of the matrices. On writing the Dirac equation as

$$i \, \frac{\hbar}{c} \, \frac{\partial \psi}{\partial t} = \left(\beta \, m \, c - i \, \hbar \, \underline{\alpha} \cdot \underline{\nabla} \right) \psi \tag{1235}$$

and iterating, one has

$$- \left(\frac{\hbar}{c} \right)^2 \frac{\partial^2 \psi}{\partial t^2} = \left(\beta \, m \, c - i \, \hbar \, \underline{\alpha} \cdot \underline{\nabla} \right)^2 \psi$$

$$= \left(\beta^2 \, m^2 \, c^2 - i \, \hbar \, m \, c \, (\, \beta \, \underline{\alpha} + \underline{\alpha} \, \beta \,) \cdot \underline{\nabla} \right.$$

$$\left. - \hbar^2 \, (\, \underline{\alpha} \cdot \underline{\nabla} \,)^2 \right) \psi \tag{1236}$$

When expressed in terms of individual matrices $\alpha^{(j)}$, the above equation becomes

$$- \left(\frac{\hbar}{c} \right)^2 \frac{\partial^2 \psi}{\partial t^2} = \left(\beta^2 \, m^2 \, c^2 - i \, \hbar \, m \, c \, \sum_j \, (\, \beta \, \alpha^{(j)} + \alpha^{(j)} \, \beta \,) \, \nabla_j \right.$$

$$\left. - \frac{\hbar^2}{2} \, \sum_{i,j} \, (\, \alpha^{(i)} \, \alpha^{(j)} + \alpha^{(j)} \, \alpha^{(i)} \,) \, \nabla_i \, \nabla_j \right) \psi \tag{1237}$$

since the derivatives commute. If the above equation is to be equivalent to the Klein-Gordon equation, then the coefficients of the various derivatives must be identical for both equations. Therefore, it is required that the constant terms are equal

$$\beta^2 = \hat{I} \tag{1238}$$

It is also required that the first-order derivative terms vanish and that the second-order derivative terms should be equal, hence the matrices must satisfy the anti-commutation relations

$$\alpha^{(i)} \, \beta + \beta \, \alpha^{(i)} = 0$$
$$\alpha^{(i)} \, \alpha^{(j)} + \alpha^{(j)} \, \alpha^{(i)} = 2 \, \delta^{i,j} \, \hat{I} \tag{1239}$$

On imposing the above conditions, Dirac's form of the relativistic Schrödinger equation is compatible with the Klein-Gordon equation.

From eqn(1238), one concludes that if the Hermitean matrices are brought to diagonal form then the diagonal elements are given by ± 1. The possible dimensions N of the matrix can be determined by considering the anti-commutation relations. On taking the determinant of eqn(1239), one finds

$$
\begin{aligned}
\det \alpha^{(i)} \ \det \beta &= (-1)^N \ \det \beta \ \det \alpha^{(i)} \\
\det \alpha^{(i)} \ \det \alpha^{(j)} &= (-1)^N \ \det \alpha^{(j)} \ \det \alpha^{(i)}
\end{aligned}
\tag{1240}
$$

Hence, on cancelling the common factors of determinants, one finds

$$
(-1)^N = 1 \tag{1241}
$$

so N must be even. Furthermore, the matrices must be traceless. This can be seen by considering

$$
\alpha^{(i)} \ \alpha^{(j)} = -\alpha^{(j)} \ \alpha^{(i)} \tag{1242}
$$

which on multiplying by $\alpha^{(i)}$, yields the relation

$$
\begin{aligned}
\alpha^{(j)} &= -\alpha^{(i)} \ \alpha^{(j)} \ \alpha^{(i)} \\
\alpha^{(j)} &= -(\alpha^{(i)})^{-1} \ \alpha^{(j)} \ \alpha^{(i)}
\end{aligned}
\tag{1243}
$$

since $\alpha^{(i)}$ is its own inverse. Apart from the negative sign, the form of the left-hand side is of the form of an equivalence transformation. By using cyclic invariance, it can be shown that the trace of a matrix is invariant under equivalence transformations. Therefore, one has

$$
\text{Trace } \alpha^{(i)} = -\text{Trace } \alpha^{(i)} \tag{1244}
$$

or

$$
\text{Trace } \alpha^{(i)} = 0 \tag{1245}
$$

which proves that the matrices are traceless.

Since the Dirac matrices satisfy

$$
\begin{aligned}
\beta^2 &= \hat{I} \\
(\alpha^{(i)})^2 &= \hat{I}
\end{aligned}
\tag{1246}
$$

then their eigenvalues must all be ± 1, as can be seen by operating on the eigenvalue equation

$$
\beta \ \phi_\beta = \lambda_\beta \ \phi_\beta \tag{1247}
$$

with β. This process yields

$$
\begin{aligned}
\beta^2 \ \phi_\beta &= \lambda_\beta \ \beta \ \phi_\beta \\
&= \lambda_\beta^2 \ \phi_\beta
\end{aligned}
\tag{1248}
$$

which with $\beta^2 = \hat{I}$, requires that the eigenvalues must satisfy the equation

$$\lambda_\beta^2 = 1 \tag{1249}$$

This and the condition that the matrices are traceless implies that the set of eigenvalues of each matrix are composed of equal numbers of $+1$ and -1, and it also confirms the conclusion that dimension N of the matrices must be even. The smallest value of the dimension for which there is a representation of the matrices is $N = 4$. The smallest even value of N, $N = 2$ cannot be used since one can only construct three linearly independent anti-commuting 2×2 matrices[92]. These three matrices are the Pauli spin matrices $\sigma^{(j)}$. Hence, Dirac constructed the relativistic theory with $N = 4$.

It is useful to find a representation in which the mass term is diagonal, since this represents the largest energy which occurs in the non-relativistic limit. When diagonalized, the β matrix has two eigenvalues of $+1$ and two eigenvalues of -1 and so β can be expressed in 2×2 block-diagonal form. We shall express the 4×4 matrices in the form of 2×2 block matrices. In this case, one can represent the matrix in the block-diagonal form

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{1250}$$

If the three matrices $\alpha^{(i)}$ are to anti-commute with β and be Hermitean, they must have the off-diagonal form

$$\alpha^{(i)} = \begin{pmatrix} 0 & A^{(i)} \\ A^{(i)\dagger} & 0 \end{pmatrix} \tag{1251}$$

where $A^{(i)}$ is an arbitrary 2×2 matrix. We shall choose all three $A^{(i)}$ matrices to be Hermitean. Since the three $\alpha^{(i)}$ matrices must anti-commute with each other, the $A^{(i)}$ must also anti-commute with each other. Since the three Pauli matrices are mutually anti-commuting, one can set

$$\alpha^{(i)} = \begin{pmatrix} 0 & \sigma^{(i)} \\ \sigma^{(i)} & 0 \end{pmatrix} \tag{1252}$$

where the $\sigma^{(i)}$ and I are, respectively, the 2×2 Pauli matrices and the 2×2 unit matrix. The Pauli matrices are given by

$$\sigma^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{1253}$$

---

[92]In $d + 1$ space-time dimensions, one can form 2^{d+1} matrices from products of the set of $d+1$ linearly independent (anti-commuting) Dirac-matrices. We shall assume that the product matrices are linearly independent. Since the number of linearly independent $N \times N$ matrices is N^2, the minimum dimension N which will yield a representation of the Dirac-matrices is $N = 2^{\frac{d+1}{2}}$.

$$\sigma^{(2)} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{1254}$$

and

$$\sigma^{(3)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1255}$$

The matrix $\alpha^{(0)}$ is defined as the 4×4 identity matrix

$$\alpha^{(0)} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{1256}$$

This set of matrices form a representation of the Dirac matrices. This can be seen by directly showing that they satisfy the appropriate relations. Many different representations of the Dirac matrices can be found, but they are all related by equivalence transformations and the physical results are independent of which choice is made.

Exercise:

By direct matrix multiplication, show that the above matrices satisfy the relations

$$(\alpha^{(j)})^2 = \beta^2 = \hat{I} \tag{1257}$$

and the anti-commutation relations

$$\begin{aligned}
\alpha^{(i)} \, \beta + \beta \, \alpha^{(i)} &= 0 \\
\alpha^{(i)} \, \alpha^{(j)} + \alpha^{(j)} \, \alpha^{(i)} &= 2 \, \delta^{i,j} \, \hat{I}
\end{aligned} \tag{1258}$$

and so form a representation of the Dirac matrices.

12.1 Conservation of Probability

One can find a conservation law for Dirac's equation

$$i \, \hbar \, \frac{\partial \psi}{\partial t} = \left(- i \, \hbar \, c \, \underline{\alpha} \, . \, \underline{\nabla} + \beta \, m \, c^2 \right) \psi \tag{1259}$$

On pre-multiplying the Dirac equation by $\psi^\dagger$, which is the Hermitean conjugate of the spinor wave function and is defined as the row vector formed by the complex conjugate of the components

$$\psi^\dagger = \left(\psi^{(0)*} , \psi^{(1)*} , \psi^{(2)*} , \psi^{(3)*} \right) , \tag{1260}$$

one obtains

$$i \, \hbar \, \psi^\dagger \, \frac{\partial \psi}{\partial t} = \left(- i \, \hbar \, c \, \psi^\dagger \, \underline{\alpha} \, . \, \underline{\nabla} \, \psi + \psi^\dagger \, \beta \, \psi \, m \, c^2 \right) \tag{1261}$$

The Hermitean conjugate of the Dirac equation is given by

$$- i \hbar \frac{\partial \psi^\dagger}{\partial t} = \left(+ i \hbar c \, \underline{\nabla} \cdot \psi^\dagger \, \underline{\alpha}^\dagger + \psi^\dagger \, \beta^\dagger \, m \, c^2 \right) \qquad (1262)$$

Since α and β are Hermitean matrices, the Hermitean conjugate equation simplifies to

$$- i \hbar \frac{\partial \psi^\dagger}{\partial t} = \left(+ i \hbar c \, \underline{\nabla} \cdot \psi^\dagger \, \underline{\alpha} + \psi^\dagger \, \beta \, m \, c^2 \right) \qquad (1263)$$

Post-multiplying the Hermitean conjugate equation by the column-vector ψ, yields

$$- i \hbar \frac{\partial \psi^\dagger}{\partial t} \, \psi = \left(+ i \hbar c \, \underline{\nabla} \cdot \psi^\dagger \, \underline{\alpha} \, \psi + \psi^\dagger \, \beta \, \psi \, m \, c^2 \right) \qquad (1264)$$

On subtracting eqn(1264) from the eqn(1261) and combining terms, one obtains

$$i \hbar \frac{\partial}{\partial t} \left(\psi^\dagger \, \psi \right) = - i \hbar c \, \underline{\nabla} \cdot \left(\psi^\dagger \, \underline{\alpha} \, \psi \right) \qquad (1265)$$

The above equation has the form of a continuity equation

$$\frac{\partial \rho}{\partial t} + \underline{\nabla} \cdot \underline{j} = 0 \qquad (1266)$$

in which the probability density is given by

$$\rho = \psi^\dagger \, \psi \qquad (1267)$$

Using the rules of matrix multiplication the probability density is a real scalar quantity, which is given by the sum of squares

$$\rho = \mid \psi^{(0)} \mid^2 + \mid \psi^{(1)} \mid^2 + \mid \psi^{(2)} \mid^2 + \mid \psi^{(3)} \mid^2 \qquad (1268)$$

and so it is positive definite. Hence, unlike the Klein-Gordon equation, the Dirac equation does not lead to negative probability densities. The probability current density $\underline{j}$ is given by

$$\underline{j} = c \, \psi^\dagger \, \underline{\alpha} \, \psi \qquad (1269)$$

In this case, the total probability

$$Q = \int d^3\underline{x} \, \psi^\dagger \, \psi = \int d^3\underline{x} \, \rho \qquad (1270)$$

is conserved, since

$$\begin{aligned} \frac{dQ}{dt} &= \int d^3\underline{x} \, \frac{\partial \rho}{\partial t} \\ &= - \int d^3\underline{x} \, \underline{\nabla} \cdot \underline{j} \\ &= - \int d^2\underline{S} \cdot \underline{j} \end{aligned} \qquad (1271)$$

where Gauss's theorem has been used to represent the volume integral as surface integral. For a sufficiently large volume, the current at the boundary vanishes, hence the total probability is conserved

$$\frac{dQ}{dt} = 0 \tag{1272}$$

12.2   Covariant Form of the Dirac Equation

In the absence of an electromagnetic field, the Dirac equation can be expressed in either of the two forms

$$\alpha^\mu \, \hat{p}_\mu \, \psi = \beta \, m \, c \, \psi$$
$$i \, \hbar \, \alpha^\mu \, \partial_\mu \, \psi = \beta \, m \, c \, \psi \tag{1273}$$

where it has been recalled that

$$\alpha^{(0)} = \hat{I} \tag{1274}$$

and the covariant momentum operator is given by

$$\hat{p}_\mu = i \, \hbar \, \left(\frac{\partial}{\partial x^\mu} \right) \tag{1275}$$

Or equivalently, after multiplying the Dirac equation by β and then introducing the four γ matrices via

$$\gamma^\mu = \beta \, \alpha^\mu \tag{1276}$$

one finds that the Dirac equation appears in the alternate forms

$$\gamma^\mu \, \hat{p}_\mu \, \psi = m \, c \, \psi$$
$$i \, \hbar \, \gamma^\mu \, \partial_\mu \, \psi = m \, c \, \psi \tag{1277}$$

The four gamma matrices satisfy the anti-commutation relations

$$\gamma^\mu \, \gamma^\nu + \gamma^\nu \, \gamma^\mu = 2 \, g^{\mu,\nu} \, \hat{I} \tag{1278}$$

where $\hat{I}$ is the 4×4 identity matrix, and $g^{\mu,\nu}$ is the Minkowski metric. The gamma matrices labelled by the spatial indices are Unitary and anti-Hermitean, as shall be proved below.

It is easy to show that the matrix with the temporal index (0) is unitary and Hermitean

$$(\gamma^{(0)})^{-1} = \gamma^{(0)}$$
$$(\gamma^{(0)})^\dagger = \gamma^{(0)} \tag{1279}$$

since β is its own inverse and β is Hermitean.

The gamma matrices with spatial indices are anti-Hermitean as

$$
\begin{aligned}
(\, \gamma^{(i)} \,)^\dagger &= (\, \beta \, \alpha^{(i)} \,)^\dagger \\
&= (\, (\, \alpha^{(i)} \,)^\dagger \, \beta^\dagger \,) \\
&= (\, \alpha^{(i)} \, \beta \,) \\
&= (\, - \, \beta \, \alpha^{(i)} \,) \\
&= - \, \gamma^{(i)}
\end{aligned}
\tag{1280}
$$

since $\alpha^{(i)}$ and β are Hermitean and, in the fourth line the operators have been anti-commuted. Now, the gamma matrices with spatial indices can be shown to be unitary since

$$
\begin{aligned}
\gamma^{(i)} \, \gamma^{(i)} &= \beta \, \alpha^{(i)} \, \beta \, \alpha^{(i)} \\
&= - \, \beta \, \beta \, \alpha^{(i)} \, \alpha^{(i)} \\
&= - \, \hat{I}
\end{aligned}
\tag{1281}
$$

where, in obtaining the second line, the anti-commutation properties of $\alpha^{(i)}$ and β have been used, and the property

$$
(\, \alpha^{(i)} \,)^2 \; = \; \beta^2 \; = \; \hat{I}
\tag{1282}
$$

was used to obtain the last line. Since it has already been demonstrated that the spatial matrices are anti-Hermitean

$$
(\, \gamma^{(i)} \,)^\dagger \; = \; - \, \gamma^{(i)}
\tag{1283}
$$

then it follows that $\gamma^{(i)}$ is unitary as

$$
(\, \gamma^{(i)} \,)^\dagger \, \gamma^{(i)} \; = \; \hat{I}
\tag{1284}
$$

which completes the proof.

The continuity equation can also be expressed in a covariant form. The covariant Dirac adjoint of ψ is defined as $\overline{\psi}^\dagger$ where

$$
\overline{\psi}^\dagger \; = \; \psi^\dagger \, \gamma^{(0)}
\tag{1285}
$$

Hence, since

$$
(\, \gamma^{(0)} \,)^2 \; = \; \hat{I}
\tag{1286}
$$

the Hermitean conjugate wave function $\psi^\dagger$ can be expressed in terms of the adjoint spinor $\overline{\psi}^\dagger$ via

$$
\psi^\dagger \; = \; \overline{\psi}^\dagger \, \gamma^{(0)}
\tag{1287}
$$

The continuity equation has the Lorentz covariant form

$$
\frac{\partial j^\mu}{\partial x^\mu} \; = \; 0
\tag{1288}
$$

224

where the four-vector conserved probability current j^μ is given by

$$j^\mu = c\,\psi^\dagger\,\alpha^\mu\,\psi \tag{1289}$$

By using the definition of the Dirac adjoint, the current density can be re-expressed as the four quantities

$$
\begin{aligned}
j^{(0)} &= c\,\overline{\psi}^\dagger\,\gamma^{(0)}\,\psi \\
j^{(i)} &= c\,\overline{\psi}^\dagger\,\gamma^{(i)}\,\psi
\end{aligned}
\tag{1290}
$$

that, respectively, represent c times the probability density and the $j^{(i)}$ are the contravariant components of the probability current density.

12.3 The Field Free Solution

In the absence of fields, the Dirac equation can be solved exactly by assuming a solution in the form of plane-waves. This is because the momentum operator $\hat{p}$ commutes with the Hamiltonian $\hat{H}$ since in the absence of fields there is no explicit dependence on position. The solution can be expressed as a momentum eigenstate in the form

$$
\psi = \begin{pmatrix} u^{(0)} \\ u^{(1)} \\ u^{(2)} \\ u^{(3)} \end{pmatrix} \exp\left[-\,i\,k_\mu\,x^\mu \right] \tag{1291}
$$

where the functions $u^\mu(k)$ are to be determined. On substituting this form in the Dirac equation, it becomes an algebraic equation of the form

$$
\left(k^{(0)}\,\hat{I} - \underline{k}\cdot\underline{\alpha} - (\frac{m\,c}{\hbar})\,\beta \right)\psi = 0 \tag{1292}
$$

where $\underline{k}$ is a three-vector with components given by the contra-variant spatial components of k^μ. In order to write this equation in two by two block-diagonal form, the four-component spinor ψ can be written in terms of two two-components spinors

$$
\psi = \begin{pmatrix} \phi^A \\ \phi^B \end{pmatrix} \tag{1293}
$$

where the two two-component spinors are given by

$$
\begin{aligned}
\phi^A &= \begin{pmatrix} u^{(0)} \\ u^{(1)} \end{pmatrix} \\
\phi^B &= \begin{pmatrix} u^{(2)} \\ u^{(3)} \end{pmatrix}
\end{aligned}
\tag{1294}
$$

225

Hence, the Dirac equation can be expressed as the block-diagonal matrix equation

$$
\begin{pmatrix} \left(-k^{(0)} + \frac{m\,c}{\hbar} \right) I & \underline{k} \cdot \underline{\sigma} \\ \underline{k} \cdot \underline{\sigma} & \left(-k^{(0)} - \frac{m\,c}{\hbar} \right) I \end{pmatrix} \begin{pmatrix} \phi^A \\ \phi^B \end{pmatrix} = 0 \qquad (1295)
$$

where the three-vector scalar product involves the contra-variant components of the momentum $k^{(i)}$ with the Pauli spin matrices $\sigma^{(i)}$. The above equation is an eigenvalue equation for $k^{(0)}$. The eigenvalues are given by the solution of the secular equation

$$
\begin{vmatrix} \left(-k^{(0)} + \frac{m\,c}{\hbar} \right) I & \underline{k} \cdot \underline{\sigma} \\ \underline{k} \cdot \underline{\sigma} & \left(-k^{(0)} - \frac{m\,c}{\hbar} \right) I \end{vmatrix} = 0 \qquad (1296)
$$

which can be written as

$$
\left(k^{(0)\,2} - \left(\frac{m\,c}{\hbar} \right)^2 \right) = \left(\underline{\sigma} \cdot \underline{k} \right)^2 \qquad (1297)
$$

Using Pauli's identity

$$
\left(\underline{\sigma} \cdot \underline{A} \right) \left(\underline{\sigma} \cdot \underline{B} \right) = \left(\underline{A} \cdot \underline{B} \right) I + i\,\underline{\sigma} \cdot \left(\underline{A} \wedge \underline{B} \right) \qquad (1298)
$$

one finds the energy eigenvalues are given by the doubly-degenerate dispersion relations

$$
k^{(0)} = \pm \sqrt{ \left(\frac{m\,c}{\hbar} \right)^2 + \underline{k}^2 } \qquad (1299)
$$

Thus, the field free relativistic electron can have positive and negative-energy eigenvalues given by

$$
E = \pm \sqrt{ m^2\,c^4 + \underline{p}^2\,c^2 } \qquad (1300)
$$

Since the solutions are degenerate, solutions can be found that are simultaneous eigenvalues of the Hamiltonian $\hat{H}$ given by

$$
\hat{H} = \begin{pmatrix} m\,c^2\,I & -\,i\,\hbar\,c\,\underline{\sigma} \cdot \underline{\nabla} \\ -\,i\,\hbar\,c\,\underline{\sigma} \cdot \underline{\nabla} & -\,m\,c^2\,I \end{pmatrix} \qquad (1301)
$$

and another operator that commutes with $\hat{H}$. It is convenient to choose the second operator to be the helicity operator.

The helicity operator $\hat{\Sigma}$ corresponds to the projection of the electron's spin along the direction of momentum. The (un-normalized) helicity operator is

$$\Sigma_k = +1 \qquad\qquad \Sigma_k = -1$$
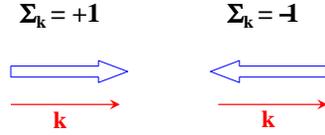
Figure 44: A cartoon depicting the two helicity states of a spin one-half particle.

given by

$$\hat{\Sigma} \;=\; -\,i\,\hbar\,\left(\begin{array}{cc} \underline{\sigma}\cdot\underline{\nabla} & 0 \\ 0 & \underline{\sigma}\cdot\underline{\nabla} \end{array}\right) \tag{1302}$$

This is the appropriate relativistic generalization of spin valid only for free particles[93], as the helicity is a conserved quantity since

$$[\,\hat{H}\,,\,\hat{\Sigma}\,]\;=\;0 \tag{1303}$$

In the absence of electromagnetic fields, the Hamiltonian is evaluated as

$$\hat{H}(\underline{k}) \;=\; \left(\begin{array}{cc} m\,c^2\,I & \hbar\,c\,\underline{\sigma}\cdot\underline{k} \\ \hbar\,c\,\underline{\sigma}\cdot\underline{k} & -\,m\,c^2\,I \end{array}\right) \tag{1304}$$

Likewise, for the source free case, the properly normalized Helicity operator is found as

$$\Lambda(\underline{k}) \;=\; \left(\begin{array}{cc} \underline{\sigma}\cdot\hat{k} & 0 \\ 0 & \underline{\sigma}\cdot\hat{k} \end{array}\right) \tag{1305}$$

which has eigenvalues of ± 1.

The axis of quantization of σ will be chosen to be along the direction of propagation $\hat{k}$. In this case, the helicity operator becomes

$$\Lambda(\underline{k}) \;=\; \left(\begin{array}{cc} \sigma^{(3)} & 0 \\ 0 & \sigma^{(3)} \end{array}\right) \tag{1306}$$

and the eigenstates of helicity with eigenvalue $+1$ are composed of a linear superposition of the spin-up eigenstates. We shall represent the two-component spinors ϕ_+^A via

$$\begin{aligned} \phi_+^A \;&=\; u^{(0)}\,\chi_+ \\ &=\; u^{(0)}\,\left(\begin{array}{c} 1 \\ 0 \end{array}\right) \end{aligned} \tag{1307}$$

---

[93]Helicity is not conserved for spherically symmetric potentials. However, if only a time-independent vector potential is present, the generalized quantity

$$\hat{\Sigma} \;=\; \underline{\sigma}\cdot(\,\hat{\underline{p}} - \frac{q}{c}\,\underline{A}\,)$$

is conserved. This conservation law implies that the spin will always retain its alignment with the velocity.

227

and ϕ^B_+ as

$$
\begin{aligned}
\phi^B_+ &= u^{(2)} \chi_+ \\
&= u^{(2)} \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\end{aligned}
\tag{1308}
$$

Therefore, one has

$$
\psi_+(x) = \begin{pmatrix} u^{(0)} \chi_+ \\ u^{(2)} \chi_+ \end{pmatrix} \exp\left[-i\, k_\mu\, x^\mu \right]
\tag{1309}
$$

Likewise, for the negative helicity states, ϕ^A_- can be represented via

$$
\begin{aligned}
\phi^A_- &= u^{(1)} \chi_- \\
&= u^{(1)} \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}
\tag{1310}
$$

and ϕ^B_+ as

$$
\begin{aligned}
\phi^B_- &= u^{(3)} \chi_- \\
&= u^{(3)} \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}
\tag{1311}
$$

Thus, the eigenstates with helicity -1 are the spin-down eigenstates

$$
\psi_-(x) = \begin{pmatrix} u^{(1)} \chi_- \\ u^{(3)} \chi_- \end{pmatrix} \exp\left[-i\, k_\mu\, x^\mu \right]
\tag{1312}
$$

Clearly, states with different helicities are orthogonal since

$$
\chi^\dagger_{\Lambda'}\, \chi_\Lambda = \delta_{\Lambda,\Lambda'}
\tag{1313}
$$

which is as it should be since they are eigenstates of a Hermitean operator.

On substituting the helicity eigenstates ψ_Λ into the Dirac equation for the free spin one-half particle

$$
i\,\hbar\, \frac{\partial}{\partial t}\, \psi_\Lambda = \hat{H}\, \psi_\Lambda
\tag{1314}
$$

one finds

$$
E \begin{pmatrix} \phi^A_\Lambda \\ \phi^B_\Lambda \end{pmatrix} = \begin{pmatrix} m\, c^2 & \sigma^{(3)}\, c\,\hbar\, k^{(3)} \\ \sigma^{(3)}\, c\,\hbar\, k^{(3)} & -\, m\, c^2 \end{pmatrix} \begin{pmatrix} \phi^A_\Lambda \\ \phi^B_\Lambda \end{pmatrix}
\tag{1315}
$$

Therefore, the complex amplitudes ϕ^A_Λ and ϕ^B_Λ are found to be related by

$$
\phi^B_\Lambda = \frac{\sigma^{(3)}\, c\,\hbar\, k^{(3)}}{E\, +\, m\, c^2}\, \phi^A_\Lambda
\tag{1316}
$$

228

This equation shows that the components ϕ_Λ^B are small for the positive-energy solutions, whereas the complementary expression

$$\phi_\Lambda^A \; = \; - \; \frac{\sigma^{(3)} \; c \; \hbar \; k^{(3)}}{m \; c^2 \; - \; E} \; \phi_\Lambda^B \tag{1317}$$

shows that ϕ_Λ^A is small for the negative-energy solutions. Hence, the two positive-energy and two negative-energy (un-normalized) solutions of the Dirac equation can be written as

$$\psi_+(x) \; = \; \mathcal{N}_e \; \left(\begin{array}{c} \chi_+ \\ \frac{c \; \hbar \; k^{(3)}}{E \; + \; m \; c^2} \; \chi_+ \end{array} \right) \; \exp\left[\; - \; i \; k_\mu \; x^\mu \; \right] \tag{1318}$$

for helicity $+1$ and

$$\psi_-(x) \; = \; \mathcal{N}_e \; \left(\begin{array}{c} \chi_- \\ - \; \frac{c \; \hbar \; k^{(3)}}{E \; + \; m \; c^2} \; \chi_- \end{array} \right) \; \exp\left[\; - \; i \; k_\mu \; x^\mu \; \right] \tag{1319}$$

for helicity -1. In this expression $\mathcal{N}_e$ is a normalization factor.

The normalization condition is

$$\int d^3\underline{r} \; \psi^\dagger \; \psi \; = \; 1 \tag{1320}$$

which determines the magnitude of the normalization constant through

$$\begin{aligned} 1 \; &= \; V \, \mathcal{N}_e{}^2 \; \left(\, 1 \; + \; \frac{c^2 \; \hbar^2 \; k^2}{(\; E \; + \; m \; c^2 \;)^2} \, \right) \\ &= \; V \, \mathcal{N}_e{}^2 \; \left(\frac{E^2 \; + \; 2 \; E \; m \; c^2 \; + \; m^2 \; c^4 \; + \; c^2 \; \hbar^2 \; k^2}{(\; E \; + \; m \; c^2 \;)^2} \right) \\ &= \; V \, \mathcal{N}_e{}^2 \; \left(\frac{2 \; E^2 \; + \; 2 \; E \; m \; c^2}{(\; E \; + \; m \; c^2 \;)^2} \right) \\ &= \; V \, \mathcal{N}_e{}^2 \; \left(\frac{2 \; E}{E \; + \; m \; c^2} \right) \end{aligned}$$

Hence, the normalization constant can be set as

$$\mathcal{N}_e \; = \; \sqrt{\frac{E \; + \; m \; c^2}{2 \; E \; V}} \tag{1321}$$

for positive E.

For states with negative energies,

$$E \; = \; - \; \sqrt{ \; m^2 \; c^4 \; + \; c^2 \; \hbar^2 \; k^2 } \tag{1322}$$

the lower components are the large components. In this case, it is more convenient to express the negative-energy solutions as

$$\psi_+(x) \;=\; \mathcal{N}_p \; \begin{pmatrix} -\, \frac{c\,\hbar\,k}{m\,c^2\,-\,E}\,\chi_+ \\[1em] \chi_+ \end{pmatrix} \; \exp\Big[\,-\,i\,k_\mu\,x^\mu\,\Big] \tag{1323}$$

for helicity +1 and

$$\psi_-(x) \;=\; \mathcal{N}_p \; \begin{pmatrix} \frac{c\,\hbar\,k}{m\,c^2\,-\,E}\,\chi_- \\[1em] \chi_- \end{pmatrix} \; \exp\Big[\,-\,i\,k_\mu\,x^\mu\,\Big] \tag{1324}$$

for helicity -1. Furthermore, in this expression the normalization constant has the form

$$\mathcal{N}_p \;=\; \sqrt{\frac{m\,c^2\,-\,E}{-\,2\,E\,V}} \tag{1325}$$

Hence, the positive and negative-energy solutions are symmetric under the interchange $E \;\to\; -\,E$, if $\Lambda \;\to\; -\,\Lambda$ and the upper and lower two-component spinors (ϕ^A, ϕ^B) are interchanged.

General Helicity Eigenstates

The Helicity operator for a particle with a momentum $\hbar\,\underline{k}$ is given by the Hermitean operator

$$\begin{aligned} \Lambda(k) \;&=\; \frac{1}{k} \begin{pmatrix} k^{(3)} & k^{(1)}\,-\,i\,k^{(2)} \\[1em] k^{(1)}\,+\,i\,k^{(2)} & -\,k^{(3)} \end{pmatrix} \\[1em] &=\; \begin{pmatrix} \cos\theta_k & \sin\theta_k\,\exp[-i\varphi_k] \\[1em] \sin\theta_k\,\exp[+i\varphi_k] & -\,\cos\theta_k \end{pmatrix} \end{aligned} \tag{1326}$$

which since

$$\Lambda(k)\,\Lambda(k) \;=\; I \tag{1327}$$

has eigenvalues Λ of ± 1. The helicity eigenstates[94] are given by the two-component spinors $\chi_{\Lambda\pm}$. The positive helicity state is given by

$$\begin{aligned} \chi_{\Lambda+} \;&=\; \frac{1}{\sqrt{2\,k\,(\,k\,-\,k^{(3)}\,)}} \begin{pmatrix} k^{(1)}\,-\,i\,k^{(2)} \\[1em] k\,-\,k^{(3)} \end{pmatrix} \\[1em] &=\; \exp[\,-\,i\,\tfrac{\varphi_k}{2}\,] \begin{pmatrix} \cos\tfrac{\theta_k}{2}\,\exp[-i\tfrac{\varphi_k}{2}] \\[1em] \sin\tfrac{\theta_k}{2}\,\exp[+i\tfrac{\varphi_k}{2}] \end{pmatrix} \end{aligned} \tag{1328}$$

[94] C. G. Darwin, Proc. Roy. Soc. **A 118**, 654 (1928).
C. G. Darwin, Proc. Roy. Soc. **A 120**, 631 (1928).

in which (k, θ_k, φ_k) are the polar coordinates of $\underline{k}$. The negative helicity eigenstate is given by the spinor $\chi_{\Lambda-}$

$$
\begin{aligned}
\chi_{\Lambda-} &= \frac{1}{\sqrt{2\,k\,(\,k\,-\,k^{(3)}\,)}} \begin{pmatrix} -\,k\,+\,k^{(3)} \\[2mm] k^{(1)}\,+\,i\,k^{(2)} \end{pmatrix} \\[3mm]
&= \exp[\,+\,i\,\frac{\varphi_k}{2}\,] \begin{pmatrix} -\,\sin\frac{\theta_k}{2}\,\exp[-i\frac{\varphi_k}{2}] \\[2mm] \cos\frac{\theta_k}{2}\,\exp[+i\frac{\varphi_k}{2}] \end{pmatrix}
\end{aligned}
\tag{1329}
$$

Therefore, the general helicity eigenstate plane-wave solutions of the Dirac equation can be written in terms of two two-component spinors as

$$
\psi_{\Lambda\pm}(x) \;=\; \mathcal{N}_e \begin{pmatrix} \chi_{\Lambda\pm} \\[2mm] \frac{c\,\hbar\,k\,\Lambda_{\pm}}{E\,+\,m\,c^2}\,\chi_{\Lambda\pm} \end{pmatrix} \exp\Big[\,-\,i\,k_\mu\,x^\mu\,\Big]
\tag{1330}
$$

In this expression $\mathcal{N}_e$ is a normalization factor

$$
\mathcal{N}_e \;=\; \sqrt{\frac{E\,+\,m\,c^2}{2\,E\,V}}
\tag{1331}
$$

These plane-wave solutions are useful in considerations of scattering processes.

12.4   Coupling to Fields

The Dirac equation describes relativistic spin one-half fermions, and their antiparticles. It describes all massive leptons such as the electron, muon and tao particle, and can be generalized to describe their interaction with the electromagnetic field, or its generalization the electro-weak field. In the limit $m \to 0$, the Dirac equation reduces to the Weyl equation[95] which describes neutrinos. The Dirac equation also describes massive quarks and the interaction can be generalized to quantum chromodynamics.

In the absence of interactions, the Dirac equation can be expressed in either of the two forms

$$
\begin{aligned}
\alpha^\mu\,\hat{p}_\mu\,\psi &= \beta\,m\,c\,\psi \\
i\,\hbar\,\alpha^\mu\,\partial_\mu\,\psi &= \beta\,m\,c\,\psi
\end{aligned}
\tag{1332}
$$

The interaction with electromagnetic field is introduced as follows. Using the minimal coupling approximation, where

$$
\hat{p}_\mu \;\to\; \hat{p}_\mu{}' \;=\; \hat{p}_\mu\,-\,\frac{q}{c}\,A_\mu
\tag{1333}
$$

---

[95] H. Weyl, Z. Physik, **56**, 330 (1929).

and q is the charge of the particle, the Dirac equation in the presence of an electromagnetic field becomes

$$\alpha^\mu \left(\hat{p}_\mu - \frac{q}{c} A_\mu \right) \psi = \beta \, m \, c \, \psi$$

$$i \, \hbar \, \alpha^\mu \left(\partial_\mu + i \, \frac{q}{\hbar \, c} A_\mu \right) \psi = \beta \, m \, c \, \psi \tag{1334}$$

This process has resulted in the inclusion of the interaction with the electromagnetic field in a gauge invariant, Lorentz covariant manner. The appearance of the gauge field together with the derivative results in local gauge invariance. Sometimes it is convenient to define a covariant derivative as the gauge-invariant combination

$$\mathcal{D}_\mu \equiv \partial_\mu + i \, \frac{q}{\hbar \, c} A_\mu \tag{1335}$$

The concept of the covariant derivative also appears in the context of other gauge field theories. Using this definition we can express the Dirac equation in the presence of an electromagnetic field in the compact covariant form

$$i \, \hbar \, \gamma^\mu \, \mathcal{D}_\mu \, \psi = m \, c \, \psi \tag{1336}$$

The presence of an electromagnetic field does not alter the form of the conserved four-vector current

$$j^\mu = c \, \overline{\psi}^\dagger \, \gamma^\mu \, \psi \tag{1337}$$

which is explicitly gauge invariant.

12.4.1 Mott Scattering

We shall consider the scattering of positive-energy electrons from a nucleus of charge Z. The initial electron beam has momentum $\hbar \, \underline{k}$ which is scattered by the target nucleus. The detector is positioned so that it detects all the scattered electrons with momentum $\hbar \, \underline{k}'$. The initial and final states of the positive-energy electron can be represented by the Dirac spinors of the form ψ_σ

$$\psi_{\underline{k},\sigma}(x) = \mathcal{N}_k \begin{pmatrix} \chi_\sigma \\ \frac{c \, \hbar \, \underline{k} \cdot \sigma}{E_k + m \, c^2} \chi_\sigma \end{pmatrix} \exp \left[- i \, k_\mu \, x^\mu \right] \tag{1338}$$

where the normalization constant is chosen as

$$\mathcal{N}_k = \sqrt{\frac{E_k + m \, c^2}{2 \, E_k \, V}} \tag{1339}$$

The interaction Hamiltonian with the electrostatic field of the nucleus is given by the diagonal matrix

$$\hat{H}_{Int} = - \frac{Z \, e^2}{r} \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{1340}$$

The flux of incident electrons is defined by

$$\mathcal{F} = \frac{|\,\underline{v}\,|}{V} \tag{1341}$$

where

$$\underline{v} = \left(\frac{\partial E}{\partial \underline{p}} \right) \tag{1342}$$

Therefore, the electron flux is given by

$$\mathcal{F} = \left(\frac{\hbar\ k\ c^2}{V\ E_k} \right) \tag{1343}$$

The elastic scattering cross-section in which the final state polarization is unmeasured is given by

$$\left(\frac{d\sigma}{d\Omega'} \right) = \frac{1}{(\,2\ \pi\,)^2} \left(\frac{E_k\ V^2}{\hbar^2\ k\ c^2} \right) \sum_{\sigma'} \int_0^\infty dk'\ k'^2\ |\ <\ \underline{k}'\sigma'\ |\ \hat{H}_{Int}\ |\ \underline{k},\sigma\ >\ |^2\ \delta(\ E_k - E_{k'}\ ) \tag{1344}$$

where the delta function ensures conservation of energy. Since the polarization of the final state electron is unmeasured, the spin σ' is summed over. The integration over k' can be performed, yielding

$$\left(\frac{d\sigma}{d\Omega'} \right) = \left(\frac{E\ V}{2\ \pi\ \hbar^2\ c^2} \right)^2 \sum_{\sigma'} |\ <\ \underline{k}'\sigma'\ |\ \hat{H}_{Int}\ |\ \underline{k},\sigma\ >\ |^2 \tag{1345}$$

where $\underline{k}$ and $\underline{k}'$ are restricted to be on the energy shell ($E = E_k = E_{k'}$). The matrix elements can then be evaluated as

$$< \underline{k}',\sigma'\ |\ \hat{H}_{Int}\ |\ \underline{k},\sigma\ > = -\left(\frac{4\ \pi\ Z\ e^2}{V\ |\ \underline{k}\ -\ \underline{k}'\ |^2} \right) \left(\frac{E\ +\ m\ c^2}{2\ E\ V} \right)$$
$$\times\ \chi^T_{\sigma'} \left(I\ +\ \frac{c^2\ \hbar^2\ (\ \underline{\sigma}\ .\ \underline{k}'\ )\ (\ \underline{\sigma}\ .\ \underline{k}\ )}{(\ E\ +\ m\ c^2\ )^2} \right) \chi_\sigma \tag{1346}$$

where the normalization constants have been combined, since energy is conserved. Likewise, the complex conjugate matrix elements are given by

$$< \underline{k},\sigma\ |\ \hat{H}_{Int}\ |\ \underline{k}',\sigma'\ > = -\left(\frac{4\ \pi\ Z\ e^2}{V\ |\ \underline{k}\ -\ \underline{k}'\ |^2} \right) \left(\frac{E\ +\ m\ c^2}{2\ E\ V} \right)$$
$$\times\ \chi^T_\sigma \left(I\ +\ \frac{c^2\ \hbar^2\ (\ \underline{\sigma}\ .\ \underline{k}\ )\ (\ \underline{\sigma}\ .\ \underline{k}'\ )}{(\ E\ +\ m\ c^2\ )^2} \right) \chi_{\sigma'} \tag{1347}$$

These expressions for the matrix elements are inserted into the scattering cross-section. Since the final state polarization is not detected, then σ' must be summed over. The trace over σ' is evaluated by using the completeness relation

$$\sum_{\sigma'} \chi_{\sigma'}\ \chi^T_{\sigma'} = I \tag{1348}$$

233

The resulting matrix elements involve the spin-dependent factor

$$\chi_\sigma^T \left(I + \frac{c^2 \, \hbar^2 \, (\, \underline{\sigma} \cdot \underline{k} \,) \, (\, \underline{\sigma} \cdot \underline{k}' \,)}{(\, E \, + \, m \, c^2 \,)^2} \right) \left(I + \frac{c^2 \, \hbar^2 \, (\, \underline{\sigma} \cdot \underline{k}' \,) \, (\, \underline{\sigma} \cdot \underline{k} \,)}{(\, E \, + \, m \, c^2 \,)^2} \right) \chi_\sigma$$

(1349)

The products of matrix elements shown above can be evaluated with the aid of the Pauli identity. The sum of the cross-terms can be evaluated directly using the Pauli identity. We note that since the vector product are antisymmetric in $\underline{k}$ and $\underline{k}'$, the sum of the vector product terms cancel. That is

$$\frac{c^2 \, \hbar^2}{(\, E \, + \, m \, c^2 \,)^2} \left[(\, \underline{\sigma} \cdot \underline{k} \,) \, (\, \underline{\sigma} \cdot \underline{k}' \,) \, + \, (\, \underline{\sigma} \cdot \underline{k}' \,) \, (\, \underline{\sigma} \cdot \underline{k} \,) \right]$$
$$= \frac{c^2 \, \hbar^2}{(\, E \, + \, m \, c^2 \,)^2} \, 2 \, (\, \underline{k} \cdot \underline{k}' \,) \, \hat{I}$$

(1350)

The remaining term is evaluated by using the Pauli identity for the inner two scalar products, and then re-using the identity for the outer two scalar products. Explicitly, this process yields

$$\frac{c^4 \, \hbar^4}{(\, E \, + \, m \, c^2 \,)^4} \left[(\, \underline{\sigma} \cdot \underline{k} \,) \, (\, \underline{\sigma} \cdot \underline{k}' \,) \, (\, \underline{\sigma} \cdot \underline{k}' \,) \, (\, \underline{\sigma} \cdot \underline{k} \,) \right]$$
$$= \frac{c^4 \, \hbar^4}{(\, E \, + \, m \, c^2 \,)^4} \, \underline{k}'^2 \, \underline{k}^2 \, \hat{I}$$

(1351)

Hence, the cross-section is given by

$$\left(\frac{d\sigma}{d\Omega'} \right) = \left(\frac{Z \, e^2}{\hbar^2 \, c^2 \, | \, \underline{k} \, - \, \underline{k}' \, |^2} \right)^2 \left[(\, E + m \, c^2 \,)^2 + 2 \, c^2 \, \hbar^2 \, \underline{k} \cdot \underline{k}' + \frac{c^4 \, \hbar^4 \, k^2 \, k'^2}{(\, E \, + \, m \, c^2 \,)^2} \right]$$

(1352)

It should be noted that the last two terms originated from the combined action of the Pauli spin operators and involved the lower two-component spinors. The last term can be simplified by using the elastic scattering condition $k = k'$ and then using the identity

$$c^4 \, \hbar^4 \, k^4 = (\, E^2 \, - \, m^2 \, c^4 \,)^2$$

(1353)

in the numerator. On canceling the factor of $(\, E + m \, c^2 \,)^2$ in the denominator of the last term with a similar factor in the numerator, the last term is recognized as being just $(\, E \, - \, m \, c^2 \,)^2$. Hence, on combining the first and last terms, one finds the result

$$\left(\frac{d\sigma}{d\Omega'} \right) = \left(\frac{Z \, e^2}{\hbar^2 \, c^2 \, | \, \underline{k} \, - \, \underline{k}' \, |^2} \right)^2 \left[2 \, (\, E^2 + m^2 \, c^4 \,) + 2 \, c^2 \, \hbar^2 \, \underline{k} \cdot \underline{k}' \right]$$

(1354)

The scattering angle θ' is introduced in the square parenthesis through

$$\underline{k} \cdot \underline{k}' = k^2 \, \cos \theta'$$

(1355)

and also in the denominator of the Coulomb interaction by

$$| \, \underline{k} \, - \, \underline{k}' \, |^2 \; = \; 4 \, k^2 \, \sin^2 \frac{\theta'}{2} \tag{1356}$$

Furthermore, the factor of $m^2 \, c^4$ in the square parenthesis can be replaced by

$$m^2 \, c^4 \; = \; E^2 \, - \, c^2 \, \hbar^2 \, k^2 \tag{1357}$$

so that the cross-section takes the form

$$\begin{aligned} \left(\frac{d\sigma}{d\Omega'} \right) \; &= \; \left(\frac{Z \, e^2}{2 \, \hbar^2 \, c^2 \, k^2 \, \sin^2 \frac{\theta'}{2}} \right)^2 \left[\, E^2 \; - \; c^2 \, \hbar^2 \, k^2 \, \sin^2 \frac{\theta'}{2} \, \right] \\ &= \; \left(\frac{2 \, Z \, e^2 \, E}{4 \, \hbar^2 \, c^2 \, k^2 \, \sin^2 \frac{\theta'}{2}} \right)^2 \left[\, 1 \; - \; \left(\frac{v}{c} \right)^2 \, \sin^2 \frac{\theta'}{2} \, \right] \end{aligned} \tag{1358}$$

where the expression for the magnitude of the velocity

$$v^2 \; = \; \left(\frac{c^2 \, \hbar \, k}{E} \right)^2 \tag{1359}$$

has been introduced. The above result is the Mott scattering cross-section[96], which describes the scattering of charged electrons. It differs from the Rutherford scattering cross-section due to the multiplicative factor of relativistic origin, which deviates from unity due to the electron's internal degree of freedom. The extra contribution to the scattering is interpreted in terms of scattering from the magnetic moment associated with the electron's spin interacting with the magnetic field of the nuclear charge that the electron experiences in its rest frame. It should be noted that even if the initial beam of electrons is un-polarized, the scattered beam will be partially spin-polarized (due to higher-order corrections).

12.4.2   Maxwell's Equations

Maxwell's equations can be written in the form of the Dirac equation. We introduce a four-component wave function ψ given by

$$\psi \; = \; \begin{pmatrix} 0 \\ B^{(1)} - i E^{(1)} \\ B^{(2)} - i E^{(2)} \\ B^{(3)} - i E^{(3)} \end{pmatrix} \tag{1360}$$

Maxwell's equations can be written in the form

$$i \; \alpha^\mu \, \partial_\mu \psi \; = \; - \; \frac{4 \, \pi}{c} \, j \tag{1361}$$

---

[96] N. F. Mott, Proc. Roy. Soc. **A 124**, 425 (1929).

where j is the contravariant form of the current four-vector

$$j = \begin{pmatrix} c\,\rho \\ j^{(1)} \\ j^{(2)} \\ j^{(3)} \end{pmatrix} \tag{1362}$$

We shall require that the matrices α^μ are Hermitean and that they satisfy the equation

$$(\alpha^\mu)^2 = \hat{I} \tag{1363}$$

On comparing with the form of Maxwell's equations[97], one finds that the Matrices are given by

$$\alpha^{(0)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1364}$$

$$\alpha^{(1)} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \tag{1365}$$

$$\alpha^{(2)} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & i \\ -1 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix} \tag{1366}$$

$$\alpha^{(3)} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{1367}$$

The matrices corresponding to the spatial indices are traceless and satisfy the anti-commutation relations

$$\alpha^{(i)} \, \alpha^{(j)} + \alpha^{(j)} \, \alpha^{(i)} = 2 \, \delta_{i,j} \tag{1368}$$

and

$$\alpha^{(i)} \, \alpha^{(j)} = i \sum_k \xi^{i,j,k} \, \alpha^{(k)} \tag{1369}$$

On pre-multiplying Maxwell's equations in the form

$$i \, \alpha^\mu \, \partial_\mu \psi = - \frac{4\,\pi}{c} \, j \tag{1370}$$

---

[97]Since the first element of ψ is zero, the first columns of the matrices are not determined directly from the comparison. The first rows are determined by demanding that the matrices are Hermitean.

with the operator

$$i \, \alpha^{\nu} \, \partial_{\nu} \tag{1371}$$

one obtains

$$- \, \alpha^{\nu} \, \alpha^{\mu} \, \partial_{\nu} \, \partial_{\mu} \psi \; = \; - \, i \, \frac{4 \, \pi}{c} \, \alpha^{\nu} \, \partial_{\nu} j \tag{1372}$$

Utilizing the anti-commutation of the spatial matrices, the left-hand side simplifies to

$$- \left[\, - \, \partial_{\mu} \, \partial^{\mu} \; + \; 2 \left(\, \alpha^{\nu} \, \partial_{\nu} \, \right) \frac{1}{c} \, \frac{\partial}{\partial t} \, \right] \psi \tag{1373}$$

On substituting the new form of Maxwell's equations in the second term, the expression reduces to

$$- \left[\, - \, \partial_{\mu} \, \partial^{\mu} \, \psi \; + \; i \, \frac{8 \, \pi}{c^2} \, \frac{\partial}{\partial t} \, j \, \right] \tag{1374}$$

Thus, the equation becomes

$$\partial_{\mu} \, \partial^{\mu} \psi \; = \; i \, \frac{4 \, \pi}{c} \left(\frac{2}{c} \, \frac{\partial}{\partial t} \; - \, \alpha^{\nu} \, \partial_{\nu} \, \right) j \tag{1375}$$

The zero-th component of the source term vanishes, due to conservation of charge.

12.4.3   The Gordon Decomposition

The interaction of the Dirac particle with the electromagnetic field is described by the interaction Hamiltonian which is described by the 4×4 matrix

$$\hat{H}_I \; = \; \left(\frac{q}{c} \right) c \, \gamma^{(0)} \, \gamma^{\mu} \, A_{\mu} \tag{1376}$$

The matrix interaction Hamiltonian operator yields an interaction Hamiltonian density $\hat{\mathcal{H}}_I$ given by

$$\begin{aligned}
\hat{\mathcal{H}}_I \; &= \; \left(\frac{q}{c} \right) c \, \overline{\psi}^{\dagger} \, \gamma^{\mu} \, \psi \, A_{\mu} \\
&= \; \left(\frac{q}{c} \right) j^{\mu} \, A_{\mu}
\end{aligned} \tag{1377}$$

where j^{μ} is the four-vector probability current density which satisfies the condition for conservation of probability. Due to the prominence of the current density operator in applications of the Dirac equation, since it naturally describes interactions with an electromagnetic field and the conservation laws, the physical content of the current densities shall be examined next.

In the presence of an electromagnetic field, the four-vector current density is given by the expression

$$j^{\nu} \; = \; c \, \overline{\psi}^{\dagger} \, \gamma^{\nu} \, \psi \tag{1378}$$

237

where

$$\overline{\psi}^{\dagger} \; = \; \psi^{\dagger} \; \gamma^{(0)} \tag{1379}$$

One can rewrite the current density by using the Dirac equation

$$i \, \hbar \, \gamma^{\mu} \left(\; \partial_{\mu} \; + \; i \, \frac{q}{\hbar \, c} \, A_{\mu} \; \right) \psi \; = \; m \, c \, \psi \tag{1380}$$

and the Hermitean conjugate equation

$$- \, i \, \hbar \left(\; \partial_{\mu} \; - \; i \, \frac{q}{\hbar \, c} \, A_{\mu} \; \right) \psi^{\dagger} \; \gamma^{\mu \dagger} \; = \; m \, c \, \psi^{\dagger} \tag{1381}$$

On symmetrizing the current density and then substituting the Dirac equation in one term and its Hermitean conjugate in the other term, one obtains

$$
\begin{aligned}
j^{\nu} \; &= \; \frac{c}{2} \left(\; \overline{\psi}^{\dagger} \; \gamma^{\nu} \; \psi \; + \; \overline{\psi}^{\dagger} \; \gamma^{\nu} \; \psi \; \right) \\
&= \; \frac{i \, \hbar}{2 \, m} \left(\; - \; (\, \partial_{\mu} \; - \; i \, \frac{q}{\hbar \, c} \, A_{\mu} \,) \psi^{\dagger} \; \gamma^{\mu \dagger} \; \gamma^{(0)} \; \gamma^{\nu} \; \psi \; + \; \psi^{\dagger} \; \gamma^{(0)} \; \gamma^{\nu} \; \gamma^{\mu} \; (\, \partial_{\mu} \; + \; i \, \frac{q}{\hbar \, c} \, A_{\mu} \,) \psi \; \right) \\
&= \; \frac{i \, \hbar}{2 \, m} \left(\; - \; (\, \partial_{\mu} \; - \; i \, \frac{q}{\hbar \, c} \, A_{\mu} \,) \overline{\psi}^{\dagger} \; \gamma^{(0)} \; \gamma^{\mu \dagger} \; \gamma^{(0)} \; \gamma^{\nu} \; \psi \; + \; \overline{\psi}^{\dagger} \; \gamma^{\nu} \; \gamma^{\mu} \; (\, \partial_{\mu} \; + \; i \, \frac{q}{\hbar \, c} \, A_{\mu} \,) \psi \; \right)
\end{aligned}
\tag{1382}
$$

where the partial derivatives only operate on the wave function immediately to the right of it. The identity

$$\gamma^{(0)} \; \gamma^{(0)} \; = \; \hat{I} \tag{1383}$$

has been used to express $\psi^{\dagger}$ in terms of $\overline{\psi}^{\dagger}$. However, since the γ matrices satisfy

$$\gamma^{(0)} \; \gamma^{\mu \dagger} \; \gamma^{(0)} \; = \; \gamma^{\mu} \tag{1384}$$

the current can be further simplified to yield

$$j^{\nu} \; = \; \frac{i \, \hbar}{2 \, m} \left(\; - \; (\, \partial_{\mu} \; - \; i \, \frac{q}{\hbar \, c} \, A_{\mu} \,) \overline{\psi}^{\dagger} \; \gamma^{\mu} \; \gamma^{\nu} \; \psi \; + \; \overline{\psi}^{\dagger} \; \gamma^{\nu} \; \gamma^{\mu} \; (\, \partial_{\mu} \; + \; i \, \frac{q}{\hbar \, c} \, A_{\mu} \,) \psi \; \right) \tag{1385}$$

where, once again, the partial derivative only operates on the wave function immediately to the right of it. Furthermore, if one sets

$$
\begin{aligned}
\frac{1}{2} \left(\; \gamma^{\mu} \; \gamma^{\nu} \; + \; \gamma^{\nu} \; \gamma^{\mu} \; \right) \; &= \; g^{\mu, \nu} \; \hat{I} \\
\frac{1}{2} \left(\; \gamma^{\mu} \; \gamma^{\nu} \; - \; \gamma^{\nu} \; \gamma^{\mu} \; \right) \; &= \; - \, i \, \sigma^{\mu, \nu}
\end{aligned}
\tag{1386}
$$

then the current density can be expressed as the sum of two contributions

$$
\begin{aligned}
j^\nu &= j_c^\nu + j_s^\nu \\
&= \frac{i\,\hbar}{2\,m} \left(-\, g^{\mu,\nu}\,(\,\partial_\mu \overline{\psi}^\dagger\, \psi \,-\, \overline{\psi}^\dagger\, \partial_\mu \psi\,) \,+\, 2\,i\,\frac{q}{\hbar\,c}\,g^{\mu,\nu}\,\overline{\psi}^\dagger\, A_\mu\,\psi \right) \\
&\quad -\, \frac{\hbar}{2\,m}\,\frac{\partial}{\partial x^\mu} \left(\overline{\psi}^\dagger\, \sigma^{\mu,\nu}\,\psi \right)
\end{aligned}
\tag{1387}
$$

where

$$
\begin{aligned}
j_c^\nu &= \frac{i\,\hbar}{2\,m} \left(-\, (\,\partial^\nu \overline{\psi}^\dagger\, \psi \,-\, \overline{\psi}^\dagger\, \partial^\nu \psi\,) \,+\, 2\,i\,\frac{q}{\hbar\,c}\,\overline{\psi}^\dagger\, A^\nu\,\psi \right) \\
j_s^\nu &= -\, \frac{\hbar}{2\,m}\,\frac{\partial}{\partial x^\mu} \left(\overline{\psi}^\dagger\, \sigma^{\mu,\nu}\,\psi \right)
\end{aligned}
\tag{1388}
$$

This is the Gordon decomposition[98] of the probability current density. A similar expression can be derived for the matrix elements of the interaction operator between states $\overline{\psi}_\beta^\dagger$ and ψ_α. As shall be shown, the first contribution in the Gordon decomposition is gauge invariant and dominates the current density in the non-relativistic limit. The second contribution involves the matrix $\sigma^{\mu,\nu}$ which is anti-symmetric in its indices and has the form of a spin contribution to the current density.

Let us examine the first term in the probability current density. If ψ represents an energy eigenstate, then $j_c^{(0)}$ is given by

$$
j_c^{(0)} = \left(\frac{E}{m\,c} \right) \overline{\psi}^\dagger\, \psi \,-\, \frac{q}{m\,c}\,\overline{\psi}^\dagger\, A^{(0)}\,\psi
\tag{1389}
$$

This contribution obviously yields the main contribution to (c times) the probability density

$$
j_c^{(0)} \approx c\,\overline{\psi}^\dagger\, \psi
\tag{1390}
$$

in the non-relativistic limit since the rest mass energy dominates the energy $E \sim m\,c^2$. The spatial components of $j_c^{(i)}$ are given by

$$
\underline{j}_c = \frac{i\,\hbar}{2\,m} \left[(\,\nabla\, \overline{\psi}^\dagger\,)\,\psi \,-\, \overline{\psi}^\dagger\,(\,\nabla\, \psi\,) \right] \,-\, \frac{q}{m\,c}\,\overline{\psi}^\dagger\underline{A}\,\psi
\tag{1391}
$$

where the derivatives have been expressed as derivatives w.r.t. the contravariant components $x^{(i)}$ of the position vector. This expression coincides with the full non-relativistic expression for the current density $j^{(i)}$.

We now examine the second term j_s^μ in the Gordon decomposition. For future reference, the anti-symmetrized products of the Dirac matrices $\sigma^{\mu,\nu}$ will

---

[98]W. Gordon, Zeit. für Physik, **50**, 630 (1928).

be expressed in 2×2 block diagonal form. Therefore, since

$$
\begin{aligned}
\gamma^{(0)} &= \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \\
\gamma^{(i)} &= \begin{pmatrix} 0 & \sigma^{(i)} \\ -\sigma^{(i)} & 0 \end{pmatrix}
\end{aligned}
\tag{1392}
$$

and

$$
\sigma^{\mu,\nu} = \frac{i}{2} \left(\gamma^\mu \, \gamma^\nu - \gamma^\nu \, \gamma^\mu \right)
\tag{1393}
$$

the matrices are found as

$$
\sigma^{0,j} = i \begin{pmatrix} 0 & \sigma^{(j)} \\ \sigma^{(j)} & 0 \end{pmatrix}
\tag{1394}
$$

and

$$
\sigma^{i,j} = \sum_k \xi^{i,j,k} \begin{pmatrix} \sigma^{(k)} & 0 \\ 0 & \sigma^{(k)} \end{pmatrix}
\tag{1395}
$$

The two by two block diagonal matrix of Pauli spin matrices will be denoted by $\hat{\underline{\sigma}}$. For an energy eigenstate, the time component of $j_s^{(0)}$ is identically zero. Hence, the spatial components of $j_s^{(i)}$ are given by

$$
\underline{j}_s = - \frac{\hbar}{2\,m} \, \underline{\nabla} \wedge (\, \overline{\psi}^\dagger \, \hat{\underline{\sigma}} \, \psi \,)
\tag{1396}
$$

where $\hat{\underline{\sigma}}$ is the 2×2 block-diagonal Pauli spin matrix

$$
\hat{\underline{\sigma}} = \begin{pmatrix} \underline{\sigma} & 0 \\ 0 & \underline{\sigma} \end{pmatrix}
\tag{1397}
$$

The additional term in the current density clearly involves the Pauli spin-matrices. To elucidate its meaning, its contribution to the energy shall be examined. On substituting this term in the interaction Hamiltonian density, one finds a contribution

$$
\begin{aligned}
\hat{\mathcal{H}}_I^{\text{spin}} &= - \frac{q}{c} \, \underline{j}_s \cdot \underline{A} \\
&= + \frac{q\,\hbar}{2\,m\,c} \, \underline{A} \cdot \left(\underline{\nabla} \wedge (\, \overline{\psi}^\dagger \, \hat{\underline{\sigma}} \, \psi \,) \right)
\end{aligned}
\tag{1398}
$$

On integrating over space, the interaction Hamiltonian density gives rise to the interactions contribution to the total energy. By integrating by parts, it can be shown that this energy contribution is equivalent to the energy contribution caused by an equivalent form of the interaction Hamiltonian density

$$
\begin{aligned}
\hat{\mathcal{H}}_I^{\text{spin}} &\equiv - \frac{q\,\hbar}{2\,m\,c} \, (\, \overline{\psi}^\dagger \, \hat{\underline{\sigma}} \, \psi \,) \cdot (\, \underline{\nabla} \wedge \underline{A} \,) \\
&\equiv - \frac{q\,\hbar}{2\,m\,c} \, (\, \overline{\psi}^\dagger \, \hat{\underline{\sigma}} \, \psi \,) \cdot \underline{B}
\end{aligned}
\tag{1399}
$$

where $\underline{B}$ is the magnetic field. Hence, the interaction energy contains a term which represents an interaction between the electron's internal degree of freedom and the magnetic field.

12.5   Lorentz Covariance of the Dirac Equation

One goal of Physics is to write the laws in a manner which are independent of any arbitrary choices that are made. Within special relativity, this implies that the laws of Physics should be written in a way which is independent of the choice of inertial reference frame. Dirac's theory is Lorentz covariant if the results are independent of the Lorentz frame used. To this end, it is required that the Dirac equation in a Lorentz transformed frame of reference has the same form as the Dirac equation in the original reference frame, and also that the solutions of these two equations describe the same physical states. That is, the two solutions must describe the same set of measurable properties in the different reference frames, and therefore the results are simply related by the Lorentz transformation.

The first step of the proof of the Lorentz covariance of the Dirac equation requires that one should be able to show that under a Lorentz transformation defined by

$$A^\mu \;\rightarrow\; A^{\mu\prime} \;=\; \Lambda^\mu{}_\nu \; A^\nu \tag{1400}$$

then the Dirac equation is transformed from

$$\gamma^\mu \left(\, \hat{p}_\mu \,-\, \frac{q}{c} \, A_\mu \, \right) \psi \;=\; m \, c \, \psi \tag{1401}$$

to an equation with an equivalent form

$$\gamma^{\mu\prime} \left(\, \hat{p}'_\mu \,-\, \frac{q}{c} \, A'_\mu \, \right) \psi' \;=\; m \, c \, \psi' \tag{1402}$$

Furthermore, the four components of the spinor wave function ψ' are assumed to be linearly related to the components of ψ by a four by four matrix $\hat{\mathcal{R}}(\Lambda)$ which is independent of x^μ

$$\psi'(x') \;=\; \hat{\mathcal{R}}(\Lambda) \, \psi(x) \tag{1403}$$

Hence, the transformed Dirac equation can be re-written in terms of the untransformed spinor

$$\gamma^{\mu\prime} \left(\, \hat{p}'_\mu \,-\, \frac{q}{c} \, A'_\mu \, \right) \psi' \;=\; m \, c \, \psi'$$

$$\gamma^{\mu\prime} \left(\, \hat{p}'_\mu \,-\, \frac{q}{c} \, A'_\mu \, \right) \hat{\mathcal{R}}(\Lambda) \, \psi \;=\; m \, c \, \hat{\mathcal{R}}(\Lambda) \, \psi \tag{1404}$$

if such an $\hat{\mathcal{R}}(\Lambda)$ exists. The $\gamma^{\mu\prime}$ matrices must satisfy the same anti-commutation relations as the γ^μ and, therefore, only differ from them by a similarity trans-

formation[99]. The transformations of $\gamma^{\mu\prime}$ just results in the set of the four linear equations that compose the Dirac equation being combined in different ways, so this rearrangement can be absorbed in the definition of $\hat{\mathcal{R}}(\Lambda)$. That is, one can choose to impose the convention that $\gamma^{\mu\prime} = \gamma^{\mu}$. The transformed Dirac equation can be expressed as

$$\gamma^{\mu\prime} \, (\, \hat{p}'_{\mu} \, - \, \frac{q}{c} \, A'_{\mu} \,) \, \hat{\mathcal{R}}(\Lambda) \, \psi \;\; = \;\; m \, c \, \hat{\mathcal{R}}(\Lambda) \, \psi$$

$$\gamma^{\mu\prime} \, \Lambda_{\mu}{}^{\nu} \, (\, \hat{p}_{\nu} \, - \, \frac{q}{c} \, A_{\nu} \,) \, \hat{\mathcal{R}}(\Lambda) \, \psi \;\; = \;\; m \, c \, \hat{\mathcal{R}}(\Lambda) \, \psi \qquad (1405)$$

where the transformation properties of the momentum four-vector have been used[100]. On multiplying by the inverse of $\hat{\mathcal{R}}(\Lambda)$, one has

$$\hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\mu\prime} \, \Lambda_{\mu}{}^{\nu} \, (\, \hat{p}_{\nu} \, - \, \frac{q}{c} \, A_{\nu} \,) \, \hat{\mathcal{R}}(\Lambda) \, \psi \;\; = \;\; m \, c \, \psi \qquad (1406)$$

$$\hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\mu\prime} \, \hat{\mathcal{R}}(\Lambda) \, \Lambda_{\mu}{}^{\nu} \, (\, \hat{p}_{\nu} \, - \, \frac{q}{c} \, A_{\nu} \,) \, \psi \;\; = \;\; m \, c \, \psi \qquad (1407)$$

where the four by four matrices $\hat{\mathcal{R}}(\Lambda)$ have been commuted with the differential operators and also with the components of the Lorentz transform. The condition for covariance as

$$\hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\mu\prime} \, \hat{\mathcal{R}}(\Lambda) \, \Lambda_{\mu}{}^{\nu} \;\; = \;\; \gamma^{\nu} \qquad (1408)$$

The transformed Dirac equation has the same form as the original equation if the transformed $\gamma^{\mu\prime}$ matrices satisfy the same anti-commutations and conditions as the unprimed matrices. This can be achieved by choosing $\gamma^{\mu\prime} = \gamma^{\mu}$. This choice yields the condition for covariance as

$$\hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\mu} \, \hat{\mathcal{R}}(\Lambda) \, \Lambda_{\mu}{}^{\nu} \;\; = \;\; \gamma^{\nu} \qquad (1409)$$

Since for a Lorentz transform one has

$$\Lambda_{\mu}{}^{\nu} \, \Lambda^{\rho}{}_{\nu} \;\; = \;\; \delta_{\mu}{}^{\rho} \qquad (1410)$$

then multiplying the above covariance condition by $\Lambda^{\rho}{}_{\nu}$ leads to

$$\hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\mu} \, \hat{\mathcal{R}}(\Lambda) \;\; = \;\; \Lambda^{\mu}{}_{\nu} \, \gamma^{\nu} \qquad (1411)$$

The above equation determines the 4×4 matrix $\hat{\mathcal{R}}(\Lambda)$. If $\hat{\mathcal{R}}(\Lambda)$ exits, the Dirac equation has the same form in the two frames of reference and the solutions are linearly related. Pauli's "fundamental theorem" guarantees that a matrix $\hat{\mathcal{R}}(\Lambda)$ exists which does satisfy the condition. Instead of following the general theorem, the solution will be inferred from consideration of infinitesimal Lorentz

---

[99] This is a statement of Pauli's fundamental theorem [W. Pauli, Ann. Inst. Henri Poincaré **6**, 109 (1936).]. For a general discussion, see R. H. Good Jr. Rev. Mod. Phys. **27**, 187 (1955).

[100] It should be noted that the matrices $\Lambda_{\mu}{}^{\nu}$ and $\hat{\mathcal{R}}$ act on totally different spaces. The matrices $\Lambda_{\mu}{}^{\nu}$ act on the components of the four-vectors x^{ν}, whereas the $\hat{\mathcal{R}}$ matrices act on the components of the four-component Dirac spinor ψ.

transformations.

The matrix $\hat{\mathcal{R}}(\Lambda)$ will be determined by considering the effect of an infinitesimal Lorentz transformation

$$\Lambda^\mu{}_\nu \;=\; \delta^\mu{}_\nu \;+\; \epsilon^\mu{}_\nu \;+\; \ldots \tag{1412}$$

where $\delta^\mu{}_\nu$ is the Kronecker delta function. The matrix $\hat{\mathcal{R}}(\Lambda)$ for the infinitesimal transformation can also be expanded as

$$\hat{\mathcal{R}} \;=\; \hat{I} \;-\; \frac{i}{4}\,\epsilon^\mu{}_\nu\,\omega_\mu{}^\nu \;+\; \ldots \tag{1413}$$

where $\omega_\mu{}^\nu$ is a four by four matrix that has yet to be determined. The inverse matrix can be written as

$$\hat{\mathcal{R}}^{-1} \;=\; \hat{I} \;+\; \frac{i}{4}\,\epsilon^\mu{}_\nu\,\omega_\mu{}^\nu \;+\; \ldots \tag{1414}$$

to first-order in the infinitesimal quantity $\epsilon^\mu{}_\nu$. On substituting the matrices for the infinitesimal transform into the equation that determines $\hat{\mathcal{R}}$, one obtains

$$\frac{i}{4}\,\epsilon^\rho{}_\sigma\left(\,\omega_\rho{}^\sigma\,\gamma^\mu \;-\; \gamma^\mu\,\omega_\rho{}^\sigma\,\right) \;=\; \epsilon^\mu{}_\nu\,\gamma^\nu \tag{1415}$$

or on raising and lowering indices

$$\frac{i}{4}\,\epsilon_{\rho\sigma}\left(\,\omega^{\rho\sigma}\,\gamma^\mu \;-\; \gamma^\mu\,\omega^{\rho\sigma}\,\right) \;=\; g^{\mu\rho}\,\epsilon_{\rho\sigma}\,\gamma^\sigma \tag{1416}$$

Thus, since $\epsilon_{\rho\sigma}$ is anti-symmetric as it represents an infinitesimal Lorentz transformation, the matrix $\omega^{\rho\sigma}$ can be restricted to be anti-symmetric in the indices, because any symmetric part does not contribute to the matrix $\hat{\mathcal{R}}$. By making specific choices for the anti-symmetric quantities $\epsilon_{\rho\sigma}$, which are zero except for a chosen pair of indices (say α and β), one finds that the anti-symmetric part of $\omega^{\alpha\beta}$ is determined from the equation

$$\frac{i}{2}\,[\,\omega^{\alpha\beta}\,,\,\gamma^\mu\,] \;=\; g^{\mu\alpha}\,\gamma^\beta \;-\; g^{\mu\beta}\,\gamma^\alpha \tag{1417}$$

These sets of equations have to be satisfied even if arbitrary choices are made for the infinitesimal Lorentz transformations $\epsilon_{\rho\sigma}$. The infinitesimal unitary matrix $\hat{\mathcal{R}}$ can be expressed in terms of six generators $\omega^{\rho\sigma}$ of the infinitesimal Lorentz transformation

$$\hat{\mathcal{R}} \;=\; \hat{I} \;-\; \frac{i}{4}\,\epsilon_{\rho\sigma}\,\omega^{\rho\sigma} \;+\; \ldots \tag{1418}$$

The set of matrices $\omega^{\rho\sigma}$ that define $\hat{\mathcal{R}}$ must satisfy the equation

$$\frac{i}{2}\,[\,\omega^{\rho\sigma}\,,\,\gamma^\mu\,] \;=\; g^{\mu\rho}\,\gamma^\sigma \;-\; g^{\mu\sigma}\,\gamma^\rho \tag{1419}$$

The set of (as yet unknown) matrices $\omega^{\rho\sigma}$ that solve the above set of equations are given by

$$\omega^{\alpha\beta} \;=\; \sigma^{\alpha\beta} \;=\; \frac{i}{2}\,[\,\gamma^\alpha\,,\,\gamma^\beta\,] \tag{1420}$$

which are the six generators of the general infinitesimal Lorentz transformation. This solution, and hence, the existence of $\hat{\mathcal{R}}(\Lambda)$ shows that the solutions of the Dirac equation and the transformed equation are in a one to one correspondence.

Proof of Solution

It can be shown that the expression for $\sigma^{\alpha,\beta}$ given in eqn(1420) satisfies the requirement of eqn(1419), by evaluating the nested commutator through repeatedly using the anti-commutation properties of the γ matrices. The commutator can be expressed as a nested commutator or as the sum of two commutators

$$
\begin{aligned}
[\,\sigma^{\alpha\beta}\,,\,\gamma^\mu\,] \;&=\; \frac{i}{2}\,[\,[\,\gamma^\alpha\,,\,\gamma^\beta\,]\,,\,\gamma^\mu\,] \\
&=\; \frac{i}{2}\,[\,\gamma^\alpha\,\gamma^\beta\,,\,\gamma^\mu\,] \;-\; \frac{i}{2}\,[\,\gamma^\beta\,\gamma^\alpha\,,\,\gamma^\mu\,]
\end{aligned}
\tag{1421}
$$

On using the anti-commutation relation for the γ matrices

$$\frac{1}{2}\left(\gamma^\alpha\,\gamma^\beta \;+\; \gamma^\beta\,\gamma^\alpha\right) \;=\; g^{\alpha,\beta}\,\hat{I} \tag{1422}$$

one can eliminate the second term leading to

$$
\begin{aligned}
[\,\sigma^{\alpha\beta}\,,\,\gamma^\mu\,] \;&=\; i\,[\,\gamma^\alpha\,\gamma^\beta\,,\,\gamma^\mu\,] \;+\; i\,g^{\alpha,\beta}\,[\,\hat{I}\,,\,\gamma^\mu\,] \\
&=\; i\,[\,\gamma^\alpha\,\gamma^\beta\,,\,\gamma^\mu\,]
\end{aligned}
\tag{1423}
$$

where the second line follows since the identity matrix commutes with γ^μ. One notices that if the γ^μ's are anti-commuted to the center of each product, some terms will cancel and there may be some simplification. On using the anti-commutation relation in the second term of the expression

$$[\,\sigma^{\alpha\beta}\,,\,\gamma^\mu\,] \;=\; i\left(\gamma^\alpha\,\gamma^\beta\,\gamma^\mu \;-\; \gamma^\mu\,\gamma^\alpha\,\gamma^\beta\right) \tag{1424}$$

one finds

$$[\,\sigma^{\alpha\beta}\,,\,\gamma^\mu\,] \;=\; i\left(\gamma^\alpha\,\gamma^\beta\,\gamma^\mu \;+\; \gamma^\alpha\,\gamma^\mu\,\gamma^\beta \;-\; 2\,g^{\mu,\alpha}\,\gamma^\beta\right) \tag{1425}$$

Likewise, the γ matrices in the first term can also be anti-commuted, leading to

$$
\begin{aligned}
[\,\sigma^{\alpha\beta}\,,\,\gamma^\mu\,] \;&=\; i\left(2\,g^{\mu,\beta}\,\gamma^\alpha \;-\; \gamma^\alpha\,\gamma^\mu\,\gamma^\beta \;+\; \gamma^\alpha\,\gamma^\mu\,\gamma^\beta \;-\; 2\,g^{\mu,\alpha}\,\gamma^\beta\right) \\
&=\; 2\,i\left(g^{\mu,\beta}\,\gamma^\alpha \;-\; g^{\mu,\alpha}\,\gamma^\beta\right)
\end{aligned}
\tag{1426}
$$

244

since the middle pair of terms cancel. Hence, one has proved that

$$\frac{i}{2} \left[\sigma^{\alpha\beta} , \gamma^{\mu} \right] = \left(g^{\mu,\alpha} \gamma^{\beta} - g^{\mu,\beta} \gamma^{\alpha} \right) \qquad (1427)$$

which completes the identification of the solution of the equation for $\omega^{\alpha,\beta}$. Therefore, since $\hat{\mathcal{R}}(\Lambda)$ exists, it has been shown that the form of the Dirac equation is maintained in the primed reference frame and that there is a one to one correspondence between the solutions of the primed and unprimed frames.

Equivalence of Physical Properties

It remains to be shown that the ψ and ψ' describe the properties of the same physical system, albeit in two different frames of reference. That is, the properties associated with ψ must be related to the properties of ψ' and the relation can be obtained by considering the Lorentz transformation. The most complete physical descriptions of a unique quantum mechanical state are related to the probability density, which can only be inferred from an infinite set of position measurements. The probability density, should behave similarly to the time component of a four-vector as was seen from the consideration of the continuity equation. Therefore, it follows that if the four-vector probability currents of ψ and ψ' are related via a Lorentz transformation, then the two spinors describe the same physical state of the system.

The probability current four-vector j^{μ} in the unprimed frame is described by

$$\begin{aligned} j^{\mu} &= c \, \overline{\psi}^{\dagger} \, \gamma^{\mu} \, \psi \\ &= c \, \psi^{\dagger} \, \gamma^{(0)} \, \gamma^{\mu} \, \psi \end{aligned} \qquad (1428)$$

and in the primed frame, one has

$$\begin{aligned} j^{\mu\prime} &= c \, \psi^{\dagger\prime} \, \gamma^{(0)} \, \gamma^{\mu} \, \psi' \\ &= c \, \psi^{\dagger} \, \hat{\mathcal{R}}^{\dagger} \, \gamma^{(0)} \, \gamma^{\mu} \, \hat{\mathcal{R}} \, \psi \end{aligned} \qquad (1429)$$

The identity

$$\hat{\mathcal{R}}^{-1} = \gamma^{(0)} \, \hat{\mathcal{R}}^{\dagger} \, \gamma^{(0)} \qquad (1430)$$

will be proved below, so on using this identity together with

$$\gamma^{(0)} \, \gamma^{(0)} = \hat{I} \qquad (1431)$$

the probability current density can be re-written as

$$\begin{aligned} j^{\mu\prime} &= c \, \psi^{\dagger} \, \gamma^{(0)} \, \gamma^{(0)} \, \hat{\mathcal{R}}^{\dagger} \, \gamma^{(0)} \, \gamma^{\mu} \, \hat{\mathcal{R}} \, \psi \\ &= c \, \psi^{\dagger} \, \gamma^{(0)} \, \hat{\mathcal{R}}^{-1} \, \gamma^{\mu} \, \hat{\mathcal{R}} \, \psi \end{aligned} \qquad (1432)$$

However, because the covariant condition is given by

$$\hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^\mu \, \hat{\mathcal{R}}(\Lambda) \;=\; \Lambda^\mu{}_\nu \, \gamma^\nu \tag{1433}$$

the current density can be expressed as

$$
\begin{aligned}
j^{\mu\prime} \;&=\; c \, \psi^\dagger \, \gamma^{(0)} \, \Lambda^\mu{}_\nu \, \gamma^\nu \, \psi \\
&=\; \Lambda^\mu{}_\nu \, c \, \overline{\psi}^\dagger \, \gamma^\nu \, \psi \\
&=\; \Lambda^\mu{}_\nu \, j^\nu
\end{aligned}
\tag{1434}
$$

Hence, the probability current densities $j^{\mu\prime}$ and j^μ found in the two reference frames are simply related via the Lorentz transformation. Therefore, the Dirac equation gives consistent results, no matter what inertial frame of reference is used.

Proof of Identity

The identity

$$\hat{\mathcal{R}}^{-1}(\Lambda) \;=\; \gamma^{(0)} \, \hat{\mathcal{R}}^\dagger(\Lambda) \, \gamma^{(0)} \tag{1435}$$

can be proved by starting from the expression for the expression for $\hat{\mathcal{R}}$ appropriate for infinitesimal transformation given by

$$\hat{\mathcal{R}} \;=\; \hat{I} \,+\, \frac{1}{8} \, \epsilon_{\mu\nu} \, [\, \gamma^\mu \,,\, \gamma^\nu \,] \,+\, \dots \tag{1436}$$

Hence, the Hermitean conjugate is given by

$$
\begin{aligned}
\hat{\mathcal{R}}^\dagger \;&=\; \hat{I} \,+\, \frac{1}{8} \, \epsilon_{\mu\nu} \, [\, \gamma^{\nu\dagger} \,,\, \gamma^{\mu\dagger} \,] \,+\, \dots \\
\hat{\mathcal{R}}^\dagger \;&=\; \hat{I} \,-\, \frac{1}{8} \, \epsilon_{\mu\nu} \, [\, \gamma^{\mu\dagger} \,,\, \gamma^{\nu\dagger} \,] \,+\, \dots
\end{aligned}
\tag{1437}
$$

since the Hermitean conjugate of a product is the product of the Hermitean conjugate of the factors taken in opposite order. On forming the product $\gamma^{(0)} \, \hat{\mathcal{R}}^\dagger \, \gamma^{(0)}$ and inserting a factor of

$$\gamma^{(0)} \, \gamma^{(0)} \;=\; \hat{I} \tag{1438}$$

between the pairs of four by four γ matrices in the commutator and noting that

$$\gamma^{(0)} \, \gamma^{\mu\dagger} \, \gamma^{(0)} \;=\; \gamma^\mu \tag{1439}$$

one finds that

$$
\begin{aligned}
\gamma^{(0)} \, \hat{\mathcal{R}}^\dagger \, \gamma^{(0)} \;&=\; \hat{I} \,-\, \frac{1}{8} \, \epsilon_{\mu\nu} \, [\, \gamma^\mu \,,\, \gamma^\nu \,] \,+\, \dots \\
&=\; \hat{\mathcal{R}}^{-1}
\end{aligned}
\tag{1440}
$$

The last line follows from the observation that on combining the expression for $\hat{\mathcal{R}}$ with the expression for $\gamma^{(0)}\,\hat{\mathcal{R}}^{\dagger}\,\gamma^{(0)}$, the terms of order ϵ cancel. Hence to the order of ϵ^2, the product $\gamma^{(0)}\,\hat{\mathcal{R}}^{\dagger}\,\gamma^{(0)}$ coincides with $\hat{\mathcal{R}}^{-1}$. This concludes the discussion of the desired identity.

Finite Rotations

Consider a finite rotation of the coordinate system specified by the transformation matrix Λ

$$x^{\mu\prime} \; = \; \Lambda^{\mu}{}_{\nu}\, x^{\nu} \tag{1441}$$

Specifically, a finite (passive) rotation through an angle φ about the $\hat{e}_3$ direction can be expressed in terms of the transformation matrix

$$\Lambda \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & +\sin\varphi & 0 \\ 0 & -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1442}$$

The above transformation represents a rotation of the coordinate system while the physical system stays put. For an infinitesimal rotation through $\delta\varphi$, the
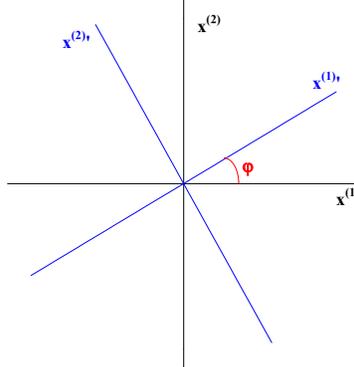
Figure 45: A passive rotation of the coordinate system through an angle φ about the $\hat{e}_3$-axis.

transformation matrix reduces to

$$\Lambda \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & +\delta\varphi & 0 \\ 0 & -\delta\varphi & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \ldots \tag{1443}$$

to first-order in the infinitesimal quantity $\delta\varphi$. Therefore, with the infinitesimal form of the general Lorentz transformation

$$\Lambda^{\mu}{}_{\nu} \;=\; \delta^{\mu}{}_{\nu} \,+\, \epsilon^{\mu}{}_{\nu} \,+\, \ldots \tag{1444}$$

247

on lowering the first index, one identifies

$$\epsilon_{12} \;=\; -\,\epsilon_{21} \;=\; -\,\delta\varphi \tag{1445}$$

The infinitesimal transformation of a Dirac spinor was determined to be given by

$$\hat{\mathcal{R}}(\delta\varphi) \;=\; \hat{I} \;-\; \frac{i}{4}\,\epsilon_{\mu\nu}\,\sigma^{\mu\nu} \;+\; \ldots \tag{1446}$$

Hence, for a infinitesimal rotation one has

$$
\begin{aligned}
\hat{\mathcal{R}}(\delta\varphi) \;&=\; \hat{I} \;+\; \frac{i}{4}\,\big(\,\delta\varphi\,\sigma^{1,2} - \delta\varphi\,\sigma^{2,1}\,\big) \;+\; \ldots \\
&=\; \hat{I} \;+\; \frac{i}{2}\,\delta\varphi\,\sigma^{1,2} \;+\; \ldots \\
&=\; \exp\!\left[\, i\,\frac{\delta\varphi}{2}\,\sigma^{1,2} \,\right]
\end{aligned}
\tag{1447}
$$

since $\sigma^{\mu,\nu}$ is anti-symmetric. On compounding N infinitesimal transformations about the same axis $\hat{\mathcal{R}}(\delta\varphi)$ using their exponential form, and defining $N\,\delta\varphi = \varphi$, one obtains the finite rotation $\hat{\mathcal{R}}(\varphi)$

$$
\begin{aligned}
\hat{\mathcal{R}}(\varphi) \;&=\; \Big(\,\hat{\mathcal{R}}(\delta\varphi)\,\Big)^{N} \\
&=\; \exp\!\left[\, i\,N\,\frac{\delta\varphi}{2}\,\sigma^{1,2} \,\right] \\
&=\; \exp\!\left[\, i\,\frac{\varphi}{2}\,\sigma^{1,2} \,\right]
\end{aligned}
\tag{1448}
$$

Therefore, for a finite rotation, the transformation matrix is given by

$$\hat{\mathcal{R}}(\varphi) \;=\; \exp\!\left[\, i\,\frac{\varphi}{2}\,\sigma^{1,2} \,\right] \tag{1449}$$

which can be expressed in terms of even and odd-powers of $\sigma^{1,2}$ via

$$\hat{\mathcal{R}}(\varphi) \;=\; \cos\!\left[\,\frac{\varphi}{2}\,\sigma^{1,2}\,\right] \;+\; i\,\sin\!\left[\,\frac{\varphi}{2}\,\sigma^{1,2}\,\right] \tag{1450}$$

but since

$$
\begin{aligned}
\sigma^{1,2} \;&=\; \frac{i}{2}\,[\,\gamma^{(1)}\,,\,\gamma^{(2)}\,] \\
&=\; \hat{\sigma}^{(3)}
\end{aligned}
\tag{1451}
$$

the transformation can be expressed as

$$\hat{\mathcal{R}}(\varphi) \;=\; \cos\!\left[\,\frac{\varphi}{2}\,\hat{\sigma}^{(3)}\,\right] \;+\; i\,\sin\!\left[\,\frac{\varphi}{2}\,\hat{\sigma}^{(3)}\,\right] \tag{1452}$$

The above expression can be simplified by expanding the trigonometric functions in series of φ and then using the property of the $\hat{\sigma}^{(j)}$ matrices

$$(\hat{\sigma}^{(3)})^2 = \hat{I} \tag{1453}$$

Since the repeated use of the above identity leads to

$$\begin{aligned}
(\hat{\sigma}^{(3)})^{2n} &= \hat{I} \\
(\hat{\sigma}^{(3)})^{2n+1} &= \hat{\sigma}^{(3)}
\end{aligned} \tag{1454}$$

the series simplify and can be re-summed leading to

$$\hat{\mathcal{R}}(\varphi) = \cos\left[\frac{\varphi}{2} \right] \hat{I} + i \, \sin\left[\frac{\varphi}{2} \right] \hat{\sigma}^{(3)} \tag{1455}$$

Therefore, under a finite rotation through angle φ around the unit vector $\hat{e}$, a spinor is rotated by the operator

$$\hat{\mathcal{R}}(\varphi) = \cos\frac{\varphi}{2} \, \hat{I} + i \, \sin\frac{\varphi}{2} \, \hat{e} \cdot \underline{\hat{\sigma}} \tag{1456}$$

From the above equation, due to the presence of the half-angle, one notes that a rotation φ and through $\varphi + 2\pi$ are not equivalent, since

$$\hat{\mathcal{R}}(\varphi + 2\pi) = - \hat{\mathcal{R}}(\varphi) \tag{1457}$$

which changes the sign of the spinor. For spin one-half electrons, it is necessary to rotate through 4π to return to the same state

$$\hat{\mathcal{R}}(\varphi + 4\pi) = \hat{\mathcal{R}}(\varphi) \tag{1458}$$

A quantity which is bi-linear in $\overline{\psi}^{\dagger}$ and ψ will remain invariant under a rotation of 2π.

Finite Lorentz Boosts

A finite Lorentz boost by velocity v along the $\hat{e}_1$ direction can be expressed in terms of the transformation

$$\Lambda = \begin{pmatrix} \cosh\chi & -\sinh\chi & 0 & 0 \\ -\sinh\chi & \cosh\chi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1459}$$

where the rapidity χ is defined by

$$\tanh\chi = \frac{v}{c} \tag{1460}$$

so

$$\cosh \chi \;\; = \;\; \frac{1}{\sqrt{1 - (\frac{v}{c})^2}}$$

$$\sinh \chi \;\; = \;\; \frac{\frac{v}{c}}{\sqrt{1 - (\frac{v}{c})^2}} \tag{1461}$$

For an infinitesimal boost through $\delta\chi$, the transformation matrix reduces to

$$\Lambda \;\; = \;\; \begin{pmatrix} 1 & -\,\delta\chi & 0 & 0 \\ -\,\delta\chi & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \; + \; \ldots \tag{1462}$$

to first-order in the infinitesimal quantity $\delta\chi$. Therefore, with the infinitesimal form of the general Lorentz transformation

$$\Lambda^\mu{}_\nu \;\; = \;\; \delta^\mu{}_\nu \; + \; \epsilon^\mu{}_\nu \; + \; \ldots \tag{1463}$$

on lowering the first index, one identifies

$$\epsilon_{01} \;\; = \;\; -\,\epsilon_{10} \;\; = \;\; -\,\delta\chi \tag{1464}$$

The infinitesimal transformation of a Dirac spinor was determined to be given by

$$\hat{\mathcal{R}}(\delta\chi) \;\; = \;\; \hat{I} \; - \; \frac{i}{4}\,\epsilon_{\mu\nu}\,\sigma^{\mu\nu} \; + \; \ldots \tag{1465}$$

Hence, for a infinitesimal Lorentz boost one has

$$\hat{\mathcal{R}}(\delta\chi) \;\; = \;\; \hat{I} \; + \; 2\,\frac{i}{4}\,\delta\chi\,\sigma^{0,1} \; + \; \ldots$$

$$= \;\; \exp\left[\, i\,\frac{\delta\chi}{2}\,\sigma^{0,1} \,\right] \tag{1466}$$

On compounding N successive infinitesimal Lorentz boosts (with parallel velocities) given by $\hat{\mathcal{R}}(\delta\chi)$ and defining $N\,\delta\chi \;=\; \chi$, one obtains the finite Lorentz boost $\hat{\mathcal{R}}(\chi)$

$$\hat{\mathcal{R}}(\chi) \;\; = \;\; \left(\, \hat{\mathcal{R}}(\delta\chi) \,\right)^N$$

$$= \;\; \exp\left[\, i\,N\,\frac{\delta\chi}{2}\,\sigma^{0,1} \,\right]$$

$$= \;\; \exp\left[\, i\,\frac{\chi}{2}\,\sigma^{0,1} \,\right] \tag{1467}$$

Therefore, for a finite Lorentz boost, the transformation matrix is given by

$$\hat{\mathcal{R}}(\chi) \;\; = \;\; \exp\left[\, i\,\frac{\chi}{2}\,\sigma^{0,1} \,\right] \tag{1468}$$

which can be expressed in terms of even and odd-powers of $\sigma^{0,1}$ via

$$\hat{\mathcal{R}}(\chi) \;=\; \cosh\left[\, i \,\frac{\chi}{2}\, \sigma^{0,1} \,\right] \;+\; \sinh\left[\, i \,\frac{\chi}{2}\, \sigma^{0,1} \,\right] \tag{1469}$$

but since

$$\begin{aligned}
\sigma^{0,1} &= \frac{i}{2}\,[\,\gamma^{(0)}\,,\,\gamma^{(1)}\,] \\
&= i\,\alpha^{(1)} \tag{1470}
\end{aligned}$$

the transformation can be expressed as

$$\hat{\mathcal{R}}(\chi) \;=\; \cosh\left[\,-\,\frac{\chi}{2}\,\alpha^{(1)}\,\right] \;+\; \sinh\left[\,-\,\frac{\chi}{2}\,\alpha^{(1)}\,\right] \tag{1471}$$

The above expression can be simplified by expanding the hyperbolic functions in series of χ and then using the property of the α matrices

$$(\,\alpha^{(1)}\,)^2 \;=\; \hat{I} \tag{1472}$$

Since the repeated use of the above identity leads to

$$\begin{aligned}
(\,\alpha^{(1)}\,)^{2n} &= \hat{I} \\
(\,\alpha^{(1)}\,)^{2n+1} &= \alpha^{(1)} \tag{1473}
\end{aligned}$$

the series simplify and can be re-summed leading to

$$\begin{aligned}
\hat{\mathcal{R}}(\chi) &= \cosh\left[\,-\,\frac{\chi}{2}\,\right]\hat{I} \;+\; \sinh\left[\,-\,\frac{\chi}{2}\,\right]\alpha^{(1)} \\
&= \cosh\left[\,\frac{\chi}{2}\,\right]\hat{I} \;-\; \sinh\left[\,\frac{\chi}{2}\,\right]\alpha^{(1)} \tag{1474}
\end{aligned}$$

Therefore, under a finite boost through velocity $\underline{v}$, a spinor is rotated by the operator

$$\begin{aligned}
\hat{\mathcal{R}}(\chi) &= \cosh\frac{\chi}{2}\left[\,\hat{I} \;-\; \tanh\frac{\chi}{2}\,\alpha^{(1)}\,\right] \\
&= \cosh\frac{\chi}{2}\left[\,\hat{I} \;-\; \tanh\frac{\chi}{2}\,\hat{v}\,.\,\underline{\alpha}\,\right] \tag{1475}
\end{aligned}$$

where the rapidity χ is determined by

$$\tanh\chi \;=\; \frac{v}{c} \tag{1476}$$

Exercise:

Determine the relationship between the rapidities for a combined Lorentz transformation consisting of two successive Lorentz boosts with parallel velocities v_0 and v_1.

Exercise:

Starting from a solution of a free stationary Dirac particle with spin σ, perform a Lorentz boost to determine the solution for a Dirac electron with momentum $\underline{p}'$.
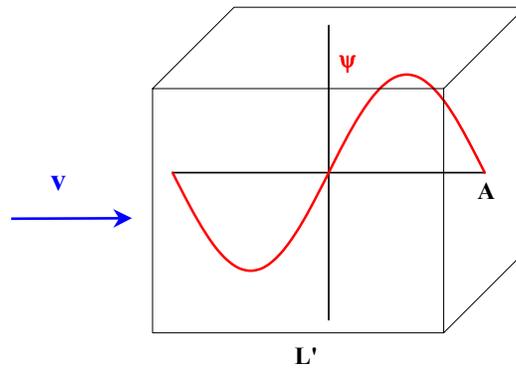
Figure 46: A cartoon depicting a stationary free-electron confined in a volume V with proper length L, viewed from a coordinate system moving with velocity v anti-parallel to the $\hat{e}_1$-axis.

Exercise:

Show that the helicity eigenvalue of a free Dirac particle can be reversed by going to a new reference frame which is "overtaking" the particle.

12.5.1   The Space of the Anti-commuting γ^μ-Matrices.

One can form sixteen matrices Γ_i from the product of the four γ matrices. Since the γ^μ matrices obey the anti-commutation relations

$$\{ \, \gamma^\mu \, , \, \gamma^\nu \, \}_+ \; = \; 2 \, g^{\mu,\nu} \, \hat{I} \tag{1477}$$

all other products can be reduced to the above products. The order of the matrices is irrelevant, since the different matrices anti-commute. Also, since $(\, \gamma^\mu \,)^2 \; = \; \pm \, \hat{I}$, one only needs to consider the products in which each matrix enters at most one time. Hence, since each of the four matrices either appear as a factor or do not, there are only 2^4 such matrices. These sixteen Γ_i matrices can be constructed from $\hat{I}$, γ^μ, $\sigma^{\mu,\nu} \; = \; i \, \gamma^\mu \, \gamma^\nu$, $\gamma^{(4)}$ and $\gamma^{(4)} \, \gamma^\mu$, by choosing

Table 6: The Set of the Sixteen Matrices Γ_n with their Phase Factors $(j > i)$

$$\hat{I}$$

$\gamma^{(0)}$	$i\,\gamma^{(i)}$
$\gamma^{(0)}\,\gamma^{(i)}$	$i\,\varepsilon^{i,j,k}\,\gamma^{(i)}\,\gamma^{(j)}$

$$\gamma^{(4)} \;=\; i\,\gamma^{(0)}\,\gamma^{(1)}\,\gamma^{(2)}\,\gamma^{(3)}$$

$-\,i\,\gamma^{(0)}\,\gamma^{(4)}$	$\gamma^{(i)}\,\gamma^{(4)}$

appropriate phase factors.

Closure under Multiplication

The set of matrices Γ_i formed from the set of γ^μ are closed under multiplication, so

$$\Gamma_i\,\Gamma_j \;=\; a_{i,j}\,\Gamma_k \tag{1478}$$

where $a_{i,j}^4 \;=\; 1$. The sixteen Γ_i matrices can be chosen as the product of the members of the above set multiplied by a phase factor taken from the set $\pm\,1$ and $\pm\,i$, such that the condition

$$(\,\Gamma_i\,)^2 \;=\; \hat{I} \tag{1479}$$

is satisfied. Furthermore, by counting the number of non-equivalent factors of the γ^μ in the products, one can show that

$$\Gamma_i\,\Gamma_j \;=\; \hat{I} \quad \text{only if} \quad i \;=\; j \tag{1480}$$

Also, by anti-commuting the factors of γ^μ in the products, one can show that

$$\Gamma_i\,\Gamma_j \;=\; \pm\,\Gamma_j\,\Gamma_i \tag{1481}$$

Specifically, for a fixed Γ_i not equal to the identity, one can always find a specific Γ_k such that

$$\Gamma_i\,\Gamma_k \;=\; -\,\Gamma_k\,\Gamma_i \tag{1482}$$

which on multiplying by Γ_k results in

$$\Gamma_k\,\Gamma_i\,\Gamma_k \;=\; -\,\Gamma_i \tag{1483}$$

Traceless Matrices

The above facts can be used to show that the Γ_i matrices, other than the identity, are traceless. This can be proved by considering

$$
\begin{aligned}
- \text{Trace } \Gamma_i &= \text{Trace}(- \Gamma_i) = \text{Trace}(\Gamma_k \, \Gamma_i \, \Gamma_k) \\
&= \text{Trace}(\Gamma_i \, \Gamma_k \, \Gamma_k) = \text{Trace } \Gamma_i
\end{aligned}
\tag{1484}
$$

in which the existence of a specific Γ_k which anti-commutes with Γ_i has been used, and where the cyclic invariance of the trace has been used as has been $(\Gamma_k)^2 = \hat{I}$. Hence, all the Γ_i matrices, other than the identity, are traceless

$$
\text{Trace } \Gamma_i = 0
\tag{1485}
$$

Linear Independence

The sixteen Γ_i matrices are linearly independent. The linear independence can be expressed in terms of the absence of any non-trivial solution of the equation

$$
\sum_i C_i \, \Gamma_i = 0
\tag{1486}
$$

other than $C_i \equiv 0$ for all i. If the Γ_i are linearly independent, the only solution of this equation is

$$
C_i \equiv 0 \qquad \text{for all } i
\tag{1487}
$$

This can be proved by multiplying eqn(1486) by any one Γ_j in the set which leads to

$$
\begin{aligned}
C_j \, \hat{I} + \sum_{i \neq j} C_i \, \Gamma_i \, \Gamma_j &= 0 \\
C_j \, \hat{I} + \sum_{i \neq j} C_i \, a_{i,j} \, \Gamma_k &= 0
\end{aligned}
\tag{1488}
$$

On taking the trace one finds

$$
\begin{aligned}
0 &= C_j \, \text{Trace } \hat{I} + \sum_{i \neq j} C_i \, a_{i,j} \, \text{Trace } \Gamma_k \\
&= C_j \, 4
\end{aligned}
\tag{1489}
$$

since the matrices Γ_k are traceless. Hence, all the C_j are zero, so the matrices are linearly independent.

Uniqueness of Expansions

The existence of sixteen linearly independent matrices require that the matrices can be represented in a space of $N \times N$ matrices, where $N \geq 4$. Any matrix A in the space of 4×4 matrices can be uniquely expressed in terms of the basis set of the Γ_i. For example, if

$$A = \sum_i C_i \, \Gamma_i \tag{1490}$$

then on multiplying by Γ_j and taking the trace, one has

$$
\begin{aligned}
\text{Trace}(\, A \, \Gamma_j \,) &= \sum_i C_i \, \text{Trace}(\, \Gamma_i \, \Gamma_j \,) \\
&= C_j \, \text{Trace}(\, \Gamma_j \, \Gamma_j \,) + \sum_{i \neq j} C_i \, \text{Trace}(\, \Gamma_i \, \Gamma_j \,) \\
&= C_j \, \text{Trace}(\, \hat{I} \,) + \sum_{i \neq j} C_i \, \text{Trace}(\, a_{i,j} \, \Gamma_k \,) \\
&= C_j \, 4
\end{aligned}
\tag{1491}
$$

Hence, the coefficients C_j in the expansion of A are uniquely determined as

$$C_j = \frac{1}{4} \, \text{Trace}(\, A \, \Gamma_j \,) \tag{1492}$$

Schur's Lemma

The uniqueness of the expansion can be used to show that the product of Γ_i for fixed i with the set of Γ_j for leads to a different Γ_k for each j. This can be shown by assuming that there exist two different (linearly independent) values Γ_j and $\Gamma_{j'}$ which lead to the same Γ_k

$$
\begin{aligned}
\Gamma_i \, \Gamma_j &= a_{i,j} \, \Gamma_k \\
\Gamma_i \, \Gamma_{j'} &= a_{i,j'} \, \Gamma_k
\end{aligned}
\tag{1493}
$$

On multiplying by Γ_i, one obtains

$$
\begin{aligned}
\Gamma_j &= a_{i,j} \, \Gamma_i \, \Gamma_k \\
\Gamma_{j'} &= a_{i,j'} \, \Gamma_i \, \Gamma_k
\end{aligned}
\tag{1494}
$$

Hence, one infers that

$$\Gamma_{j'} = \frac{a_{i,j'}}{a_{i,j}} \, \Gamma_j \tag{1495}$$

which contradicts the assumption that Γ_j and $\Gamma_{j'}$ are linearly independent. Therefore for fixed i, the product of $\Gamma_i \, \Gamma_j$ leads to a different result Γ_k for the different Γ_j.

One can also prove Schur's lemma. Schur's Lemma states that if a matrix A commutes with all the γ^μ's, then A is a multiple of the identity. If A commutes with the γ^μ's, it also commutes with all the Γ_i's. Schur's lemma follows from the expansion of A as

$$A \;=\; C_i \, \Gamma_i \;+\; \sum_{j\neq i} C_j \, \Gamma_j \tag{1496}$$

for any i such that $\Gamma_i \;\neq\; \hat{I}$. Then, one notes that there exits a Γ_k such that

$$\Gamma_k \, \Gamma_i \, \Gamma_k \;=\; -\,\Gamma_i \tag{1497}$$

Since it has been assumed that A commutes with all the Γ_i, for the specific Γ_k one has

$$
\begin{aligned}
A \;&=\; \Gamma_k \, A \, \Gamma_k \\
&=\; C_i \, \Gamma_k \, \Gamma_i \, \Gamma_k \;+\; \sum_{j\neq i} C_j \, \Gamma_k \, \Gamma_j \, \Gamma_k \\
&=\; -\, C_i \, \Gamma_i \;+\; \sum_{j\neq i} C_j \, \Gamma_k \, \Gamma_j \, \Gamma_k
\end{aligned}
\tag{1498}
$$

Furthermore, since the Γ_i matrices either commute or anti-commute

$$\Gamma_k \, \Gamma_j \, \Gamma_k \;=\; (\,\pm 1\,)_{j,k} \, \Gamma_j \tag{1499}$$

the above equation reduces to

$$A \;=\; -\, C_i \, \Gamma_i \;+\; \sum_{j\neq i} C_j \, (\,\pm 1\,)_{j,k} \, \Gamma_j \tag{1500}$$

which should be compared with the assumed form of the expansion

$$A \;=\; C_i \, \Gamma_i \;+\; \sum_{j\neq i} C_j \, \Gamma_j \tag{1501}$$

Since the expansion is unique, the coefficients of the Γ_j are unique and in particular

$$C_i \;=\; -\, C_i \tag{1502}$$

so $C_i \;=\; 0$ for any i such that $\Gamma_i \;\neq\; \hat{I}$. Hence, if A commutes with all the Γ_i then A must be proportional to the identity.

Pauli's Fundamental Theorem

Pauli's fundamental theorem states that if there are two representations of the algebra of anti-commuting γ-matrices, say $\gamma^{\mu\prime}$ and γ^{μ}, then these representations are related via a similarity transformation

$$\gamma^{\mu\prime} = \hat{S} \, \gamma^{\mu} \, \hat{S}^{-1} \tag{1503}$$

where $\hat{S}$ is a non-singular matrix.

The theorem requires that one constructs a set of sixteen matrices $\Gamma_i{}'$ from the $\gamma^{\mu\prime}$ following the same rules with which the Γ_i were constructed from γ^{μ}. Then one can describe the non-singular matrix by

$$\hat{S} = \sum_i \Gamma_i{}' \, F \, \Gamma_i \tag{1504}$$

where F is an arbitrary 4×4 matrix.

First one notes that

$$\Gamma_i \, \Gamma_j = a_{i,j} \, \Gamma_k \tag{1505}$$

so on iterating, one has

$$\Gamma_i \, \Gamma_j \, \Gamma_i \, \Gamma_j = a_{1,j}^2 \, \Gamma_k^2 = a_{i,j}^2 \, \hat{I} \tag{1506}$$

since

$$\Gamma_k^2 = \hat{I} \tag{1507}$$

On pre-multiplying eqn(1506) by $\Gamma_j \, \Gamma_i$, one obtains

$$\Gamma_j \, \Gamma_i \, \Gamma_i \, \Gamma_j \, \Gamma_i \, \Gamma_j = a_{i,j}^2 \, \Gamma_j \, \Gamma_i \tag{1508}$$

but since

$$\Gamma_j \, \Gamma_i \, \Gamma_i \, \Gamma_j = \hat{I} \tag{1509}$$

eqn(1508) reduces to

$$\Gamma_i \, \Gamma_j = a_{i,j}^2 \, \Gamma_j \, \Gamma_i \tag{1510}$$

However, as

$$\Gamma_i \, \Gamma_j = a_{i,j} \, \Gamma_k \tag{1511}$$

the equation becomes

$$a_{i,j} \, \Gamma_k = a_{i,j}^2 \, \Gamma_j \, \Gamma_i \tag{1512}$$

or since $a_{i,j}^4 = 1$, the equation can be expressed as

$$\Gamma_j \, \Gamma_i = a_{i,j}^3 \, \Gamma_k \tag{1513}$$

257

The $\Gamma_i^{'}$ matrices are constructed so that they satisfy similar relations to the Γ_i. In particular, the $\Gamma_i^{'}$ matrices satisfy

$$\Gamma_i{}' \, \Gamma_j{}' \;=\; a_{i,j} \, \Gamma_k{}' \tag{1514}$$

with the same constants $a_{i,j}$ as the unprimed matrices.

Pauli's theorem follows from the above relations by noting that

$$\Gamma_i{}' \, \hat{S} \, \Gamma_i \;=\; \Gamma_i{}' \, \Big(\sum_j \Gamma_j{}' \, F \, \Gamma_j \Big) \, \Gamma_i \tag{1515}$$

but on recalling that

$$\Gamma_j \, \Gamma_i \;=\; a_{i,j}^3 \, \Gamma_k \tag{1516}$$

and

$$\Gamma_i{}' \, \Gamma_j{}' \;=\; a_{i,j} \, \Gamma_k{}' \tag{1517}$$

one finds

$$\Gamma_i{}' \, \hat{S} \, \Gamma_i \;=\; \sum_j a_{i,j}^4 \, \Gamma_k{}' \, F \, \Gamma_k \tag{1518}$$

Therefore, with $a_{i,j}^4 \;=\; 1$, the above equation reduces to

$$\Gamma_i{}' \, \hat{S} \, \Gamma_i \;=\; \sum_j \Gamma_k{}' \, F \, \Gamma_k \tag{1519}$$

However, since i is fixed and j is being summed over, every Γ_k appears once and only once in the product. Therefore, the sum can be performed over k

$$\Gamma_i{}' \, \hat{S} \, \Gamma_i \;=\; \sum_k \Gamma_k{}' \, F \, \Gamma_k \;=\; \hat{S} \tag{1520}$$

If one can show that the matrix $\hat{S}$ has an inverse, then on post-multiplying by $\hat{S}^{-1}$, one finds

$$\Gamma_i{}' \, \hat{S} \, \Gamma_i \, \hat{S}^{-1} \;=\; \hat{I} \tag{1521}$$

Furthermore, since $\Gamma_i{}'$ is its own inverse, then on pre-multiplying by $\Gamma_i{}'$ the equation reduces to

$$\hat{S} \, \Gamma_i \, \hat{S}^{-1} \;=\; \Gamma_i{}' \tag{1522}$$

This is a generalization of the statement of the theorem. As a particular case, one may choose $\Gamma_i \;=\; \gamma^\mu$ in which case the theorem becomes

$$\gamma^{\mu\prime} \;=\; \hat{S} \, \gamma^\mu \, \hat{S}^{-1} \tag{1523}$$

which was the initial statement of Pauli's fundamental theorem made above.

The matrix $\hat{S}$ is non-singular and has an inverse. This can be shown by using Schur's Lemma. One can construct a matrix $\hat{S}'$ in a manner which is symmetrical to the construction of $\hat{S}$. That is

$$\hat{S}' = \sum_i \Gamma_i \, G \, \Gamma_i{}' \tag{1524}$$

From symmetry it follows that since eqn(1520) is given by

$$\hat{S} = \Gamma_i{}' \, \hat{S} \, \Gamma_i \tag{1525}$$

one also has

$$\hat{S}' = \Gamma_i \, \hat{S}' \, \Gamma_i{}' \tag{1526}$$

Therefore, on taking the product, one obtains

$$\begin{aligned} \hat{S}' \, \hat{S} &= \Gamma_i \, \hat{S}' \, \Gamma_i{}' \, \Gamma_i{}' \, \hat{S} \, \Gamma_i \\ &= \Gamma_i \, \hat{S}' \, \hat{S} \, \Gamma_i \end{aligned} \tag{1527}$$

Hence, by Schur's Lemma one sees that $\hat{S}' \, \hat{S}$ commutes with all the matrices in the space, therefore it must be a multiple of the identity

$$\hat{S}' \, \hat{S} = \kappa \, \hat{I} \tag{1528}$$

where κ is a constant. By a judicious choice of the magnitude of the elements of F, the constant κ can be set to unity, yielding

$$\hat{S}' \, \hat{S} = \hat{I} \tag{1529}$$

Thus, $\hat{S}$ is non-singular so the inverse exists and is given by $\hat{S}^{-1} = \hat{S}'$.

12.5.2 Polarization in Mott Scattering

When evaluated in the Born Approximation, Mott scattering does not result in the polarization of an unpolarized beam. However, when higher-order corrections are included, Mott scattering produces a partially polarization of the scattered electrons[101]. If the incident beam is polarized by having a definite helicity, it is expected that the helicity may change as a result of the scattering.

The probability of non-helicity flip scattering and helicity flip scattering can be evaluated using the Born approximation. The initial beam will be considered as having a momentum $\underline{p}$ parallel to the $\hat{e}_3$-axis and as having a helicity of $+1$. The initial spinor is proportional to

$$\psi_{\underline{p},+}(\underline{r}) = \sqrt{\frac{E_p + m \, c^2}{2 \, E_p \, V}} \left(\begin{array}{c} \chi_+ \\ \frac{c \, p}{E_p + m \, c^2} \, \chi_+ \end{array} \right) \exp\left[i \, \frac{\underline{p} \cdot \underline{r}}{\hbar} \right] \tag{1530}$$

---

[101] N. F. Mott, Proc. Roy. Soc. A **124**, 425 (1929).
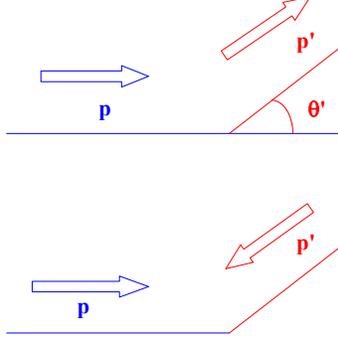
Figure 47: Helcity non-flip and helicity flip Mott scattering of an electron with helicity $+1$. The scattering angle is $\theta_{p'}$.

The electrons are assumed to be elastically scattered to a state with final momentum $\underline{p}'$. The scattering is defined to occur through an angle $\theta_{p'}$ in the $z - x$ plane. The final state is composed of a linear-superposition of states with different helicities. Since the final state helicities are specified relative to the final momentum, the final state helicity eigenstates can be obtained by rotating the initial state helicity eigenstates through an angle $\theta_{p'}$ around the $\hat{e}_2$-axis

$$
\begin{aligned}
\psi'_{\underline{p}',\Lambda'}(\underline{r}) &= \hat{\mathcal{R}}(\theta_{p'}) \, \psi_{\underline{p}',\Lambda'}(x) \\
&= \sqrt{\frac{E_p + m \, c^2}{2 \, E_p \, V}} \, \hat{\mathcal{R}}(\theta_{p'}) \left(\begin{array}{c} \chi_{\Lambda'} \\ \frac{c \, p \, \Lambda'}{E_p + m \, c^2} \, \chi_{\Lambda'} \end{array} \right) \exp \left[\, i \, \frac{\underline{p}' \cdot \underline{r}}{\hbar} \, \right]
\end{aligned}
$$

(1531)

where the rotation operator is given by

$$
\hat{\mathcal{R}}(\theta_{p'}) = \left[\cos \frac{\theta_{p'}}{2} \, \hat{I} - i \, \sin \frac{\theta_{p'}}{2} \, \hat{\sigma}^{(2)} \right]
$$

(1532)

which does not mix the upper and lower two-component spinors. Therefore, one finds that the final state two-component spinors representing helicity eigenstates are given by

$$
\begin{aligned}
\chi'_+ &= \left[\cos \frac{\theta_{p'}}{2} \, I - i \, \sin \frac{\theta_{p'}}{2} \, \sigma^{(2)} \right] \chi_+ \\
&= \left(\begin{array}{c} \cos \frac{\theta_{p'}}{2} \\ \sin \frac{\theta_{p'}}{2} \end{array} \right)
\end{aligned}
$$

(1533)

and

$$
\chi'_- = \left[\cos \frac{\theta_{p'}}{2} \, I - i \, \sin \frac{\theta_{p'}}{2} \, \sigma^{(2)} \right] \chi_-
$$

$$
= \begin{pmatrix} -\sin\frac{\theta_{p'}}{2} \\ \cos\frac{\theta_{p'}}{2} \end{pmatrix} \tag{1534}
$$

Therefore, the final state basis states are given by

$$
\psi'_{\underline{p}',\Lambda'}(\underline{r}) = \sqrt{\frac{E_p + m\,c^2}{2\,E_p\,V}} \begin{pmatrix} \chi'_{\Lambda'} \\ \frac{c\,p\,\Lambda'}{E_p + m\,c^2}\,\chi'_{\Lambda'} \end{pmatrix} \exp\left[\,i\,\frac{\underline{p}'\cdot\underline{r}}{\hbar}\,\right] \tag{1535}
$$

The Born approximation scattering cross-section can be expressed in terms of the modulus squared matrix elements

$$
\left| \int d^3\underline{r}\; \overline{\psi}'^{\dagger}_{\underline{p}',\Lambda'}(\underline{r})\; \gamma^{(0)}\; \frac{Z\,e^2}{|\,\underline{r}\,|}\; \psi_{\underline{p},+}(\underline{r}) \right|^2 \tag{1536}
$$

which is evaluated as

$$
\left(\frac{4\,\pi\,Z\,e^2}{V\,|\,\underline{p}\,-\,\underline{p}'\,|^2}\right)^2 \left(1\,+\,\Lambda'\,\frac{c^2\,p^2}{(\,E_p\,+\,m\,c^2\,)^2}\right)^2 \left(\frac{E_p\,+\,m\,c^2}{2\,E_p}\right)^2 \left|\chi'^{\dagger}_{\Lambda'}\,\chi_+\right|^2
$$

$$
= \left(\frac{4\,\pi\,Z\,e^2}{V\,|\,\underline{p}\,-\,\underline{p}'\,|^2}\right)^2 \left(\frac{(\,E_p\,+\,m\,c^2\,)^2\,+\,\Lambda'\,c^2\,p^2}{2\,E_p\,(\,E_p\,+\,m\,c^2\,)}\right)^2 \left|\chi'^{\dagger}_{\Lambda'}\,\chi_+\right|^2 \tag{1537}
$$

Therefore, the probability for non-helicity flips scattering is proportional to

$$
\propto \left(\frac{4\,\pi\,Z\,e^2}{V\,|\,\underline{p}\,-\,\underline{p}'\,|^2}\right)^2 \cos^2\frac{\theta_{p'}}{2} \tag{1538}
$$

whereas the probability for helicity flip scattering is given by

$$
\propto \left(\frac{4\,\pi\,Z\,e^2}{V\,|\,\underline{p}\,-\,\underline{p}'\,|^2}\right)^2 \left(\frac{m\,c^2}{E_p}\right)^2 \sin^2\frac{\theta_{p'}}{2} \tag{1539}
$$

It is seen that the probability for helicity flip scattering vanishes in the ultra-relativistic limit. Also, in the non-relativistic limit, a static charge cannot flip the spin. Therefore, in the non-relativistic limit, if one expresses the spin eigenstate as a linear superposition of the final helicity eigenstates

$$
\chi_+ = \cos\frac{\theta_{p'}}{2}\,\chi'_+ \,-\, \sin\frac{\theta_{p'}}{2}\,\chi'_- \tag{1540}
$$

one is lead to expect that the relative probability of helicity flip to non-helicity flip will be governed by a factor of

$$
\tan^2\frac{\theta_{p'}}{2} \tag{1541}
$$

which agrees with the above matrix elements evaluated in the non-relativistic limit. The cross-section for non-flip scattering is determined as

$$
\left(\frac{d\sigma}{d\Omega'}\right)_{+,+} = \left(\frac{2\,Z\,e^2\,E_p}{4\,c^2\,p^2\,\sin^2\frac{\theta_{p'}}{2}}\right)^2 \cos^2\frac{\theta_{p'}}{2} \tag{1542}
$$

whereas the cross-section for spin flip scattering is given by

$$
\left(\frac{d\sigma}{d\Omega'} \right)_{+,-} = \left(\frac{2\, Z\, e^2\, m\, c^2}{4\, c^2\, p2\, \sin^2 \frac{\theta_{p'}}{2}} \right)^2 \sin^2 \frac{\theta_{p'}}{2}
\tag{1543}
$$

The Born approximation to the total cross-section for scattering of polarized electrons, in which the final polarization is not measured, is given by

$$
\begin{aligned}
\left(\frac{d\sigma}{d\Omega'} \right)_{+} &= \left(\frac{d\sigma}{d\Omega'} \right)_{+,+} + \left(\frac{d\sigma}{d\Omega'} \right)_{+,-} \\
&= \left(\frac{2\, Z\, e^2\, E_p}{4\, c^2\, p2\, \sin^2 \frac{\theta_{p'}}{2}} \right)^2 \left[\cos^2 \frac{\theta_{p'}}{2} + \left(\frac{m\, c^2}{E_p} \right)^2 \sin^2 \frac{\theta_{p'}}{2} \right] \\
&= \left(\frac{2\, Z\, e^2\, E_p}{4\, c^2\, p2\, \sin^2 \frac{\theta_{p'}}{2}} \right)^2 \left[1 - \left(\frac{p\, c}{E_p} \right)^2 \sin^2 \frac{\theta_{p'}}{2} \right]
\end{aligned}
\tag{1544}
$$

which is the same as the cross-section as calculated for unpolarized electrons. The degree of polarization of the scattered beam is given

$$
P(\theta_{p'}) = \left(\frac{E_p^2\, \cos^2 \frac{\theta_{p'}}{2} - m^2\, c^4\, \sin^2 \frac{\theta_{p'}}{2}}{E_p^2\, \cos^2 \frac{\theta_{p'}}{2} + m^2\, c^4\, \sin^2 \frac{\theta_{p'}}{2}} \right)
\tag{1545}
$$

If the initial beam of electrons is unpolarized, the scattered electrons would be observed to be partially polarized, where the net polarization is in the plane perpendicular to the scattering plane. However, the polarization is due to processes of higher-order than the Born approximation and is governed by the factor $\left(\frac{Z\, e^2}{\hbar\, c} \right)$.

12.6 The Non-Relativistic Limit

The non-relativistic limit of the Dirac equation should reduce to the Schrödinger equation. As shall be seen, the appropriate Schrödinger equation for a particle with positive-energy is modified due to the existence of spin. The non-relativistic limit is described by the Pauli equation[102].

The Dirac equation can be written as

$$
\left(i\, \frac{\hbar}{c}\, \frac{\partial}{\partial t} - \frac{q}{c}\, A_0 \right) \psi = \left[\underline{\alpha} \cdot (\underline{\hat{p}} - \frac{q}{c}\, \underline{A}) + \beta\, m\, c \right] \psi
\tag{1546}
$$

The equation can be written in 2×2 block diagonal form, if the wave function is expressed in the form of two two-component spinors. We shall mainly focus on the positive-energy solutions and recognize that, in the non-relativistic limit, the largest component of the wave function is ϕ^A and the largest term in the

---

[102]W. Pauli, Z, Phys. **44**, 601 (1927).

energy is the rest mass energy $m\,c^2$. Therefore, the spinor wave function will be expressed as

$$\psi \;=\; \begin{pmatrix} \phi^A \\ \phi^B \end{pmatrix} \, \exp\left[\, -\, i\, \frac{m\,c^2}{\hbar}\, t \, \right] \tag{1547}$$

The above form explicitly displays the rest-mass energy of the positive-energy solution of the Dirac equation. Hence, the Dirac equation takes the form

$$\left(i\,\hbar\, \frac{\partial}{\partial t} \;-\; q\,A_0 \right) \begin{pmatrix} \phi^A \\ \phi^B \end{pmatrix} \;=\; \begin{pmatrix} c\,\underline{\sigma}\cdot(\,\hat{\underline{p}}\,-\,\frac{q}{c}\,\underline{A}\,)\,\phi^B \\ c\,\underline{\sigma}\cdot(\,\hat{\underline{p}}\,-\,\frac{q}{c}\,\underline{A}\,)\,\phi^A \end{pmatrix}$$
$$-\; 2\,m\,c^2 \begin{pmatrix} 0 \\ \phi^B \end{pmatrix} \tag{1548}$$

where the rest mass has been eliminated from the equation for the large component ϕ^A of the positive-energy solution. Since the kinetic energy and the potential energy are assumed to be smaller than the rest mass energy, the equation for the small component

$$\left(i\,\hbar\, \frac{\partial}{\partial t} \;-\; q\,A_0 \right) \phi^B \;=\; c\,\underline{\sigma}\cdot\left(\hat{\underline{p}} \;-\; \frac{q}{c}\,\underline{A} \right) \phi^A \;-\; 2\,m\,c^2\,\phi^B \tag{1549}$$

can be expressed as

$$\phi^B \;=\; \frac{1}{2\,m\,c}\, \underline{\sigma}\cdot\left(\hat{\underline{p}} \;-\; \frac{q}{c}\,\underline{A} \right) \phi^A \tag{1550}$$

Substituting the expression for the small component into the equation for the large component, hence eliminating ϕ^B, one finds the equation

$$\left(i\,\hbar\, \frac{\partial}{\partial t} \;-\; q\,A_0 \right) \phi^A \;=\; \frac{1}{2\,m}\left(\underline{\sigma}\cdot\left(\hat{\underline{p}} \;-\; \frac{q}{c}\,\underline{A} \right) \right)^2 \phi^A \tag{1551}$$

which is the Pauli equation. The equation can be simplified by expanding the terms involving the Pauli spin matrices. The Pauli identity can be used to obtain

$$\left(\underline{\sigma}\cdot\left(\hat{\underline{p}} \;-\; \frac{q}{c}\,\underline{A} \right) \right)^2 \;=\; I\left(\hat{\underline{p}} \;-\; \frac{q}{c}\,\underline{A} \right)^2 \;+\; i\,\underline{\sigma}\cdot\left(\left(\hat{\underline{p}} \;-\; \frac{q}{c}\,\underline{A} \right) \wedge \left(\hat{\underline{p}} \;-\; \frac{q}{c}\,\underline{A} \right) \right)$$
$$=\; I\left(\hat{\underline{p}} \;-\; \frac{q}{c}\,\underline{A} \right)^2 \;-\; \frac{q\,\hbar}{c}\,\underline{\sigma}\cdot\left(\underline{\nabla} \wedge \underline{A} \right) \tag{1552}$$

where the last term originates from the non-commutativity of the components of $\hat{\underline{p}}$ and $\underline{A}$. Since the magnetic field $\underline{B}$ is given by

$$\underline{B} \;=\; \underline{\nabla} \wedge \underline{A} \tag{1553}$$

the Pauli equation can be expressed as

$$i\,\hbar\, \frac{\partial}{\partial t}\, \phi^A \;=\; \frac{1}{2\,m}\left(\hat{\underline{p}} \;-\; \frac{q}{c}\,\underline{A} \right)^2 \phi^A \;+\; q\,A_0\,\phi^A \;-\; \frac{q\,\hbar}{2\,m\,c}\,\underline{\sigma}\cdot\underline{B}\,\phi^A \tag{1554}$$

263

The Pauli equation[103] is the non-relativistic limit of the Dirac equation. It represents the Schrödinger equation for a charged particle with spin one-half. The two components of the spinor ϕ^A in the Pauli equation represent the internal spin of the electron. The last term represents the anomalous Zeeman interaction between the magnetic field and the electron's spin.

The other contribution to the Zeeman interaction originates with the electron's orbital angular momentum $\underline{L}$. The ordinary Zeeman interaction occurs between the constant magnetic field $\underline{B}$ and the orbital angular momentum and originates from the gauge-invariant term in the Hamiltonian

$$\frac{1}{2\,m}\left(\hat{\underline{p}} - \frac{q}{c}\,\underline{A}\right)^2 = \frac{1}{2\,m}\left(\hat{\underline{p}} - \frac{q}{2\,c}\,\underline{B}\wedge\underline{r}\right)^2 \tag{1555}$$

where the vector potential has been expressed in terms of the uniform magnetic field via

$$\underline{A} = \frac{1}{2}\,\underline{B}\wedge\underline{r} \tag{1556}$$

The expression for the energy term can be further simplified to

$$\begin{aligned}
\frac{1}{2\,m}\left(\hat{\underline{p}} - \frac{q}{c}\,\underline{A}\right)^2 &= \frac{\hat{p}^2}{2\,m} - \frac{q}{4\,m\,c}\left(\hat{\underline{p}}\cdot(\underline{B}\wedge\underline{r}) + (\underline{B}\wedge\underline{r})\cdot\hat{\underline{p}}\right) \\
&\quad + \frac{q^2}{2\,m\,c^2}\,\underline{A}^2 \\
&= \frac{\hat{p}^2}{2\,m} - \frac{q}{2\,m\,c}\left((\underline{B}\wedge\underline{r})\cdot\hat{\underline{p}}\right) + \frac{q^2}{2\,m\,c^2}\,\underline{A}^2 \\
&= \frac{\hat{p}^2}{2\,m} - \frac{q}{2\,m\,c}\left(\underline{B}\cdot(\underline{r}\wedge\hat{\underline{p}})\right) + \frac{q^2}{2\,m\,c^2}\,\underline{A}^2 \\
&= \frac{\hat{p}^2}{2\,m} - \frac{q}{2\,m\,c}\left(\underline{B}\cdot\hat{\underline{L}}\right) + \frac{q^2}{2\,m\,c^2}\,\underline{A}^2
\end{aligned} \tag{1557}$$

In obtaining the second line, the i-th component of $\hat{\underline{p}}$ has been commuted with the i-th component of ($\underline{B}\wedge\underline{r}$). In obtaining the third line, the (cyclic) vector identity

$$(\underline{A}\wedge\underline{B})\cdot\underline{C} = (\underline{B}\wedge\underline{C})\cdot\underline{A} \tag{1558}$$

has been used. The first term in eqn(1557) represents the usual non-relativistic expression for the kinetic energy of the electrons, the second term represents the ordinary Zeeman interaction which originates from the paramagnetic interaction. The last term represents the diamagnetic interaction.

The total Zeeman interaction is the energy of the total magnetic moment M in the field $\underline{B}$

$$\hat{H}_{Zeeman} = -\underline{M}\cdot\underline{B} \tag{1559}$$

---

[103] W. Pauli, Z, Phys. **44**, 601 (1927).

The Dirac equation results in the Zeeman interaction of the form

$$
\begin{aligned}
\hat{H}_{Zeeman} &= -\frac{q}{2\,m\,c}\,\underline{B}\cdot\left(\hat{\underline{L}} + \hbar\,\underline{\sigma}\right) \\
&= -\frac{q}{2\,m\,c}\,\underline{B}\cdot\left(\hat{\underline{L}} + 2\,\underline{S}\right)
\end{aligned}
\tag{1560}
$$

where the spin angular momentum $\underline{S}$ has been identified as

$$
\underline{S} = \frac{\hbar}{2}\,\underline{\sigma}
\tag{1561}
$$

It is seen that both the spin angular momentum and the orbital angular momentum of the charged particle interacts with the magnetic field, therefore, both contribute to the magnetic moment. However, it is noted that the magnetic moment can be written in the form

$$
\underline{M} = \frac{q}{2\,m\,c}\left(\hat{\underline{L}} + g\,\underline{S}\right)
\tag{1562}
$$

where the magnitude of the magnetic moment is determined by the factor $\frac{q\,\hbar}{2\,m\,c}$ which is the Bohr magneton. The Dirac equation shows that the spin angular momentum couples with a different strength to orbital angular momentum, and the relative coupling strength g (the gyromagnetic ratio) is given by $g = 2$. The existence of spin and the value of 2 for the gyromagnetic ratio were the first successes of Dirac's theory. Quantum Electrodynamics[104] yields a small correction to the gyromagnetic ratio of

$$
g = 2\left(1 + \frac{1}{2\,\pi}\left(\frac{e^2}{\hbar\,c}\right) + \dots\right)
\tag{1563}
$$

which has been experimentally verified to incredible precision[105]. Using the features associated with spin, Dirac's theory correctly described the fine structure of the Hydrogen atom. The second success of the Dirac equation followed Dirac's physical interpretation of the negative-energy states in terms of anti-particles[106]. The second round of success came with the discovery of the positron by Anderson[107].

Exercise:

The Dirac equation can be phenomenologically modified to describe particles with anomalous magnetic moments. The Dirac equation is modified to

$$
\left[i\,\hbar\,\gamma^{\mu}\left(\partial_{\mu} + i\,\frac{q}{\hbar\,c}\,A_{\mu}\right) + \kappa\,\frac{q\,\hbar}{4\,m\,c^2}\,\sigma^{\mu,\nu}\,F_{\mu,\nu} - m\,c\,\hat{I}\right]\psi = 0
\tag{1564}
$$

[104] J. S. Schwinger, Phys. Rev. **73**, 416 (1948).

[105] H. M. Foley and P. Kusch, Phys. Rev. **73**, 412 (1948).
R. S. Van Dyck Jr., P. B. Schwinberg and H. G. Dehmelt, Phys. Rev. Lett. **59**, 26 (1987).

[106] P. A. M. Dirac, Proc. Roy. Soc. **A 126**, 360 (1930).

[107] C. D. Anderson, Phys. Rev. **43**, 491 (1933).

Show that the modified equation is Lorentz covariant and that the Hamiltonian is Hermitean. Also derive the corrections to the magnetic moment due to the spin by examining the non-relativistic limit.

12.7   Conservation of Angular Momentum

The law of conservation of angular momentum will now be examined. For a relativistic electron the orbital angular momentum and the spin angular momentum are not separately conserved. However, the total angular momentum which is the sum of the orbital angular momentum and spin angular momentum is conserved.

The orbital angular momentum $\hat{\underline{L}}$ defined by

$$\hat{\underline{L}} = \underline{r} \wedge \hat{\underline{p}} \tag{1565}$$

is not conserved for a spherically symmetric potential. The Dirac Hamiltonian is given by

$$\hat{H} = c\,\underline{\alpha} \cdot \hat{\underline{p}} + \beta\,m\,c^2 + \hat{I}\,V(r) \tag{1566}$$

The matrices shall be expressed in a 2×2 block diagonal form. Therefore, the identity matrix is written as

$$\hat{I} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{1567}$$

and

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{1568}$$

Finally, the α matrices are of off-diagonal form

$$\begin{aligned} \underline{\alpha} &= \begin{pmatrix} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \hat{\underline{\sigma}} \end{aligned} \tag{1569}$$

where $\hat{\underline{\sigma}}$ is the 2×2 block-diagonal Pauli spin matrix. The rate of change of orbital angular momentum is given by the Heisenberg equation of motion

$$i\,\hbar\,\frac{\partial}{\partial t}\,\hat{\underline{L}} = [\,\hat{\underline{L}}\,,\,\hat{H}\,] \tag{1570}$$

The orbital angular momentum operator commutes with the mass term and with the spherically symmetric potential $V(r)$. The orbital angular momentum does not commute with the momentum. Thus,

$$i\,\hbar\,\frac{\partial}{\partial t}\,\hat{\underline{L}} = c\,[\,\hat{\underline{L}}\,,\,\underline{\alpha} \cdot \hat{\underline{p}}\,] \tag{1571}$$

266

Hence, the Heisenberg equation of motion can be expressed in the form

$$i \, \hbar \, \frac{\partial}{\partial t} \, \underline{\hat{L}} \; = \; c \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \, [\, \underline{\hat{L}} \, , \, \underline{\hat{\sigma}} \cdot \underline{\hat{p}} \,]$$
$$= \; - \, c \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \, \underline{\hat{\sigma}} \cdot [\, \underline{\hat{p}} \, , \, \underline{\hat{L}} \,] \tag{1572}$$

However, the components of the orbital angular momentum $\hat{L}^{(i)}$ and momenta $p^{(j)}$ satisfy the commutation relations

$$[\, \hat{L}^{(i)} \, , \, p^{(j)} \,] \; = \; i \, \hbar \, \sum_k \xi^{i,j,k} \, p^{(k)} \tag{1573}$$

Therefore, one finds

$$i \, \hbar \, \frac{\partial}{\partial t} \, \underline{\hat{L}} \; = \; i \, \hbar \, c \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \, (\, \underline{\hat{\sigma}} \wedge \underline{\hat{p}} \,) \tag{1574}$$

which shows that orbital angular momentum is not conserved for a relativistic electron with a central potential.

The spin angular momentum is also not conserved. This can be seen by examining the Heisenberg equation of motion for the Pauli spin operator

$$i \, \hbar \, \frac{\partial}{\partial t} \, \underline{\hat{\sigma}} \; = \; [\, \underline{\hat{\sigma}} \, , \, \hat{H} \,] \tag{1575}$$

The spin operator commutes with $\hat{I}$ and β but does not commute with the α matrices. Hence,

$$i \, \hbar \, \frac{\partial}{\partial t} \, \underline{\hat{\sigma}} \; = \; c \, [\, \underline{\hat{\sigma}} \, , \, \underline{\alpha} \cdot \underline{\hat{p}} \,]$$
$$= \; c \, [\, \underline{\hat{\sigma}} \, , \, \underline{\hat{\sigma}} \,] \cdot \underline{\hat{p}} \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{1576}$$

The components of the Pauli spin operators satisfy the commutation relations

$$[\, \sigma^{(i)} \, , \, \sigma^{(j)} \,] \; = \; 2 \, i \, \sum_k \xi^{i,j,k} \, \sigma^{(k)} \tag{1577}$$

which, clearly, have a similar form to the commutation relations for the orbital angular momentum. Hence, spin angular momentum is not conserved since

$$i \, \hbar \, \frac{\partial}{\partial t} \, \underline{\hat{\sigma}} \; = \; - \, 2 \, i \, c \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \, (\, \underline{\hat{\sigma}} \wedge \underline{\hat{p}} \,) \tag{1578}$$

The total angular momentum $\underline{\hat{J}}$ is defined via

$$
\begin{aligned}
\underline{\hat{J}} &= \underline{\hat{L}} + \underline{\hat{S}} \\
&= \underline{\hat{L}} + \frac{\hbar}{2}\,\underline{\hat{\sigma}}
\end{aligned}
\tag{1579}
$$

The total angular momentum is conserved since

$$
\begin{aligned}
i\,\hbar\,\frac{\partial}{\partial t}\,\underline{\hat{J}} &= i\,\hbar\,\frac{\partial}{\partial t}\,\underline{\hat{L}} + i\,\hbar\,\frac{\partial}{\partial t}\,\underline{\hat{S}} \\
&= i\,\hbar\,c\left[\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}(\underline{\hat{\sigma}} \wedge \underline{\hat{p}}) - \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}(\underline{\hat{\sigma}} \wedge \underline{\hat{p}})\right] \\
&= 0
\end{aligned}
\tag{1580}
$$

which follows from combining eqn(1574) and eqn(1578). This confirms the interpretation of the quantity $\underline{\hat{S}}$ defined by

$$
\underline{\hat{S}} = \frac{\hbar}{2}\,\underline{\hat{\sigma}}
\tag{1581}
$$

as the spin angular momentum of the electron.

12.8   Conservation of Parity

Dirac was very conscious that his book "Principles of Quantum Mechanics" never contained any mention of parity. It seems that he had questioned the requirement of parity invariance[108] since biological systems are not parity invariant. Dirac's viewpoint was vindicated by the discovery that the weak interaction violates parity.

The parity transform $\mathcal{P}$ acting on the coordinates $(t,\underline{r})$ has the effect

$$
\mathcal{P}\,(t,\underline{r}) \rightarrow (t',\underline{r}') = (t,-\underline{r})
\tag{1582}
$$

which is an inversion of the spatial coordinates. Thus, the parity reverse the space-like components of vectors, so the effects of the parity operation on the position and momentum vectors are given by

$$
\begin{aligned}
\hat{\mathcal{P}}\,\underline{r}\,\hat{\mathcal{P}}^{-1} &= -\,\underline{r} \\
\hat{\mathcal{P}}\,\underline{p}\,\hat{\mathcal{P}}^{-1} &= -\,\underline{p}
\end{aligned}
\tag{1583}
$$

However, the effect of the parity transform on pseudo-vectors such as orbital angular momentum $\underline{L} = \underline{r} \wedge \underline{p}$ is such that

$$
\hat{\mathcal{P}}\,\underline{L}\,\hat{\mathcal{P}}^{-1} = \underline{L}
\tag{1584}
$$

---

[108]The question of parity conservation in weak interactions was raised subsequently by T. D. Lee and C. N. Yang [T. D. Lee and C. N. Yang, Phys. Rev. **104**, 254 (1956).]

which is unchanged. This implies that spin angular momentum should also be invariant under the parity transform

$$\hat{\mathcal{P}} \, \underline{\sigma} \, \hat{\mathcal{P}}^{-1} \;\; = \;\; \underline{\sigma} \tag{1585}$$

If the Hamiltonian $\hat{H}$ is invariant under a parity transform, one requires that

$$\hat{H} \;\; = \;\; \hat{\mathcal{P}} \, \hat{H} \, \hat{\mathcal{P}}^{-1} \tag{1586}$$

Imposing parity invariance of the Dirac Hamiltonian

$$\hat{H} \;\; = \;\; c \, \underline{\alpha} \cdot \hat{\underline{p}} \; + \; \beta \, m \, c^2 \; + \; \hat{I} \, V(\underline{r}) \tag{1587}$$

yields a condition on the potential

$$V(\underline{r}) \;\; = \;\; V(-\underline{r}) \tag{1588}$$

and also to conditions on the Dirac matrices

$$\begin{aligned} \hat{\mathcal{P}} \, \alpha \, \hat{\mathcal{P}}^{-1} \;\; &= \;\; - \, \alpha \\ \hat{\mathcal{P}} \, \beta \, \hat{\mathcal{P}}^{-1} \;\; &= \;\; \beta \end{aligned} \tag{1589}$$

The condition on the potential is the familiar condition for parity invariance in classical mechanics. In the standard representation, in 2×2 block diagonal form, the requirement of parity invariance on the Dirac matrices become the matrix equations

$$\begin{aligned} \hat{\mathcal{P}} \begin{pmatrix} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{pmatrix} \hat{\mathcal{P}}^{-1} \;\; &= \;\; - \begin{pmatrix} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{pmatrix} \\ \hat{\mathcal{P}} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \hat{\mathcal{P}}^{-1} \;\; &= \;\; \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \end{aligned} \tag{1590}$$

The above equation shows that, in the standard representation, the parity operator can be uniquely factorized as

$$\hat{\mathcal{P}} \;\; = \;\; \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \hat{P} \tag{1591}$$

where the operator $\hat{P}$ only acts on the coordinates $\underline{r}$. The presence of the matrix in the parity operation on the Dirac spinor should be compared with the effect of the parity operator on the four-vector potential of Electrodynamics $A^{\mu}(\underline{r})$ which is given by the product of spatial inversion and a matrix operation

$$\begin{aligned} \hat{\mathcal{P}} \, A^{\mu}(\underline{r}, t) \;\; &= \;\; \gamma^{\mu}{}_{\nu} \, \hat{P} \, A^{\nu}(\underline{r}, t) \\ &= \;\; \gamma^{\mu}{}_{\nu} \, A^{\nu}(-\underline{r}, t) \end{aligned} \tag{1592}$$

where the matrix $\gamma^{\mu}{}_{\nu}$ given by

$$\gamma^{\mu}{}_{\nu} \;\; = \;\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1593}$$

reverses the direction of the spatial components of the vector field.

The effect of the parity operator on the Dirac four-component spinor wave function can be computed from

$$
\begin{aligned}
\hat{\mathcal{P}} \, \psi(t, \underline{r}) & = \hat{\mathcal{P}} \begin{pmatrix} \phi^A(t, \underline{r}) \\ \phi^B(t, \underline{r}) \end{pmatrix} \\
& = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} \phi^A(t, -\underline{r}) \\ \phi^B(t, -\underline{r}) \end{pmatrix} \\
& = \begin{pmatrix} \phi^A(t, -\underline{r}) \\ -\phi^B(t, -\underline{r}) \end{pmatrix}
\end{aligned}
\tag{1594}
$$

Hence, in the standard representation, the parity operator changes the relative sign of the two two-component spinors. Due to the presence of the term $-I$ in the lower diagonal block of the parity matrix, the lower two-component spinor ϕ^B in the Dirac spinor is said to have a negative intrinsic parity.

The parity eigenstates satisfy the eigenvalue equation

$$
\hat{\mathcal{P}} \, \psi = \eta_p \, \psi
\tag{1595}
$$

with eigenvalues $\eta_p = \pm 1$, since $\hat{\mathcal{P}}^2 = \hat{I}$. The application of the parity operator on the Dirac spinor leads to the equation

$$
\begin{pmatrix} \hat{P} \, \phi^A \\ -\hat{P} \, \phi^B \end{pmatrix} = \begin{pmatrix} \eta_p \, \phi^A \\ \eta_p \, \phi^B \end{pmatrix}
\tag{1596}
$$

Hence, the two-component spinors $\phi^A(\underline{r})$ and $\phi^B(\underline{r})$ have opposite parities under spatial inversion

$$
\begin{aligned}
\hat{P} \, \phi^A(\underline{r}) & = \eta_p \, \phi^A(\underline{r}) \\
\hat{P} \, \phi^B(\underline{r}) & = -\eta_p \, \phi^B(\underline{r})
\end{aligned}
\tag{1597}
$$

In polar coordinates, the spatial part of the parity operation $\hat{P}$ is equivalent to a reflection

$$
\theta \rightarrow \pi - \theta
\tag{1598}
$$

followed by a rotation

$$
\varphi \rightarrow \varphi + \pi
\tag{1599}
$$

which has the effect that

$$
\begin{aligned}
\sin\theta & \rightarrow \sin\theta \\
\cos\theta & \rightarrow -\cos\theta \\
\exp\left[i \, m \, \varphi \right] & \rightarrow (-1)^m \, \exp\left[i \, m \, \varphi \right]
\end{aligned}
\tag{1600}
$$

Hence, the spherical harmonics with $m = l$

$$Y_l^l(\theta, \varphi) \;=\; \frac{(-1)^l}{2^l \, l!} \, \sqrt{\frac{2\,l\,+\,1}{4\,\pi}} \; \sin^l \theta \; \exp\left[i\,l\,\varphi \right] \tag{1601}$$

are eigenstates of the parity operator and have parity eigenvalues of $(-1)^l$. The lowering operator $\hat{L}^-$, defined via

$$\hat{L}^- \;=\; -\,\hbar\,\exp\left[-\,i\,\varphi \right] \left(\frac{\partial}{\partial \theta} \,-\, i\,\cot\theta\,\frac{\partial}{\partial \varphi} \right) \tag{1602}$$

is invariant under the parity transformation

$$\hat{\mathcal{P}}\,\hat{L}^-\,\hat{\mathcal{P}}^{-1} \;=\; \hat{L}^- \tag{1603}$$

Therefore, on repeatedly operating on $Y_l^l(\theta, \varphi)$ with the lowering operator $\hat{L}^-$ $(l - m)$ times, one finds that under the parity transformation

$$Y_m^l(\theta, \varphi) \;\rightarrow\; (-1)^l\,Y_m^l(\theta, \varphi) \tag{1604}$$

which shows that all states with a definite magnitude of the orbital angular momentum l are eigenstates of the parity operator and have the same eigenvalue.

Exercise:

Show that under a parity transformation the positive-energy solution for the free Dirac particle $\psi_{\underline{k},\sigma}^+(x)$ transforms as

$$\hat{\mathcal{P}}\,\psi_{\underline{k},\sigma}^+(x) \;=\; \psi_{-\underline{k},\sigma}^+(x) \tag{1605}$$

while the negative-energy solutions $\psi_{\underline{k},\sigma}^-(x)$ transform as

$$\hat{\mathcal{P}}\,\psi_{\underline{k},\sigma}^-(x) \;=\; -\,\psi_{-\underline{k},\sigma}^-(x) \tag{1606}$$

Hence, the parity operation reverses the momentum and keeps the spin invariant for the positive-energy and negative-energy solutions solution. The extra negative sign implies that the negative-energy solution has opposite intrinsic parity to the positive-energy solution.

Exercise:

Consider the parity transform as an example of an improper Lorentz transformation Λ, for which $\det |\,\Lambda\,| \;=\; -\,1$. If the Lorentz transform is given by

$$x^{\mu\prime} \;=\; \Lambda^\mu{}_\nu\,x^\nu \tag{1607}$$

the spinor wave function transforms via

$$\psi'(x') \;=\; \hat{\mathcal{R}}(\Lambda)\,\psi(x) \tag{1608}$$

271

where $\hat{\mathcal{R}}(\Lambda)$ "rotates" the spinor. The covariant condition for the Dirac equation is

$$\hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^\mu \, \hat{\mathcal{R}}(\Lambda) \;=\; \Lambda^\mu{}_\nu \, \gamma^\nu \tag{1609}$$

For a parity transformation, one has

$$x^{\mu\prime} \;=\; x_\mu \tag{1610}$$

since the spatial components of x^μ change sign. Hence, for a parity transformation, the transformation matrix is determined as

$$\Lambda^\mu{}_\nu \;=\; g_{\mu,\nu} \tag{1611}$$

which is an improper Lorentz transformation since

$$\det | \, g \, | \;=\; -1 \tag{1612}$$

Therefore, the covariant condition reduces to

$$\hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^\mu \, \hat{\mathcal{R}}(\Lambda) \;=\; g_{\mu,\nu} \, \gamma^\nu \tag{1613}$$

Solve for the matrix $\hat{\mathcal{R}}(\Lambda)$ which shuffles the components of the Dirac spinor.

12.9   Bi-linear Covariants

Under an Lorentz transformation

$$x^{\mu\prime} \;=\; \Lambda^\mu{}_\nu \, x^\nu \tag{1614}$$

(where $\Lambda^0{}_0 \; > \; 0$ for an orthochronous transformation), the Dirac spinor ψ transforms according to

$$\psi'(x') \;=\; \hat{\mathcal{R}}(\Lambda) \, \psi(x) \tag{1615}$$

and the condition that the Dirac equation is covariant under the orthochronous Lorentz transformation is

$$\hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^\mu \, \hat{\mathcal{R}}(\Lambda) \;=\; \Lambda^\mu{}_\nu \, \gamma^\nu \tag{1616}$$

From the transformational properties of the Dirac spinors, together with the identity

$$\gamma^{(0)} \, \hat{\mathcal{R}}^\dagger(\Lambda) \, \gamma^{(0)} \;=\; \hat{\mathcal{R}}^{-1}(\Lambda) \tag{1617}$$

one can find the transformational properties of quantities that are bi-linear in the Dirac spinors.

Thus, for example, the bi-linear quantity $\overline{\psi}^\dagger \, \psi$ transforms according to

$$
\begin{aligned}
\overline{\psi}^{\dagger\prime} \, \psi' \;&=\; \psi^{\dagger\prime} \, \gamma^{(0)} \, \psi' \\
&=\; \psi^\dagger \, \hat{\mathcal{R}}^\dagger(\Lambda) \, \gamma^{(0)} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&=\; \psi^\dagger \, (\, \gamma^{(0)} \,)^2 \, \hat{\mathcal{R}}^\dagger(\Lambda) \, \gamma^{(0)} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&=\; \psi^\dagger \, \gamma^{(0)} \, \hat{\mathcal{R}}^{-1}(\Lambda) \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&=\; \overline{\psi}^\dagger \, \psi
\end{aligned}
\tag{1618}
$$

Table 7: The sixteen bi-linear covariants $\overline{\psi}^{\dagger} \hat{Q} \, \psi$ for the Dirac equation.

Quantity	Bi-linear Covariant $\overline{\psi}^{\dagger} \, \hat{Q} \, \psi$	Transformed Matrix $\hat{\mathcal{R}}^{-1}(\Lambda) \, \hat{Q} \, \hat{\mathcal{R}}(\Lambda)$	Number of Matrices
Scalar	$\overline{\psi}^{\dagger} \, \hat{I} \, \psi$	$\hat{I}$	1
Vector	$\overline{\psi}^{\dagger} \, \gamma^{\mu} \, \psi$	$\Lambda^{\mu}{}_{\nu} \, \gamma^{\nu}$	4
Anti-symmetric Tensor	$\overline{\psi}^{\dagger} \, \sigma^{\mu,\nu} \, \psi$	$\Lambda^{\mu}{}_{\rho} \, \Lambda^{\nu}{}_{\tau} \, \sigma^{\rho,\tau}$	6
Pseudo-scalar	$\overline{\psi}^{\dagger} \, \gamma^{(4)} \, \psi$	$\det | \Lambda | \, \gamma^{(4)}$	1
Axial-Vector	$\overline{\psi}^{\dagger} \, \gamma^{(4)} \, \gamma^{\mu} \, \psi$	$\det | \Lambda | \, \Lambda^{\mu}{}_{\nu} \, \gamma^{(4)} \, \gamma^{\nu}$	4

where a factor of $(\, \gamma^{(0)} \,)^2 \; = \; \hat{I}$ has been used in the third line and the identity has been used in the fourth. Thus, one finds that $\overline{\psi}^{\dagger} \, \psi$ transforms like a scalar.

Likewise, one can show that the bi-linear quantities $\overline{\psi}^{\dagger} \, \gamma^{\mu} \, \psi$ transform like the components of a four-vector. That is

$$
\begin{aligned}
\overline{\psi}^{\dagger \prime} \, \gamma^{\mu} \, \psi' \;\; &= \;\; \psi^{\dagger \prime} \, \gamma^{(0)} \, \gamma^{\mu} \, \psi' \\
&= \;\; \psi^{\dagger} \, \hat{\mathcal{R}}^{\dagger}(\Lambda) \, \gamma^{(0)} \, \gamma^{\mu} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&= \;\; \psi^{\dagger} \, (\, \gamma^{(0)} \,)^2 \, \hat{\mathcal{R}}^{\dagger}(\Lambda) \, \gamma^{(0)} \, \gamma^{\mu} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&= \;\; \psi^{\dagger} \, \gamma^{(0)} \, \hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\mu} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&= \;\; \overline{\psi}^{\dagger} \, \hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\mu} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&= \;\; \Lambda^{\mu}{}_{\nu} \, \overline{\psi}^{\dagger} \, \gamma^{\nu} \, \psi
\end{aligned} \tag{1619}
$$

where the covariant condition has been used in obtaining the last line. Since this relation holds for Lorentz boosts, rotations and spatial inversions, $\overline{\psi}^{\dagger} \, \gamma^{\mu} \, \psi$ is a four-vector.

The anti-symmetric quantity $\sigma^{\mu,\nu}$ defined as

$$
\sigma^{\mu,\nu} \;\; = \;\; \frac{i}{2} \, [\, \gamma^{\mu} \, , \, \gamma^{\nu} \,] \tag{1620}
$$

can be used to form a bi-linear quantity $\overline{\psi}^{\dagger} \, \sigma^{\mu,\nu} \, \psi$. This bi-linear quantity

transforms like a second-rank anti-symmetric tensor, since

$$
\begin{aligned}
\overline{\psi}^{\dagger\prime} \, \sigma^{\mu,\nu} \, \psi' &= \psi^{\dagger\prime} \, \gamma^{(0)} \, \sigma^{\mu,\nu} \, \psi' \\
&= \psi^{\dagger} \, \hat{\mathcal{R}}^{\dagger}(\Lambda) \, \gamma^{(0)} \, \sigma^{\mu,\nu} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&= \psi^{\dagger} \, (\, \gamma^{(0)} \,)^2 \, \hat{\mathcal{R}}^{\dagger}(\Lambda) \, \gamma^{(0)} \, \sigma^{\mu,\nu} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&= \psi^{\dagger} \, \gamma^{(0)} \, \hat{\mathcal{R}}^{-1}(\Lambda) \, \sigma^{\mu,\nu} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&= \overline{\psi}^{\dagger} \, \hat{\mathcal{R}}^{-1}(\Lambda) \, \sigma^{\mu,\nu} \, \hat{\mathcal{R}}(\Lambda) \, \psi
\end{aligned}
\tag{1621}
$$

For $\mu \neq \nu$, the antisymmetric quantity $\sigma^{\mu,\nu}$ can be written as

$$
\sigma^{\mu,\nu} = i \, \gamma^{\mu} \, \gamma^{\nu}
\tag{1622}
$$

Therefore, one may re-express the bi-linear quantity as

$$
\begin{aligned}
\overline{\psi}^{\dagger\prime} \, \sigma^{\mu,\nu} \, \psi' &= i \, \overline{\psi}^{\dagger} \, \hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\mu} \, \gamma^{\nu} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&= i \, \overline{\psi}^{\dagger} \, \hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\mu} \, \hat{\mathcal{R}}(\Lambda) \, \hat{\mathcal{R}}^{-1}(\Lambda) \, \gamma^{\nu} \, \hat{\mathcal{R}}(\Lambda) \, \psi \\
&= i \, \overline{\psi}^{\dagger} \, \Lambda^{\mu}{}_{\rho} \, \gamma^{\rho} \, \Lambda^{\nu}{}_{\tau} \, \gamma^{\tau} \, \psi \\
&= \Lambda^{\mu}{}_{\rho} \, \Lambda^{\nu}{}_{\tau} \, \overline{\psi}^{\dagger} \, \sigma^{\rho,\tau} \, \psi
\end{aligned}
\tag{1623}
$$

where we have inserted a factor of $\hat{I} = \hat{\mathcal{R}}(\Lambda) \, \hat{\mathcal{R}}^{-1}(\Lambda)$ in the second line, and used the covariant condition (twice) in the third line. Hence, the bi-linear quantity $\overline{\psi}^{\dagger} \, \sigma^{\mu,\nu} \, \psi$ transforms like an anti-symmetric second-rank tensor.

One can define a quantity $\gamma^{(4)}$ in terms of a product of all the γ-matrices

$$
\gamma^{(4)} = i \, \gamma^{(0)} \, \gamma^{(1)} \, \gamma^{(2)} \, \gamma^{(3)}
\tag{1624}
$$

It is easily verified that $\gamma^{(4)}$ anti-commutes with all the γ^{μ},

$$
\{ \, \gamma^{\mu} \, , \, \gamma^{(4)} \, \}_{+} = 0
\tag{1625}
$$

Furthermore, one has

$$
(\, \gamma^{(4)} \,)^2 = \hat{I}
\tag{1626}
$$

In the standard representation of the Dirac matrices, the matrix $\gamma^{(4)}$ has the two by two block diagonal form

$$
\gamma^{(4)} = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}
\tag{1627}
$$

The quantity $\gamma^{(4)}$ can be used to construct a bi-linear covariant quantity $\overline{\psi}^{\dagger} \, \gamma^{(4)} \, \psi$. Under an orthochronous Lorentz transformation, the bi-linear quantity transform according to

$$
\overline{\psi}^{\dagger\prime} \, \gamma^{(4)} \, \psi' = \psi^{\dagger\prime} \, \gamma^{(0)} \, \gamma^{(4)} \, \psi'
$$

$$\begin{aligned}
&= \ \psi^\dagger \ \hat{\mathcal{R}}^\dagger(\Lambda) \ \gamma^{(0)} \ \gamma^{(4)} \ \hat{\mathcal{R}}(\Lambda) \ \psi \\
&= \ \psi^\dagger \ (\ \gamma^{(0)} \ )^2 \ \hat{\mathcal{R}}^\dagger(\Lambda) \ \gamma^{(0)} \ \gamma^{(4)} \ \hat{\mathcal{R}}(\Lambda) \ \psi \\
&= \ \overline{\psi}^\dagger \ \gamma^{(0)} \ \hat{\mathcal{R}}^\dagger(\Lambda) \ \gamma^{(0)} \ \gamma^{(4)} \ \hat{\mathcal{R}}(\Lambda) \ \psi \\
&= \ \overline{\psi}^\dagger \ \hat{\mathcal{R}}^{-1}(\Lambda) \ \gamma^{(4)} \ \hat{\mathcal{R}}(\Lambda) \ \psi
\end{aligned} \tag{1628}$$

A proper Lorentz transformation, such as a boost and a rotation, are generated by the quantities $\sigma^{\mu,\nu}$ which involves the anti-symmetrized product of the two Dirac matrices γ^μ and γ^ν. Since the matrices γ^μ and γ^ν individual anti-commute with $\gamma^{(4)}$, their product commutes with $\gamma^{(4)}$. Hence, one can commute $\gamma^{(4)}$ and $\hat{\mathcal{R}}(\Lambda)$. Therefore, the bi-linear quantity transforms as

$$\overline{\psi}^{\dagger\prime}(x') \ \gamma^{(4)} \ \psi'(x') \ = \ \overline{\psi}^\dagger(x) \ \gamma^{(4)} \ \psi(x) \tag{1629}$$

which behaves like a scalar under a proper orthochronous Lorentz transformation for which $\det | \Lambda | \ = \ 1$. However, for an inversion where $\det | \Lambda | \ = \ -1$, one has $\hat{\mathcal{R}}(\mathcal{P}) \ = \ \gamma^{(0)}$ which anti-commutes with $\gamma^{(4)}$. Hence, for an inversion, one has

$$\overline{\psi}^{\dagger\prime}(x') \ \gamma^{(4)} \ \psi'(x') \ = \ - \ \overline{\psi}^\dagger(x) \ \gamma^{(4)} \ \psi(x) \tag{1630}$$

so the quantity changes sign. In general, for an orthochronous transformation one can show that

$$\hat{\mathcal{R}}^{-1}(\Lambda) \ \gamma^{(4)} \ \hat{\mathcal{R}}(\Lambda) \ = \ \det | \Lambda | \ \gamma^{(4)} \tag{1631}$$

so one has

$$\overline{\psi}^{\dagger\prime}(x') \ \gamma^{(4)} \ \psi'(x') \ = \ \det | \Lambda | \ \overline{\psi}^\dagger(x) \ \gamma^{(4)} \ \psi(x) \tag{1632}$$

Therefore, the quantity $\overline{\psi}^\dagger \ \gamma^{(4)} \ \psi$ transforms as a pseudo-scalar.

One can also define the bi-linear axial-vector $\overline{\psi}^\dagger \ \gamma^{(4)} \ \gamma^\mu \ \psi$. From considerations similar to those used previously, one can show that these quantities transform according to

$$\overline{\psi}^{\dagger\prime}(x') \ \gamma^{(4)} \ \gamma^\mu \ \psi'(x') \ = \ \det | \Lambda | \ \Lambda^\mu{}_\nu \ \overline{\psi}^\dagger(x) \ \gamma^{(4)} \ \gamma^\nu \ \psi(x) \tag{1633}$$

Hence, $\overline{\psi}^\dagger \ \gamma^{(4)} \ \gamma^\mu \ \psi$ transforms like a four-vector under proper orthochronous Lorentz transformations. However, the spatial components do not change sign under an inversion, but the time components do change sign. Therefore, $\overline{\psi}^\dagger \ \gamma^{(4)} \ \gamma^\mu \ \psi$ transforms like an axial-vector.

Exercise:

Show that a modified Dirac equation described by

$$\left[i \ \hbar \ \gamma^\mu \ (\ \partial_\mu \ + \ i \ \frac{q}{\hbar \ c} \ A_\mu \ ) \ - \ i \ \frac{\kappa \ q \ \hbar}{4 \ m \ c^2} \ \sigma^{\mu,\nu} \ \gamma^{(4)} \ F_{\mu,\nu} \ - \ m \ c \ \hat{I} \ \right] \psi \ = \ 0 \tag{1634}$$

is covariant under proper Lorentz transformations, but is not covariant under improper transformations.

Show, by considering the non-relativistic limit, that the above equation describes an electron with an electric dipole moment. Determine an expression for the electric dipole moment.

12.10 The Spherically Symmetric Dirac Equation

The Dirac Hamiltonian for a (electrostatic) spherically symmetric potential is given by

$$\hat{H} \; = \; c \, \underline{\alpha} \cdot \underline{\hat{p}} \; + \; \beta \, m \, c^2 \; + \; \hat{I} \, V(r) \tag{1635}$$

The angular momentum operator $\underline{\hat{J}}$ and the parity operator $\hat{\mathcal{P}}$ commute with the Hamiltonian $\hat{H}$. Therefore, one can find simultaneous eigenstates of the three operators $\hat{H}$, $\hat{J}^2$, $\hat{J}_z$ and $\hat{\mathcal{P}}$. The energy eigenstates satisfy the equation

$$\left[\; c \, \underline{\alpha} \cdot \underline{\hat{p}} \; + \; \beta \, m \, c^2 \; + \; \hat{I} \, V(r) \; \right] \psi \; = \; E \, \psi \tag{1636}$$

On writing the four-component spinor in terms of the two two-component spinors ϕ^A and ϕ^B the energy eigenvalue equation reduces to the set of coupled equations

$$
\begin{aligned}
(\; E \; - \; V(r) \; - \; m \, c^2 \;) \, \phi^A(\underline{r}) \; &= \; c \, (\; \underline{\sigma} \cdot \underline{\hat{p}} \;) \, \phi^B(\underline{r}) \\
(\; E \; - \; V(r) \; + \; m \, c^2 \;) \, \phi^B(\underline{r}) \; &= \; c \, (\; \underline{\sigma} \cdot \underline{\hat{p}} \;) \, \phi^A(\underline{r})
\end{aligned}
\tag{1637}
$$

In spherical polar coordinates, the operator $(\; \underline{\sigma} \cdot \underline{\hat{p}} \;)$ can be expressed as

$$
\begin{aligned}
(\; \underline{\sigma} \cdot \underline{\hat{p}} \;) \; = \; &- \, i \, \hbar \left(\begin{array}{cc} \cos\theta & \sin\theta \, \exp[-i\varphi] \\ \sin\theta \, \exp[+i\varphi] & - \, \cos\theta \end{array} \right) \frac{\partial}{\partial r} \\
&- \frac{i \, \hbar}{r} \left(\begin{array}{cc} - \, \sin\theta & \cos\theta \, \exp[-i\varphi] \\ \cos\theta \, \exp[+i\varphi] & \sin\theta \end{array} \right) \frac{\partial}{\partial\theta} \\
&- \frac{i \, \hbar}{r \, \sin\theta} \left(\begin{array}{cc} 0 & - \, i \, \exp[-i\varphi] \\ i \, \exp[+i\varphi] & 0 \end{array} \right) \frac{\partial}{\partial\varphi}
\end{aligned}
\tag{1638}
$$

which has a quite complicated structure. For future reference, it shall be noted that the matrix part of the coefficient of the partial derivative w.r.t. r is simply equal to

$$\left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \tag{1639}$$

which is independent of the radial coordinate r. The operator $(\; \underline{\sigma} \cdot \underline{\hat{p}} \;)$ can be cast in a more convenient form through the repeated use of the Pauli identity.

First, the 2×2 unit matrix can be written as

$$I \ = \ \left(\ \frac{\underline{r} \cdot \underline{\sigma}}{r} \ \right)^2 \tag{1640}$$

since different Pauli spin matrices anti-commute

$$\{ \ \sigma^{(i)} \ , \ \sigma^{(j)} \ \}_+ \ = \ 2 \ \delta^{i,j} \ \hat{I} \tag{1641}$$

and are their own inverses. Therefore, one can express the operator $(\ \underline{\sigma} \cdot \underline{\hat{p}} \ )$ as

$$
\begin{aligned}
(\ \underline{\sigma} \cdot \underline{\hat{p}} \ ) \ &= \ \left(\ \frac{\underline{r} \cdot \underline{\sigma}}{r} \ \right)^2 (\ \underline{\sigma} \cdot \underline{\hat{p}} \ ) \\
&= \ \left(\ \frac{\underline{r} \cdot \underline{\sigma}}{r^2} \ \right) (\ \underline{r} \cdot \underline{\sigma} \ ) (\ \underline{\sigma} \cdot \underline{\hat{p}} \ ) \\
&= \ \left(\ \frac{\underline{r} \cdot \underline{\sigma}}{r^2} \ \right) \left(\ \underline{r} \cdot \underline{\hat{p}} \ + \ i \ \underline{\sigma} \cdot (\ \underline{r} \wedge \underline{\hat{p}} \ ) \ \right) \\
&= \ \left(\ \frac{\underline{r} \cdot \underline{\sigma}}{r^2} \ \right) \left(\ - \ i \ \hbar \ r \ \frac{\partial}{\partial r} \ + \ i \ \underline{\sigma} \cdot \underline{\hat{L}} \ \right) \\
&= \ \left(\ \frac{\underline{r} \cdot \underline{\sigma}}{r^2} \ \right) \left(\ - \ i \ \hbar \ r \ \frac{\partial}{\partial r} \ + \ \frac{2 \ i}{\hbar} \ \underline{S} \cdot \underline{\hat{L}} \ \right) \tag{1642}
\end{aligned}
$$

where the Pauli identity has been used in going between the second and third lines. Therefore, the two-component spinors satisfy the set of coupled equations

$$
\begin{aligned}
(\ E \ - \ V(r) \ - \ m \ c^2 \ ) \ \phi^A(\underline{r}) \ &= \ c \left(\frac{\underline{r} \cdot \underline{\sigma}}{r^2} \right) \left(\ - \ i \ \hbar \ r \ \frac{\partial}{\partial r} \ + \ \frac{2 \ i}{\hbar} \ \underline{S} \cdot \underline{\hat{L}} \ \right) \phi^B(\underline{r}) \\
(\ E \ - \ V(r) \ + \ m \ c^2 \ ) \ \phi^B(\underline{r}) \ &= \ c \left(\frac{\underline{r} \cdot \underline{\sigma}}{r^2} \right) \left(\ - \ i \ \hbar \ r \ \frac{\partial}{\partial r} \ + \ \frac{2 \ i}{\hbar} \ \underline{S} \cdot \underline{\hat{L}} \ \right) \phi^A(\underline{r})
\end{aligned} \tag{1643}
$$

It is seen that, due to the effect of special relativity, the Dirac equation results in the coupling of the spin and the orbital angular momentum.

Two-Component Spinor Spherical Harmonics

The angular dependence of the two-component wave functions $\phi^A(\underline{r})$ and $\phi^B(\underline{r})$ are determined by the eigenvalue equations for the magnitude and the z-components of the total angular momentum

$$\underline{J} \ = \ \underline{L} \ + \ \underline{S} \tag{1644}$$

Thus, the two-component spinor eigenstates of total angular momentum $\Omega^l_{j,j_z}(\theta, \varphi)$ which describes the angular dependence, are formed by combining states of orbital angular momentum l, represented by $Y^l_m(\theta, \varphi)$, and the spin eigenfunction

Table 8: The Clebsch-Gordon Coefficients for adding orbital angular momentum (l, m) with spin quantum numbers $(\frac{1}{2}, s_z)$ to yield a state with total angular momentum quantum numbers (j, j_z). The allowed values of m are given by $j_z = m + s_z$.

	$s_z = +\frac{1}{2}$	$s_z = -\frac{1}{2}$
$j = l + \frac{1}{2}$	$\sqrt{\dfrac{l + j_z + \frac{1}{2}}{2\,l + 1}}$	$\sqrt{\dfrac{l - j_z + \frac{1}{2}}{2\,l + 1}}$
$j = l - \frac{1}{2}$	$-\sqrt{\dfrac{l - j_z + \frac{1}{2}}{2\,l + 1}}$	$\sqrt{\dfrac{l + j_z + \frac{1}{2}}{2\,l + 1}}$

$\chi_\pm$. On combining states with orbital angular momentum l and spin $s = \frac{1}{2}$, one finds states with total angular momentum which satisfy

$$l + \frac{1}{2} \geq j \geq l - \frac{1}{2} \tag{1645}$$

Thus, it is found that the possible eigenstates correspond to $j = l + \frac{1}{2}$ and $j = l - \frac{1}{2}$. Furthermore, the corresponding eigenfunctions are expressed as

$$
\begin{aligned}
\Omega^l_{l+\frac{1}{2}, j_z}(\theta, \varphi) &= \sqrt{\frac{l + \frac{1}{2} + j_z}{2l + 1}}\; Y^l_{j_z - \frac{1}{2}}(\theta, \varphi)\; \chi_+ \;+\; \sqrt{\frac{l + \frac{1}{2} - j_z}{2l + 1}}\; Y^l_{j_z + \frac{1}{2}}(\theta, \varphi)\; \chi_- \\
\Omega^l_{l-\frac{1}{2}, j_z}(\theta, \varphi) &= -\sqrt{\frac{l + \frac{1}{2} - j_z}{2l + 1}}\; Y^l_{j_z - \frac{1}{2}}(\theta, \varphi)\; \chi_+ \;+\; \sqrt{\frac{l + \frac{1}{2} + j_z}{2l + 1}}\; Y^l_{j_z + \frac{1}{2}}(\theta, \varphi)\; \chi_-
\end{aligned}
\tag{1646}
$$

where the coefficients are identified with the Clebsch-Gordon coefficients given in Table(8). The functions $\Omega^l_{j, j_z}(\theta, \varphi)$ are the analogue of the spherical harmonics $Y^l_m(\theta, \varphi)$ in relativistic problems where spin and orbital angular momentum are coupled.

However, since orbital angular momentum is not a good quantum number, the angular dependence of the eigenstates of the Dirac Hamiltonian can be expressed as a linear superposition of states with different values of the orbital angular momentum l. For a fixed value of j, one finds that the possible values of the orbital angular momentum l are determined by

$$
\begin{aligned}
j &= l + \frac{1}{2} \\
j &= l' - \frac{1}{2}
\end{aligned}
\tag{1647}
$$

278

where $l' = l + 1$. The appropriate two-component spinor angular momentum eigenstate with quantum numbers (j, j_z) found by combining a spin one-half and orbital angular momentum $l' = (l + 1)$ is given by

$$\Omega^{l+1}_{l+\frac{1}{2},j_z}(\theta,\varphi) \; = \; -\sqrt{\frac{l + \frac{3}{2} - j_z}{2l + 3}} \; Y^{l+1}_{j_z - \frac{1}{2}}(\theta,\varphi) \; \chi_+ \; + \; \sqrt{\frac{l + \frac{3}{2} + j_z}{2l + 3}} \; Y^{l+1}_{j_z + \frac{1}{2}}(\theta,\varphi) \; \chi_-$$

(1648)

As shall be seen later, the two-component spinors $\Omega^{l'}_{l' - \frac{1}{2},j_z}(\theta,\varphi)$ and $\Omega^{l}_{l+\frac{1}{2},j_z}(\theta,\varphi)$ have opposite parities. In fact, the two-component spinors generated by angular momentum l and $l' = (l + 1)$ are related by the action of the pseudo-scalar

$$\left(\frac{r \cdot \sigma}{r} \right) \; = \; \left(\begin{array}{cc} \cos\theta & \sin\theta \; \exp[-i\varphi] \\ \sin\theta \; \exp[+i\varphi] & - \cos\theta \end{array} \right)$$

(1649)

which changes sign under a parity transformation, $(\theta,\varphi) \to (\pi - \theta, \varphi + \pi)$. The explicit relationship is given by

$$\left(\frac{r \cdot \sigma}{r} \right) \; \Omega^{l}_{j,j_z}(\theta,\varphi) \; = \; - \; \Omega^{l+1}_{j,j_z}(\theta,\varphi)$$

(1650)

as can be shown by examination of Table(1). Likewise, on using the identity

$$\left(\frac{r \cdot \sigma}{r} \right)^2 \; = \; I$$

(1651)

one finds that the inverse relationship between the two-component spinors is also given by

$$\left(\frac{r \cdot \sigma}{r} \right) \; \Omega^{l+1}_{j,j_z}(\theta,\varphi) \; = \; - \; \Omega^{l}_{j,j_z}(\theta,\varphi)$$

(1652)

Therefore, one concludes that the two angular momentum eigenstates have different properties under the spatial inversion transformation $r \to -r$.

Mathematical Interlude:
The Action of the Operator ($\hat{r} \cdot \sigma$) on the Spinor Spherical Harmonics $\Omega^{j \pm \frac{1}{2}}_{j,j_z}(\theta,\varphi)$.

Here, it will be argued that the spinor spherical harmonics satisfy the equations

$$\left(\frac{r \cdot \sigma}{r} \right) \; \Omega^{j + \frac{1}{2}}_{j,j_z}(\theta,\varphi) \; = \; - \; \Omega^{j - \frac{1}{2}}_{j,j_z}(\theta,\varphi)$$

$$\left(\frac{r \cdot \sigma}{r} \right) \; \Omega^{j - \frac{1}{2}}_{j,j_z}(\theta,\varphi) \; = \; - \; \Omega^{j + \frac{1}{2}}_{j,j_z}(\theta,\varphi)$$

(1653)

279

The components of the total angular momentum

$$\hat{J}^{(i)} \; = \; \hat{L}^{(i)} \; + \; \hat{S}^{(i)} \tag{1654}$$

commute with ($\underline{r}$. $\hat{\underline{S}}$). That is

$$[\; \hat{J}^{(i)} \; , \; (\; \underline{r} \; . \; \hat{\underline{S}} \;) \;] \; = \; 0 \tag{1655}$$

The complete proof of this statement immediately follows from the proof of the relation for any one component $\hat{J}^{(i)}$, since ($\underline{r}$. $\hat{\underline{S}}$) is spherically symmetric. Thus, for $i = 1$, one has

$$\begin{aligned}
[\; \hat{J}^{(1)} \; , \; (\; \underline{r} \; . \; \hat{\underline{S}} \;) \;] \; &= \; [\; \hat{L}^{(1)} \; + \; \hat{S}^{(1)} \; , \; x^{(1)} \; \hat{S}^{(1)} \; + \; x^{(2)} \; \hat{S}^{(2)} \; + \; x^{(3)} \; \hat{S}^{(3)} \;] \\
&= \; \hat{S}^{(2)} \; [\; \hat{L}^{(1)} \; , \; x^{(2)} \;] \; + \; \hat{S}^{(3)} \; [\; \hat{L}^{(1)} \; , \; x^{(3)} \;] \\
&\quad + \; x^{(2)} \; [\; \hat{S}^{(1)} \; , \; \hat{S}^{(2)} \;] \; + \; x^{(3)} \; [\; \hat{S}^{(1)} \; , \; \hat{S}^{(3)} \;]
\end{aligned} \tag{1656}$$

Using the commutation relations

$$[\; \hat{S}^{(i)} \; , \; \hat{S}^{(j)} \;] \; = \; i \; \hbar \; \varepsilon^{i,j,k} \; \hat{S}^{(k)} \tag{1657}$$

and

$$[\; \hat{L}^{(i)} \; , \; x^{(j)} \;] \; = \; i \; \hbar \; \varepsilon^{i,j,k} \; x^{(k)} \tag{1658}$$

one finds that

$$\begin{aligned}
[\; \hat{J}^{(1)} \; , \; (\; \underline{r} \; . \; \hat{\underline{S}} \;) \;] \; &= \; i \; \hbar \; \left(\; \hat{S}^{(2)} \; x^{(3)} \; - \; \hat{S}^{(3)} \; x^{(2)} \; + \; x^{(2)} \; \hat{S}^{(3)} \; - \; x^{(3)} \; \hat{S}^{(2)} \; \right) \\
&= \; 0
\end{aligned} \tag{1659}$$

which was to be shown. From repeated use of the above commutation relations which involve the components $\hat{J}^{(i)}$, it immediately follows that

$$[\; \hat{\underline{J}}^2 \; , \; (\; \underline{r} \; . \; \hat{\underline{S}} \;) \;] \; = \; 0 \tag{1660}$$

Thus, since $\Omega^{j \pm \frac{1}{2}}_{j,j_z}$ is a simultaneous eigenstate of $\hat{\underline{J}}^2$ and $\hat{J}^{(3)}$ and because these operators commute with ($\underline{r}$. $\hat{\underline{S}}$), then ($\underline{r}$. $\hat{\underline{S}}$) $\Omega^{j \pm \frac{1}{2}}_{j,j_z}$ is also a simultaneous eigenstate with eigenvalues (j, j_z).

Since the states ($\underline{r}$. $\hat{\underline{S}}$) $\Omega^{j \pm \frac{1}{2}}_{j,j_z}$ are simultaneous eigenstates of $\hat{\underline{J}}^2$ and $\hat{J}^{(3)}$ with eigenvalues (j, j_z), and because this subspace is spanned by the basis composed of the two states $\Omega^{j \pm \frac{1}{2}}_{j,j_z}(\theta, \varphi)$, the transformed states can be decomposed as

$$\begin{aligned}
\left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \Omega^{j + \frac{1}{2}}_{j,j_z}(\theta, \varphi) \; &= \; C_{++}(j, j_z) \; \Omega^{j + \frac{1}{2}}_{j,j_z}(\theta, \varphi) \; + \; C_{+-}(j, j_z) \; \Omega^{j - \frac{1}{2}}_{j,j_z}(\theta, \varphi) \\
\left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \Omega^{j - \frac{1}{2}}_{j,j_z}(\theta, \varphi) \; &= \; C_{-+}(j, j_z) \; \Omega^{j + \frac{1}{2}}_{j,j_z}(\theta, \varphi) \; + \; C_{--}(j, j_z) \; \Omega^{j - \frac{1}{2}}_{j,j_z}(\theta, \varphi)
\end{aligned}$$

$$\tag{1661}$$

where the coefficients $C_{\pm,\pm}(j, j_z)$ will be determined below.

First, we shall show that the coefficients $C_{\pm,\pm}(j, j_z)$ are independent of j_z. This follows as $\hat{J}^{\pm}$ commutes with $(\underline{r} \cdot \underline{\hat{S}})$ since all the components $\hat{J}^{(i)}$ commute with $(\underline{r} \cdot \underline{\hat{S}})$. Thus, one has

$$\hat{J}^{\pm} \left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta, \varphi) \;=\; \left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \hat{J}^{\pm} \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta, \varphi) \qquad (1662)$$

and

$$\hat{J}^{\pm} \left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta, \varphi) \;=\; C_{++}(j, j_z) \, \hat{J}^{\pm} \, \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta, \varphi) \;+\; C_{+-}(j, j_z) \, \hat{J}^{\pm} \, \Omega_{j,j_z}^{j-\frac{1}{2}}(\theta, \varphi)$$

$$\left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \hat{J}^{\pm} \, \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta, \varphi) \;=\; C_{++}(j, j_z \pm 1) \, \hat{J}^{\pm} \, \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta, \varphi) \;+\; C_{+-}(j, j_z \pm 1) \, \hat{J}^{\pm} \, \Omega_{j,j_z}^{j-\frac{1}{2}}(\theta, \varphi)$$

$$(1663)$$

Hence, on comparing the linearly-independent terms on the left-hand sides, one concludes that

$$\begin{aligned} C_{++}(j, j_z \pm 1) &\;=\; C_{++}(j, j_z) \\ C_{+-}(j, j_z \pm 1) &\;=\; C_{+-}(j, j_z) \end{aligned} \qquad (1664)$$

etc. Therefore, the coefficients $C_{\pm,\pm}(j, j_z)$ are independent of the value of j_z. Henceforth, we shall omit the index j_z in $C_{\pm,\pm}(j, j_z)$.

From considerations of parity, it can be determined that $C_{++}(j) = C_{--}(j) = 0$. Under the parity transformation $\underline{r} \to - \underline{r}$, one has

$$\Omega_{j,j_z}^{j\pm\frac{1}{2}}(\theta, \varphi) \;\to\; (- 1)^{j\pm\frac{1}{2}} \, \Omega_{j,j_z}^{j\pm\frac{1}{2}}(\theta, \varphi) \qquad (1665)$$

which follows from the properties of the spherical harmonics $Y_m^l(\theta, \varphi)$ under the parity transformation. Also one has

$$\left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \;\to\; - \left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \qquad (1666)$$

under the parity transform. Thus, after the parity transform, one finds that the transformed states have the decompositions

$$\begin{aligned} \left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta, \varphi) &\;=\; - \, C_{++}(j) \, \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta, \varphi) \;+\; C_{+-}(j) \, \Omega_{j,j_z}^{j-\frac{1}{2}}(\theta, \varphi) \\ \left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \Omega_{j,j_z}^{j-\frac{1}{2}}(\theta, \varphi) &\;=\; C_{-+}(j) \, \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta, \varphi) \;-\; C_{--}(j) \, \Omega_{j,j_z}^{j-\frac{1}{2}}(\theta, \varphi) \end{aligned}$$

$$(1667)$$

which by comparison with eqn(1661) leads to the identification

$$C_{++}(j) \;=\; C_{--}(j) \;=\; 0 \qquad (1668)$$

Therefore, recalling that the coefficients are independent of j_z, one can express the effect of the operator on the spinor spherical harmonics as

$$\left(\frac{r \cdot \sigma}{r}\right) \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta,\varphi) = C_{+-}(j) \, \Omega_{j,j_z}^{j-\frac{1}{2}}(\theta,\varphi)$$

$$\left(\frac{r \cdot \sigma}{r}\right) \Omega_{j,j_z}^{j-\frac{1}{2}}(\theta,\varphi) = C_{-+}(j) \, \Omega_{j,j_z}^{j+\frac{1}{2}}(\theta,\varphi) \tag{1669}$$

Furthermore, since

$$\left(\frac{r \cdot \sigma}{r}\right)^2 = I \tag{1670}$$

one obtains the condition

$$C_{+-}(j) \, C_{-+}(j) = 1 \tag{1671}$$

This condition can be made more restrictive as $\left(\frac{r \cdot \sigma}{r}\right)$ is Hermitean, which leads to

$$C_{+-}(j) = C_{-+}(j)^* \tag{1672}$$

The above two equations suggest that $C_{-+}(j)$ and $C_{+-}(j)$ are pure phase factors, such as

$$C_{+-}(j) = \exp\left[+ i \, \phi(j) \right]$$

$$C_{-+}(j) = \exp\left[- i \, \phi(j) \right] \tag{1673}$$

The phase factor can be completely determined by considering the relations (1669) with specific choices of the values of (θ,φ). As can be seen by examining the case where $\varphi = 0$, the phase $\phi(j)$ is either zero or π. For the case $\varphi = 0$, the operator simplifies to

$$\left(\frac{r \cdot \sigma}{r}\right) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \tag{1674}$$

The spinor spherical harmonics are given by

$$\Omega_{j,j_z}^{j+\frac{1}{2}}(\theta,\varphi) = \begin{pmatrix} -\sqrt{\frac{j+1-j_z}{2j+2}} \, Y_{j_z-\frac{1}{2}}^{j+\frac{1}{2}}(\theta,\varphi) \\ +\sqrt{\frac{j+1+j_z}{2j+2}} \, Y_{j_z+\frac{1}{2}}^{j+\frac{1}{2}}(\theta,\varphi) \end{pmatrix}$$

$$\Omega_{j,j_z}^{j-\frac{1}{2}}(\theta,\varphi) = \begin{pmatrix} \sqrt{\frac{j+j_z}{2j}} \, Y_{j_z-\frac{1}{2}}^{j-\frac{1}{2}}(\theta,\varphi) \\ \sqrt{\frac{j-j_z}{2j}} \, Y_{j_z+\frac{1}{2}}^{j-\frac{1}{2}}(\theta,\varphi) \end{pmatrix} \tag{1675}$$

which becomes real for $\varphi = 0$ since the spherical harmonics become real. Hence, on inspecting eqn(1669) with $\varphi = 0$, one concludes that the phase factors are equal and are purely real. That is

$$C_{+-}(j) = C_{-+}(j) = \pm 1 \tag{1676}$$

Finally, by considering $\theta = 0$, for which

$$\left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) = \left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right) \tag{1677}$$

and the spherical harmonics reduce

$$Y^{j \pm \frac{1}{2}}_{j_z \mp \frac{1}{2}}(0, \varphi) = \sqrt{\frac{2j + 1 \pm 1}{4 \pi}} \, \delta_{j_z \mp \frac{1}{2}, 0} \tag{1678}$$

one finds that, for fixed j, only the four spinor spherical harmonics $\Omega^{j \pm \frac{1}{2}}_{j, j_z}(0, \varphi)$

$$\Omega^{j + \frac{1}{2}}_{j, j_z}(0, \varphi) = \left(\begin{array}{c} -\sqrt{\frac{2j+1}{8\pi}} \, \delta_{j_z - \frac{1}{2}, 0} \\ +\sqrt{\frac{2j+1}{8\pi}} \, \delta_{j_z + \frac{1}{2}, 0} \end{array} \right)$$

$$\Omega^{j - \frac{1}{2}}_{j, j_z}(0, \varphi) = \left(\begin{array}{c} \sqrt{\frac{2j+1}{8\pi}} \, \delta_{j_z - \frac{1}{2}, 0} \\ \sqrt{\frac{2j+1}{8\pi}} \, \delta_{j_z + \frac{1}{2}, 0} \end{array} \right) \tag{1679}$$

are non-zero. The spinor spherical harmonics with $\theta = 0$ are connected via

$$\left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right) \Omega^{j \pm \frac{1}{2}}_{j, \pm \frac{1}{2}}(0, \varphi) = - \Omega^{j \mp \frac{1}{2}}_{j, \pm \frac{1}{2}}(0, \varphi) \tag{1680}$$

Hence, one has determined that

$$C_{+-}(j) = C_{-+}(j) = -1 \tag{1681}$$

which holds independent of the values of θ and j, so the effect of the operator on the spinor spherical harmonics is completely specified by

$$\left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \Omega^{j + \frac{1}{2}}_{j, j_z}(\theta, \varphi) = - \Omega^{j - \frac{1}{2}}_{j, j_z}(\theta, \varphi)$$

$$\left(\frac{\underline{r} \cdot \underline{\sigma}}{r} \right) \Omega^{j - \frac{1}{2}}_{j, j_z}(\theta, \varphi) = - \Omega^{j + \frac{1}{2}}_{j, j_z}(\theta, \varphi) \tag{1682}$$

as was to be shown.

The Ansatz

If one only considers the spatial part of the parity operator, $\hat{P}$, the two-component spinor states $\Omega^{l'}_{l' \pm \frac{1}{2}, j_z}(\theta, \varphi)$ have parities $(-1)^{l'}$

$$\hat{P} \, \Omega^{l'}_{l' \pm \frac{1}{2}, j_z}(\theta, \varphi) = (-1)^{l'} \, \Omega^{l'}_{l' \pm \frac{1}{2}, j_z}(\theta, \varphi) \tag{1683}$$

283

Furthermore, as has been seen, the upper and lower two-component spinors of the four-component Dirac spinor must have opposite intrinsic parity. Therefore, the desired simultaneous eigenstates for the relativistic electron can be either represented by the four-component Dirac spinor $\psi^-_{j,j_z}(\underline{r})$ with parity $(-1)^l = (-1)^{(l+\frac{1}{2}-\frac{1}{2})}$ of the form

$$\psi^-_{l+\frac{1}{2},j_z}(\underline{r}) = \begin{pmatrix} \frac{f^-(r)}{r} \, \Omega^l_{l+\frac{1}{2},j_z}(\theta,\varphi) \\ i \, \frac{g^-(r)}{r} \, \Omega^{l+1}_{l+\frac{1}{2},j_z}(\theta,\varphi) \end{pmatrix} \tag{1684}$$

or by $\psi^+_{j,j_z}(\underline{r})$

$$\psi^+_{l+\frac{1}{2},j_z}(\underline{r}) = \begin{pmatrix} \frac{f^+(r)}{r} \, \Omega^{l+1}_{l+\frac{1}{2},j_z}(\theta,\varphi) \\ i \, \frac{g^+(r)}{r} \, \Omega^{l}_{l+\frac{1}{2},j_z}(\theta,\varphi) \end{pmatrix} \tag{1685}$$

which has parity $(-1)^{l+\frac{1}{2}+\frac{1}{2}}$. In these expressions $f^\pm(r)$ and $g^\pm(r)$ are scalar radial functions that have to be determined as solutions of the radial equation. These states do not correspond to definite values of the orbital angular momentum since the upper and lower two-component spinors correspond to the different values of either l or $l' = l + 1$ for the orbital angular momentum.

To condense the notation, the energy eigenstates will be written in the compact form

$$\psi^\pm_{j,j_z}(\underline{r}) = \begin{pmatrix} \frac{f^\pm(r)}{r} \, \Omega^{l_A}_{j,j_z}(\theta,\varphi) \\ i \, \frac{g^\pm(r)}{r} \, \Omega^{l_B}_{j,j_z}(\theta,\varphi) \end{pmatrix} \tag{1686}$$

where $l_A = j \pm \frac{1}{2}$ and $l_B = j \mp \frac{1}{2}$.

The Radial Equation

We shall find the radial Dirac equation for the solution $\psi^\pm_{j,j_z}(\underline{r})$. The Dirac spinor wave functions in eqn(1685) and eqn(1684) are substituted into eqns(1643). The spin-orbit interaction term can be evaluated by squaring the expression

$$\hat{\underline{J}} = \hat{\underline{L}} + \underline{S} \tag{1687}$$

which leads to the identity

$$\underline{S} \cdot \hat{\underline{L}} = \frac{1}{2} \left(\hat{\underline{J}}^2 - \hat{\underline{L}}^2 - \underline{S}^2 \right) \tag{1688}$$

When this operator acts on the relativistic two-component spinor spherical harmonic $\Omega^{l_A}_{j,j_z}$, one finds

$$\underline{S} \cdot \hat{\underline{L}} \, \Omega^{l_A}_{j,j_z} = \frac{\hbar^2}{2} \left(j \, (\, j \, + \, 1 \,) \, - \, l_A \, (\, l_A \, + \, 1 \,) \, - \, \frac{3}{4} \right) \Omega^{l_A}_{j,j_z} \tag{1689}$$

which for $j = l_A + \frac{1}{2}$ yields

$$\underline{S} \cdot \hat{\underline{L}} \, \Omega_{j,j_z}^{l_A} \;=\; \frac{\hbar^2}{2} \, (\, j \,-\, \frac{1}{2} \,) \, \Omega_{j,j_z}^{l_A} \tag{1690}$$

and for $j = l_A - \frac{1}{2}$, one obtains

$$\underline{S} \cdot \hat{\underline{L}} \, \Omega_{j,j_z}^{l_A} \;=\; -\, \frac{\hbar^2}{2} \, (\, j \,+\, \frac{3}{2} \,) \, \Omega_{j,j_z}^{l_A} \tag{1691}$$

The Dirac equation can be written in the general form

$$(\, E \,-\, V(r) \,-\, m\,c^2 \,) \, \frac{f(r)}{r} \, \Omega_{j,j_z}^{l_A} \;=\; c\,\hbar \left(\frac{\underline{r} \cdot \underline{\sigma}}{r^2} \right) \left(r \, \frac{\partial}{\partial r} \,-\, \frac{2}{\hbar^2} \, \underline{S} \cdot \hat{\underline{L}} \right) \frac{g(r)}{r} \, \Omega_{j,j_z}^{l_B}$$

$$(\, E \,-\, V(r) \,+\, m\,c^2 \,) \, \frac{g(r)}{r} \, \Omega_{j,j_z}^{l_B} \;=\; -\, c\,\hbar \left(\frac{\underline{r} \cdot \underline{\sigma}}{r^2} \right) \left(r \, \frac{\partial}{\partial r} \,-\, \frac{2}{\hbar^2} \, \underline{S} \cdot \hat{\underline{L}} \right) \frac{f(r)}{r} \, \Omega_{j,j_z}^{l_A}$$

$$\tag{1692}$$

Following Dirac, it is customary to define an integer κ in terms of the eigenvalues of $\underline{S} \cdot \hat{\underline{L}}$ via

$$(\, \underline{S} \cdot \hat{\underline{L}} \,) \, \Omega_{j,j_z}^{l_A} \;=\; -\, \frac{\hbar^2}{2} \, (\, 1 \,+\, \kappa \,) \, \Omega_{j,j_z}^{l_A}$$

$$(\, \underline{S} \cdot \hat{\underline{L}} \,) \, \Omega_{j,j_z}^{l_B} \;=\; -\, \frac{\hbar^2}{2} \, (\, 1 \,-\, \kappa \,) \, \Omega_{j,j_z}^{l_B} \tag{1693}$$

Therefore, if $\Omega_{j,j_z}^{l_A} \;=\; \Omega_{j,j_z}^{j+\frac{1}{2}}$, i.e. $j = l_A - \frac{1}{2}$, then $\kappa \;=\; (j + \frac{1}{2})$.

Otherwise, if $\Omega_{j,j_z}^{l_A} \;=\; \Omega_{j,j_z}^{j-\frac{1}{2}}$, i.e. $j = l_A + \frac{1}{2}$, then $\kappa \;=\; -\,(j + \frac{1}{2})$.

On substituting the above expressions in the Dirac energy eigenvalue equation for $\psi_{j,j_z}(\underline{r})$ one finds

$$(\, E \,-\, V(r) \,-\, m\,c^2 \,) \, \frac{f(r)}{r} \, \Omega_{j,j_z}^{l_A} \;=\; c\,\hbar \left(\frac{\underline{r} \cdot \underline{\sigma}}{r^2} \right) \left(r \, \frac{\partial}{\partial r} \,+\, 1 \,-\, \kappa \right) \frac{g(r)}{r} \, \Omega_{j,j_z}^{l_B}$$

$$(\, E \,-\, V(r) \,+\, m\,c^2 \,) \, \frac{g(r)}{r} \, \Omega_{j,j_z}^{l_B} \;=\; -\, c\,\hbar \left(\frac{\underline{r} \cdot \underline{\sigma}}{r^2} \right) \left(r \, \frac{\partial}{\partial r} \,+\, 1 \,+\, \kappa \right) \frac{f(r)}{r} \, \Omega_{j,j_z}^{l_A}$$

$$\tag{1694}$$

Since the radial spin projection operator is independent of r, it can be commuted to the right of the differential operator in the large parenthesis. Then on using either the relation given in eqn(1650) or in eqn(1651), one finds that the relativistic spherical harmonics factor out of the equations, leading to

$$(\, E \,-\, V(r) \,-\, m\,c^2 \,) \, \frac{f(r)}{r} \;=\; -\, c\,\hbar \left(\frac{\partial}{\partial r} \,+\, \frac{1 \,-\, \kappa}{r} \right) \frac{g(r)}{r}$$

$$(\, E \,-\, V(r) \,+\, m\,c^2 \,) \, \frac{g(r)}{r} \;=\; c\,\hbar \left(\frac{\partial}{\partial r} \,+\, \frac{1 \,+\, \kappa}{r} \right) \frac{f(r)}{r}$$

$$\tag{1695}$$

Table 9: The Relationship between j, l^A, l^B, κ and Parity. The parity eigenvalue is given by $\eta_p = (-1)^{l^A}$ and $\kappa = \pm(j + \frac{1}{2})$.

κ	l^A	l^B	Parity
$\kappa = (j + \frac{1}{2})$	$j + \frac{1}{2}$	$j - \frac{1}{2}$	$(-1)^{\kappa}$
$\kappa = -(j + \frac{1}{2})$	$j - \frac{1}{2}$	$j + \frac{1}{2}$	$(-1)^{1-\kappa}$

Therefore, the Dirac radial equation consists of the two coupled first-order differential equations for $f(r)$ and $g(r)$. On multiplying by a factor of r and simplifying the derivatives of $f(r)/r$, one finds the pair of more symmetrical equations

$$
(E - V(r) - m\,c^2)\, f(r) + c\,\hbar \left(\frac{\partial}{\partial r} - \frac{\kappa}{r} \right) g(r) = 0
$$
$$
(E - V(r) + m\,c^2)\, g(r) - c\,\hbar \left(\frac{\partial}{\partial r} + \frac{\kappa}{r} \right) f(r) = 0
$$

$$(1696)$$

The above pair of equations are the central result of this lecture.

The Probability Density in Spherical Polar Coordinates.

The probability density $P(\underline{r})$ that an electron, in an energy eigenstate of a spherically symmetric potential, is found in the vicinity of the point (r, θ, φ) is given by

$$
\begin{aligned}
P(\underline{r}) &= \psi^\dagger(\underline{r})\, \psi(\underline{r}) \\
&= \left(\frac{|f(r)|^2}{r^2} \right) \Omega^{l_A}_{j,j_z}(\theta, \varphi)^\dagger \, \Omega^{l_A}_{j,j_z}(\theta, \varphi) + \left(\frac{|g(r)|^2}{r^2} \right) \Omega^{l_B}_{j,j_z}(\theta, \varphi)^\dagger \, \Omega^{l_B}_{j,j_z}(\theta, \varphi)
\end{aligned}
$$

$$(1697)$$

However, due to the identity

$$
\Omega^{l_A}_{j,j_z}(\theta, \varphi)^\dagger \, \Omega^{l_A}_{j,j_z}(\theta, \varphi) = \Omega^{l_B}_{j,j_z}(\theta, \varphi)^\dagger \, \Omega^{l_B}_{j,j_z}(\theta, \varphi) = A_{j,|j_z|}(\theta) \qquad (1698)
$$

the probability is independent of the azimuthal angle φ and the sign of j_z (just like in the non-relativistic case) and has a common angular factor of $A_{j,|j_z|}(\theta)$. Thus, the probability distribution factorizes into a radial and the angular factor

$$
P(\underline{r}) = \left(\frac{|f(r)|^2}{r^2} + \frac{|g(r)|^2}{r^2} \right) A_{j,|j_z|}(\theta) \qquad (1699)
$$

Table 10: Relativistic Angular Distribution Functions

j	$|j_z|$	$A_{j,|j_z|}(\theta)$
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4\,\pi}$
$\frac{3}{2}$	$\frac{1}{2}$	$\frac{1}{8\,\pi}\,(\,1\,+\,3\,\cos^2\theta\,)$
$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{8\,\pi}\,\sin^2\theta$
$\frac{5}{2}$	$\frac{1}{2}$	$\frac{3}{16\,\pi}\,(\,1\,-\,2\,\cos^2\theta\,+\,5\,\cos^4\theta\,)$
$\frac{5}{2}$	$\frac{3}{2}$	$\frac{3}{32\,\pi}\,\sin^2\theta\,(\,1\,+\,15\,\cos^2\theta\,)$
$\frac{5}{2}$	$\frac{5}{2}$	$\frac{15}{32\,\pi}\,\sin^4\theta$

The angular distribution function for a closed shell is given by the sum over the angular distribution functions. Due to the identity,

$$\sum_{j_z=-j}^{j} A_{j,|j_z|}(\theta) \;=\; \frac{2\,j\,+\,1}{4\,\pi} \tag{1700}$$

one finds that closed shells are spherically symmetric, as is expected. The first few angular dependent factors $A_{j,|j_z|}(\theta)$ are given in Table(10) and the corresponding non-relativistic angular factors are given in Table(11). On comparing the relativistic angular dependent factors with the non-relativistic factors $|Y_m^l(\theta,\varphi)|^2$, one finds that they are identical for $|j_z| = j$. Since the relativistic distribution is the sum of two generally different positive definite forms originally associated with the two spinors χ_+ and χ_-, it generally does not go to zero for non-zero values of θ.

12.10.1 The Hydrogen Atom

The radial energy eigenvalue equation for a hydrogenic-like atom is given by

$$(\,E\,+\,\frac{Z\,e^2}{r}\,-\,m\,c^2\,)\,f(r)\,+\,c\,\hbar\,\left(\,\frac{\partial}{\partial r}\,-\,\frac{\kappa}{r}\,\right)\,g(r)\;=\;0$$
$$(\,E\,+\,\frac{Z\,e^2}{r}\,+\,m\,c^2\,)\,g(r)\,-\,c\,\hbar\,\left(\,\frac{\partial}{\partial r}\,+\,\frac{\kappa}{r}\,\right)\,f(r)\;=\;0$$
$$\tag{1701}$$
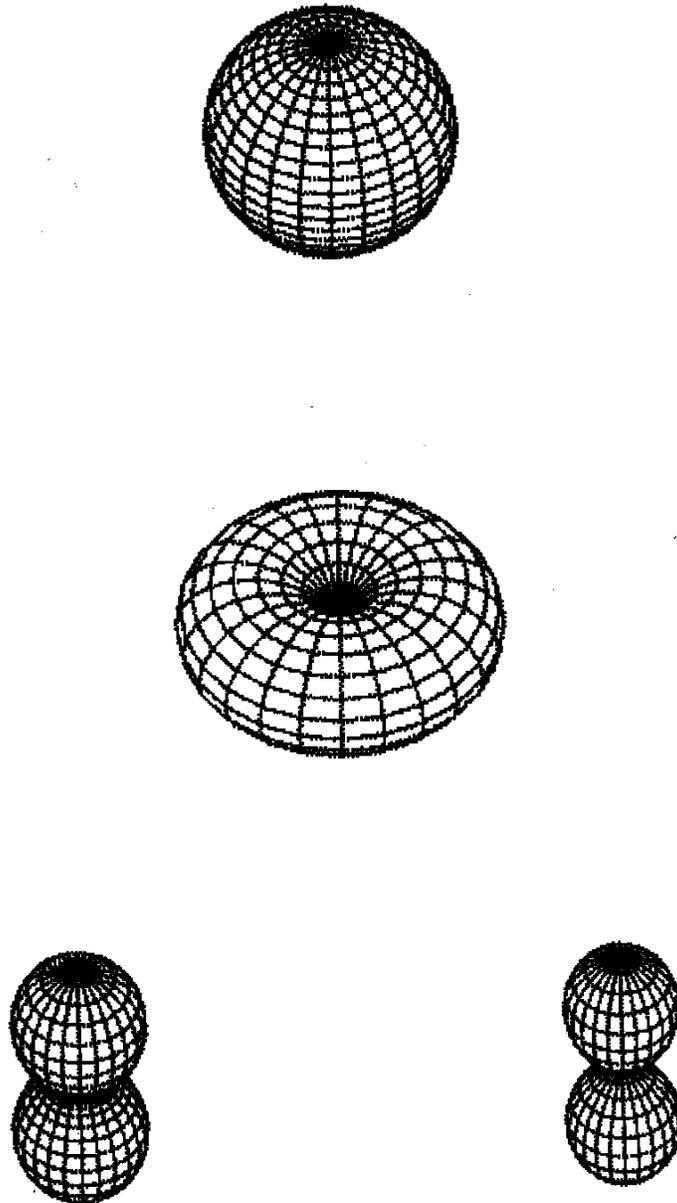
Figure 48: The relativistic (left) and non-relativistic (right) angular distributions $A_{j,|j_z|}(\theta)$ for $j = \frac{1}{2}$ and $j = \frac{3}{2}$.
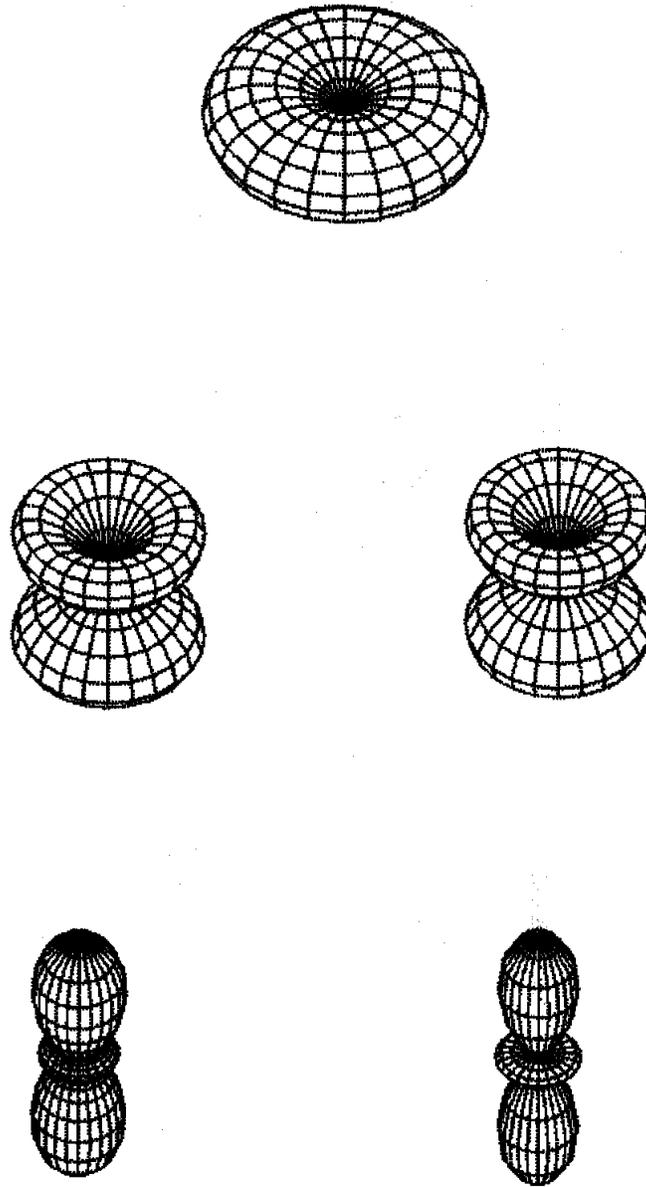
Figure 49: The relativistic (left) and non-relativistic (right) angular distributions $A_{j,|j_z|}(\theta)$ for $j = \frac{5}{2}$.

Table 11: Non-Relativistic Angular Distribution Functions

l	$|m|$	$A_{l,|m|}(\theta)$
0	0	$\frac{1}{4\,\pi}$
1	0	$\frac{3}{4\,\pi}\,\cos^2\theta$
1	1	$\frac{3}{8\,\pi}\,\sin^2\theta$
2	0	$\frac{5}{16\,\pi}\,(\,1\,-\,3\,\cos^2\theta\,)^2$
2	1	$\frac{15}{8\,\pi}\,\sin^2\theta\,\cos^2\theta$
2	2	$\frac{15}{32\,\pi}\,\sin^4\theta$

The above equations will be written in dimensionless units, where the energy is expressed in terms of the rest mass $m\,c^2$ and lengths are expressed in terms of the Compton wave length $\frac{\hbar}{m\,c}$. A dimensionless energy ϵ is defined as the ratio of E to the rest mass energy

$$\epsilon \;=\; \frac{E}{m\,c^2} \tag{1702}$$

For a bound state, $m\,c^2 \;>\; E \;>\; -\,m\,c^2$ so the value of the magnitude of ϵ is expected to be a little less than unity. A dimensionless radial variable ρ is introduced which governs the asymptotic large r decay of the bound state wave function. The variable is defined by

$$\rho \;=\; \sqrt{1\,-\,\epsilon^2}\,\left(\frac{r\,m\,c}{\hbar}\right) \tag{1703}$$

In terms of these dimensionless variables, the Dirac radial equations for the hydrogen-like atom become

$$\left(\,-\,\sqrt{\frac{1\,-\,\epsilon}{1\,+\,\epsilon}}\,+\,\frac{\gamma}{\rho}\,\right)f\,+\,\left(\frac{\partial}{\partial\rho}\,-\,\frac{\kappa}{\rho}\right)g \;=\; 0$$

$$\left(\,\sqrt{\frac{1\,+\,\epsilon}{1\,-\,\epsilon}}\,+\,\frac{\gamma}{\rho}\,\right)g\,-\,\left(\frac{\partial}{\partial\rho}\,+\,\frac{\kappa}{\rho}\right)f \;=\; 0 \tag{1704}$$

where

$$\gamma \;=\; \left(\frac{Z\,e^2}{\hbar\,c}\right) \tag{1705}$$

is a small number.

Boundary Conditions

The asymptotic $\rho \to \infty$ form of the solution can be found from the asymptotic form of the equations

$$
-\sqrt{\frac{1-\epsilon}{1+\epsilon}}\, f \; + \; \frac{\partial}{\partial \rho}\, g \;\; \sim \;\; 0
$$

$$
\sqrt{\frac{1+\epsilon}{1-\epsilon}}\, g \; - \; \frac{\partial}{\partial \rho}\, f \;\; \sim \;\; 0 \tag{1706}
$$

Hence, on combing these equations, one sees that the asymptotic form of the equation is given by

$$
\frac{\partial^2 f}{\partial \rho^2} \;=\; f \tag{1707}
$$

Therefore, one has

$$
f \;\sim\; A \, \exp\left[-\rho \right] \;+\; B \, \exp\left[+\rho \right] \tag{1708}
$$

and, likewise, g has a similar exponential form. If the solution is to be normalizable, then the coefficient B in front of the increasing exponential must be exactly zero $(B \equiv 0)$.

The asymptotic $\rho \to 0$ behavior of the solution can be found from

$$
\gamma\, f \; + \; \left(\rho\, \frac{\partial}{\partial \rho} \; - \; \kappa \right) g \;\; = \;\; 0
$$

$$
\gamma\, g \; - \; \left(\rho\, \frac{\partial}{\partial \rho} \; + \; \kappa \right) f \;\; = \;\; 0 \tag{1709}
$$

where it has been noted that both the angular momentum term κ and the Coulomb potential γ govern the small ρ variation, while the mass and energy terms are negligible. This is in contrast to the case of the non-relativistic Schrödinger equation with the Coulomb potential, where for small r the Coulomb potential term is negligible in comparison with the centrifugal potential. We shall make the ansatz for the asymptotic small ρ variation

$$
f \;\;\sim\;\; A\, \rho^s
$$

$$
g \;\;\sim\;\; B\, \rho^s \tag{1710}
$$

where the exponent s is an unknown constant and then substitute the ansatz in the above equations. This procedure yields the coupled algebraic equations

$$
\gamma\, A \; + \; (\, s \, - \, \kappa \,)\, B \;\; = \;\; 0
$$

$$
\gamma\, B \; - \; (\, s \, + \, \kappa \,)\, A \;\; = \;\; 0 \tag{1711}
$$

Hence, it is found that the exponent s is determined as solutions of the indicial equation which is a quadratic equation. The solutions are given by

$$s = \pm \sqrt{\kappa^2 - \gamma^2} \tag{1712}$$

Since, the wave function must be normalizable near $\rho \to 0$,

$$\lim_{\eta \to 0} \int_\eta dr \left(\mid f \mid^2 + \mid g \mid^2 \right) < \infty \tag{1713}$$

one must choose the positive solution for s. Normalizability near the origin requires that $2\,s\,>\,-1$. Hence, one may set

$$s = \sqrt{\kappa^2 - \gamma^2} \tag{1714}$$

This will be a good solution for $\kappa = -1$ if Z does not exceed a critical value. For values of Z greater than ≈ 172, the point charge can spark the vacuum and spontaneously generate electron-positron pairs[109]. The solution with the negative value of s given by

$$s = -\sqrt{\kappa^2 - \gamma^2} \tag{1715}$$

could also possibly exist and be normalizable if γ is greater than a critical value γ_c determined as

$$\frac{1}{2} = \sqrt{1 - \gamma_c^2} \tag{1716}$$

This critical value of γ is found from

$$\gamma_c = \frac{\sqrt{3}}{2} \tag{1717}$$

which corresponds to $Z_c \sim 118$. The solutions corresponding to negative s are, infact, un-physical and do not survive if the nucleus is considered to have a finite spatial extent.

The Fröbenius Method

We shall use the Fröbenius method to find a solution. The solutions of the radial equation shall be written in the form

$$f(r) = \exp\left[-\rho \right] \rho^s F(\rho)$$

$$g(r) = \exp\left[-\rho \right] \rho^s G(\rho) \tag{1718}$$

---

[109]H. Backe, L. Handschug, F. Hessberger, E. Kankeleit, L. Richter, F. Weik, R. Willwater, H. Bokemeyer, P. Vincent, Y. Nakayama, and J. S. Greenberg, Phys. Rev. Lett. **40**, 1443 (1978).

This form incorporates the appropriate boundary conditions at $\rho \to 0$ and $\rho \to \infty$. The coupled radial equations are transformed to

$$
\left(-\sqrt{\frac{1-\epsilon}{1+\epsilon}}\, \rho + \gamma \right) F + \left(\rho\, \frac{\partial}{\partial \rho} + s - \kappa - \rho \right) G = 0
$$

$$
\left(\sqrt{\frac{1+\epsilon}{1-\epsilon}}\, \rho + \gamma \right) G - \left(\rho\, \frac{\partial}{\partial \rho} + s + \kappa - \rho \right) F = 0
$$

$$(1719)$$

The functions $F(\rho)$ and $G(\rho)$ can be expressed as an infinite power series in ρ

$$
F(\rho) = \sum_{n=0}^{\infty} a_n\, \rho^n
$$

$$
G(\rho) = \sum_{n=0}^{\infty} b_n\, \rho^n \qquad (1720)
$$

where the coefficients a_n and b_n are constants which have still to be determined. The coefficients are determined by substituting the series in the differential equation and then equating the coefficients of the same power in ρ. Equating the coefficient of ρ^n yields the set of relations

$$
\left(-\sqrt{\frac{1-\epsilon}{1+\epsilon}}\, a_{n-1} + \gamma\, a_n \right) + (\, n + s - \kappa\,)\, b_n - b_{n-1} = 0
$$

$$
\left(\sqrt{\frac{1+\epsilon}{1-\epsilon}}\, b_{n-1} + \gamma\, b_n \right) - (\, n + s + \kappa\,)\, a_n + a_{n-1} = 0
$$

$$(1721)$$

This equation is automatically satisfied for $n = 0$, since by definition $a_{-1} = b_{-1} \equiv 0$ so the equation reduces to the indicial equation for s. These relations yield recursion relations between the coefficients (a_n, b_n) with different values of n. The form of the recursion relation can be made explicit by using a relation between a_n and b_n valid for any n. This relation is found by multiplying the first relation of eqn(1721) by the factor

$$
\sqrt{\frac{1+\epsilon}{1-\epsilon}} \qquad (1722)
$$

and adding it to the second, one sees that the coefficients with index $n - 1$ vanish. This process results in the equation

$$
\left(\sqrt{\frac{1+\epsilon}{1-\epsilon}}\, \gamma - (\, n + s + \kappa\,) \right) a_n + \left(\gamma + \sqrt{\frac{1+\epsilon}{1-\epsilon}}\, (\, n + s - \kappa\,) \right) b_n = 0
$$

$$(1723)$$

valid for any n. The above equation can be used to eliminate the coefficients b_n and yield a recursion relation between a_n and a_{n-1}. The ensuing recursion

relation will enable us to explicitly calculate the wave functions $G(\rho)$ and hence $F(\rho)$.

Truncation of the Series

The behavior of the recursion relation for large values of n can be found by noting that eqn(1723) yields

$$n \; a_n \; \sim \; \sqrt{\frac{1 \; + \; \epsilon}{1 \; - \; \epsilon}} \; n \; b_n \tag{1724}$$

which when substituted back into the large n limit of the first relation of eqn(1721) yields

$$n \; a_n \; \sim \; 2 \; a_{n-1} \tag{1725}$$

Since the large ρ limit of the function is dominated by the highest powers of ρ, it is seen that if the series does not terminate, the functions $F(\rho)$ and $G(\rho)$ would be exponentially growing functions of ρ

$$F(\rho) \; \sim \; \exp\left[\; + \; 2 \; \rho \; \right]$$

$$G(\rho) \; \sim \; \exp\left[\; + \; 2 \; \rho \; \right] \tag{1726}$$

Therefore, the set of recursion relations must terminate, since if the series does not terminate, the large ρ behavior of the functions $F(\rho)$ and $G(\rho)$ would governed by the growing exponentials. Even when combined with the decaying exponential term that appear in the relations

$$f(r) \; = \; \rho^s \; F(\rho) \; \exp\left[\; - \; \rho \; \right]$$

$$g(r) \; = \; \rho^s \; G(\rho) \; \exp\left[\; - \; \rho \; \right] \tag{1727}$$

the resulting functions $f(r)$ and $g(r)$ would not satisfy the required boundary conditions at $\rho \to \infty$. We shall assume that the series truncate after the n_r-th terms. That is, it is possible to set

$$a_{n_r+1} \; = \; 0$$

$$b_{n_r+1} \; = \; 0 \tag{1728}$$

Thus, the components of the radial wave function may have n_r nodes. Assuming that the coefficients with indices $n_r + 1$ vanish and using the first relation in eqn(1721) with $n = n_r + 1$, one obtains the condition

$$\sqrt{\frac{1 \; + \; \epsilon}{1 \; - \; \epsilon}} \; b_{n_r} \; = \; - \; a_{n_r} \tag{1729}$$

A second condition is given by the relation between a_n and b_n

$$\left(\sqrt{\frac{1 + \epsilon}{1 - \epsilon}} \, \gamma - (n + s + \kappa) \right) a_n + \left(\gamma + \sqrt{\frac{1 + \epsilon}{1 - \epsilon}} \, (n + s - \kappa) \right) b_n = 0 \tag{1730}$$

valid for any n. We shall set $n = n_r$ and then eliminate a_{n_r} using the termination condition expressed by eqn(1729). After some simplification, this leads to the equation

$$\epsilon \, \gamma = (n_r + s) \sqrt{1 - \epsilon^2} \tag{1731}$$

This equation determines the square of the dimensionless energy eigenvalue ϵ^2. On squaring this equation, simplifying and taking the square root, one finds

$$\epsilon = \pm \frac{(n_r + s)}{\sqrt{(n_r + s)^2 + \gamma^2}} \tag{1732}$$

or, equivalently, the energy of the hydrogen atom[110] is given by

$$E = \pm \frac{m \, c^2}{\sqrt{1 + \frac{\gamma^2}{(n_r + s)^2}}} \tag{1733}$$

where

$$s = \sqrt{(j + \frac{1}{2})^2 - \gamma^2} \tag{1734}$$

This expression for the energy eigenvalue is independent of the sign of κ and, therefore, it holds for both cases

$$j = (l + 1) - \frac{1}{2}$$
$$j = l + \frac{1}{2} \tag{1735}$$

Hence, the energy eigenstates are predicted to be doubly degenerate (in addition to the $(2j + 1)$ degeneracy associated with j_3), since states with the same j but have different values of l' have the same energy. If the positive-energy eigenvalue is expanded in powers of γ, one obtains

$$E \approx m \, c^2 - \frac{1}{2} \, m \, c^2 \, \frac{\gamma^2}{(n_r + j + \frac{1}{2})^2} + \dots \tag{1736}$$

which agrees with the energy eigenvalues found from the non-relativistic Schrödinger equation. However, as has been seen, the exact energy eigenvalue depends on n_r and $(j + \frac{1}{2})$ separately, as opposed to being a function of the principle quantum number n which is defined as the sum $n = n_r + j + \frac{1}{2}$. Hence, the Dirac

[110]C. G. Darwin, Proc. Roy. Soc. **A 118**, 654 (1928).
C. G. Darwin, Proc. Roy. Soc. **A 120**, 621 (1928).
W. Gordon, Zeit. für Physik, **48**, 11 (1928).

equation lifts the degeneracy between states with different values of the angular momentum. The energy levels together with their quantum numbers are shown in Table(12). The energy splitting between states with the same n and different j values has a magnitude which is governed by the square of the fine structure constant Z $(\frac{e^2}{\hbar c})$. That is

$$
\begin{aligned}
E \quad \approx \quad m c^2 & \left[1 \; - \; \frac{1}{2} \; \frac{\gamma^2}{(\, n_r \; + \; j \; + \; \frac{1}{2} \,)^2} \right. \\
& \left. - \; \frac{1}{2} \; \frac{\gamma^4}{(\, n_r \; + \; j \; + \; \frac{1}{2} \,)^3} \; \left(\frac{1}{(\, j \; + \; \frac{1}{2} \,)} \; - \; \frac{3}{4 \, (\, n_r \; + \; j \; + \; \frac{1}{2} \,)} \right) + \; \ldots \right]
\end{aligned}
$$

$$(1737)$$

The fine structure splittings for H-like atoms was first observed by Michelson[111] and the theoretical prediction is in agreement with the accurate measurements of Paschen[112]. The fine structure splitting is important for atoms with larger Z. This observation has a classical interpretation which reflects the fact that for large Z the electrons move in orbits with smaller radii and, therefore, the electrons must move faster. Relativistic effects become more important for electrons which move faster, and this occurs for atoms with larger values of Z. Although the fine structure splitting does remove some degeneracy, the two states with the same principle quantum number n and the same angular momentum j but which have different values of l are still predicted to be degenerate. Thus, for example, the $2S_{j=\frac{1}{2}}$ and the $2P_{j=\frac{1}{2}}$ states of Hydrogen are predicted to be degenerate by the Dirac equation. It has been shown that this degeneracy is removed by the Lamb shift, which is due to the interaction of an electron with its own radiation field. The Lamb shift is smaller than the fine structure shifts discussed above because it involves an extra factor of $\frac{e^2}{\hbar c}$.

The Ground State Wave Function

The ground state wave function of the hydrogen atom is slightly singular at the origin. This can be seen by noting that it corresponds to $n_r \; = \; 0$ and $\kappa \; = \; -1$. Since the dimensionless energy is given by the expression

$$\epsilon \; = \; \sqrt{1 \; - \; \gamma^2} \tag{1738}$$

one finds that the dimensionless radial distance ρ is simply given by

$$
\begin{aligned}
\rho \quad &= \quad \sqrt{1 \; - \; \epsilon^2} \; \left(\frac{r \, m \, c}{\hbar} \right) \\
&= \quad \gamma \; \left(\frac{m \, c}{\hbar} \right) \, r \\
&= \quad \frac{Z \, e^2 \, m}{\hbar^2} \, r
\end{aligned}
\tag{1739}
$$

[111]A. A. Michelson, Phil. Mag. **31**, 338 (1891).
[112]F. Paschen, Ann. Phys. **50**, 901 (1916).

Table 12: The Equivalence between Relativistic and Spectroscopic Quantum Numbers.

$n = n_r + |\kappa|$	n_r	$\kappa = \pm(j + \frac{1}{2})$	nL_j	Degenerate Partner	$\left(\frac{E}{m\ c^2}\right)$
1	0	-1	$1S_{\frac{1}{2}}$		$\sqrt{1 - \gamma^2}$
2	1	-1	$2S_{\frac{1}{2}}$	$2P_{\frac{1}{2}}$	$\sqrt{1 - \frac{\gamma^2}{2 + 2\sqrt{1-\gamma^2}}}$
2	1	+1	$2P_{\frac{1}{2}}$	$2S_{\frac{1}{2}}$	- -
2	0	-2	$2P_{\frac{3}{2}}$		$\sqrt{1 - \frac{1}{4}\ \gamma^2}$
3	2	-1	$3S_{\frac{1}{2}}$	$3P_{\frac{1}{2}}$	$\sqrt{1 - \frac{\gamma^2}{5 + 4\sqrt{1-\gamma^2}}}$
3	2	+1	$3P_{\frac{1}{2}}$	$3S_{\frac{1}{2}}$	- -
3	1	-2	$3P_{\frac{3}{2}}$	$3D_{\frac{3}{2}}$	$\sqrt{1 - \frac{\gamma^2}{5 + 2\sqrt{4-\gamma^2}}}$
3	1	+2	$3D_{\frac{3}{2}}$	$3P_{\frac{3}{2}}$	- -
3	0	-3	$3D_{\frac{5}{2}}$		$\sqrt{1 - \frac{1}{9}\ \gamma^2}$

where the characteristic length scale is just the non-relativistic Bohr radius divided by Z. The wave functions are written as

$$f(\rho) = \exp\left[-\rho\right] \rho^s F(\rho)$$

$$g(\rho) = \exp\left[-\rho\right] \rho^s G(\rho) \tag{1740}$$

Since $n_r = 0$, the recursion relations terminate immediately leading to the functions $F(\rho)$ and $G(\rho)$ being given, respectively by constants a_0 and b_0. Since $\kappa = -1$ for the ground state, the recursion relations are simply

$$\gamma\, a_0 + (s + 1)\, b_0 = 0$$
$$\gamma\, b_0 - (s - 1)\, a_0 = 0 \tag{1741}$$

The solution of the equations results in the index s being given by

$$s = \sqrt{1 - \gamma^2} \tag{1742}$$

and

$$\frac{b_0}{a_0} = \frac{s + \kappa}{\gamma} = \frac{\sqrt{1 - \gamma^2} - 1}{\gamma} \tag{1743}$$

This shows that the lower component is smaller than the upper constant by approximately γ, which has the magnitude of $\frac{v}{c}$ where v is the velocity in Bohr's theory. The ratio of b_0 to a_0 determines the radial functions as

$$\frac{f(r)}{r} = a_0\, \rho^{s-1}\, \exp\left[-\rho\right]$$

$$\frac{g(r)}{r} = a_0\, \frac{\sqrt{1 - \gamma^2} - 1}{\gamma}\, \rho^{s-1}\, \exp\left[-\rho\right] \tag{1744}$$

Since $Y_0^0(\theta, \varphi) = \frac{1}{\sqrt{4\,\pi}}$, the angular spherical harmonics for the upper components are just

$$\Omega^A(\theta, \varphi) = \frac{1}{\sqrt{4\,\pi}}\, \chi_\sigma \tag{1745}$$

and the lower components are given by

$$\Omega^B(\theta, \varphi) = -\frac{r \cdot \sigma}{r}\, \frac{1}{\sqrt{4\,\pi}}\, \chi_\sigma$$

$$= -\frac{1}{\sqrt{4\,\pi}} \begin{pmatrix} \cos\theta & \sin\theta \exp[-i\varphi] \\ \sin\theta \exp[+i\varphi] & -\cos\theta \end{pmatrix} \chi_\sigma \tag{1746}$$

Thus, apart from an over all normalization factor, the four-component spinor Dirac wave function ψ is given by

$$\psi = \frac{\mathcal{N}}{\sqrt{4\,\pi}}\, \rho^{\sqrt{1-\gamma^2}-1}\, \exp\left[-\rho\right] \begin{pmatrix} \chi_\sigma \\ -i\left(\frac{r \cdot \sigma}{r}\right)\chi_\sigma \end{pmatrix} \tag{1747}$$

298
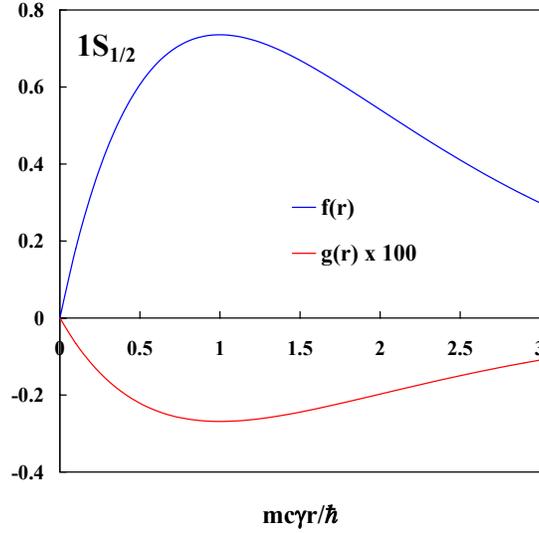
Figure 50: The large $f(r)$ and small component $g(r)$ radial wave functions for the $1S_{\frac{1}{2}}$ ground state of Hydrogen.

Hence, it is seen that as ρ approaches the origin, at first the wave function is slowly varying since

$$\rho^{\sqrt{1-\gamma^2}-1} \;\sim\; \exp\left[\; - \; \frac{\gamma^2}{2} \; \ln\rho \; \right] \;\sim\; 1 \; - \; \frac{\gamma^2}{2} \; \ln\rho \qquad (1748)$$

but for distances smaller than the characteristic length scale

$$r_c \; = \; \frac{\hbar}{m\,c\,\gamma} \; \exp\left[\; - \; \frac{2}{\gamma^2} \; \right] \qquad (1749)$$

the wave function exhibits a slight singularity. This length scale is much smaller that the nuclear radius so, due to the spatial distribution of the nuclear charge, the singularity is largely irrelevant. This singularity is not present in the non-relativistic limit, since in this limit one assumes that the inequality $\mid V(r) \mid \ll m\,c^2$ always holds, although this assumption is invalid for $r \sim 0$. Therefore, one concludes that the relativistic theory differs from the non-relativistic theory at small distances, which could have been discerned from the use of the Heisenberg uncertainty principle.
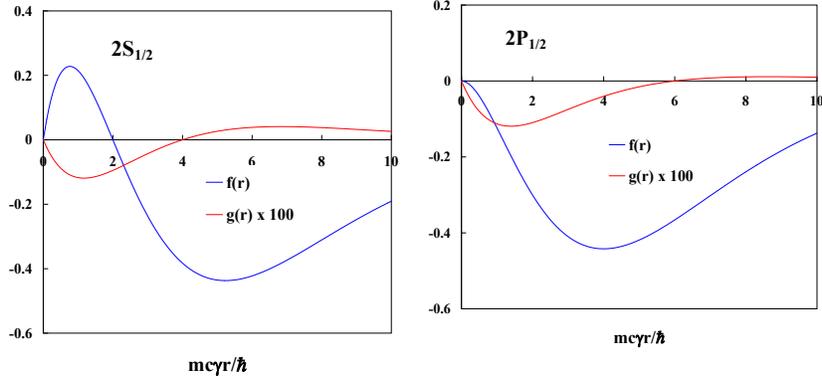
Figure 51: The radial wave functions for the $2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ states of Hydrogen.

12.10.2 Lowest-Order Radial Wavefunctions

The first few radial functions for the hydrogen atom can be expressed in the form

$$
\begin{aligned}
f(r) &= \mathcal{N} \sqrt{1 + \left(\frac{E}{m\,c^2} \right)} \left(\frac{r}{a} \right)^s \exp\left[-\frac{r}{a} \right] \left[c_0 - 2\,c_1 \left(\frac{r}{a} \right) \right] \\
g(r) &= -\mathcal{N} \sqrt{1 - \left(\frac{E}{m\,c^2} \right)} \left(\frac{r}{a} \right)^s \exp\left[-\frac{r}{a} \right] \left[d_0 - 2\,d_1 \left(\frac{r}{a} \right) \right]
\end{aligned}
\tag{1750}
$$

where the above form is restricted to the case where the radial quantum number n_r take on the values of 0 or 1. The index s is the same as that which occurs in the Frobenius method and is given by the positive solution

$$
s = \sqrt{\kappa^2 - \gamma^2}
\tag{1751}
$$

It is seen that the radial wavefunction depend on the dimensionless variable ρ defined by

$$
\rho = \frac{r}{a}
\tag{1752}
$$

where the length scale a is given in terms of the energy E and the Compton wavelength by

$$
a = \left[1 - \left(\frac{E}{m\,c^2} \right)^2 \right]^{-\frac{1}{2}} \left(\frac{\hbar}{m\,c} \right)
\tag{1753}
$$

The values of the indices s, energy E, length scale a and normalization $\mathcal{N}$ are given in Table(13). Since the two-component spinor spherical harmonics $\Omega^l_{j,j_z}(\theta, \varphi)$ are normalized to unity, the normalization condition is determined from the integral

$$
\mathcal{N}^2 \int_0^\infty dr \left(\mid f \mid^2 + \mid g \mid^2 \right) = 1
\tag{1754}
$$

300

involving the radial wave functions. The integral is evaluated with the aid of the identity

$$\int_0^\infty d\rho \, \rho^{a+b} \, \exp\left[-2\rho\right] = 2^{-(a+b+1)} \, \Gamma(a+b+1) \qquad (1755)$$

The coefficients c_n and d_n in the above expansion of the radial functions differ

Table 13: Parameters specifying the Radial Functions for the Hydrogen atom.

State	s	$\left(\dfrac{E}{m\,c^2}\right)$	$a\left(\dfrac{\gamma\,m\,c}{\hbar}\right)$	$\mathcal{N}$
$\kappa = -1$ $1S_{\frac{1}{2}}$	$\sqrt{1-\gamma^2}$	$\sqrt{1-\gamma^2}$	1	$\dfrac{1}{\sqrt{a}}\dfrac{2^{s+\frac{1}{2}}}{\sqrt{2\Gamma(2s+1)}}$
$\kappa = -1$ $2S_{\frac{3}{2}}$	$\sqrt{1-\gamma^2}$	$\sqrt{\dfrac{1+\sqrt{1-\gamma^2}}{2}}$	$2\left(\dfrac{E}{m\,c^2}\right)$	$\dfrac{1}{2}\sqrt{\dfrac{\left(\frac{2\,E}{m\,c^2}-1\right)}{\left(\frac{2\,E}{m\,c^2}\right)}}\dfrac{1}{\sqrt{a}}\dfrac{2^{s+\frac{1}{2}}}{\sqrt{\Gamma(2s+1)}}$
$\kappa = 1$ $2P_{\frac{1}{2}}$	$\sqrt{1-\gamma^2}$	$\sqrt{\dfrac{1+\sqrt{1-\gamma^2}}{2}}$	$2\left(\dfrac{E}{m\,c^2}\right)$	$\dfrac{1}{2}\sqrt{\dfrac{\left(\frac{2\,E}{m\,c^2}+1\right)}{\left(\frac{2\,E}{m\,c^2}\right)}}\dfrac{1}{\sqrt{a}}\dfrac{2^{s+\frac{1}{2}}}{\sqrt{\Gamma(2s+1)}}$
$\kappa = -2$ $2P_{\frac{3}{2}}$	$\sqrt{4-\gamma^2}$	$\sqrt{1-\dfrac{\gamma^2}{4}}$	2	$\dfrac{1}{\sqrt{a}}\dfrac{2^{s+\frac{1}{2}}}{\sqrt{2\Gamma(2s+1)}}$

from the coefficients a_n and b_n that occur in the Frobenius expansion, since the values of the ratio c_{n_r}/d_{n_r} has been chosen to simplify in the limit of large n. In particular at the value of n_r (at which the series terminates), the ratio is chosen to satisfy

$$\frac{c_{n_r}}{d_{n_r}} = 1 \qquad (1756)$$

instead of the condition

$$\frac{a_{n_r}}{b_{n_r}} = -\sqrt{\frac{1+\left(\frac{E}{m\,c^2}\right)}{1-\left(\frac{E}{m\,c^2}\right)}} \qquad (1757)$$

301

The relative negative sign and the square root factors in the coefficients have been absorbed into the expressions for the upper and lower components $f(r)$ and $g(r)$. The square root factors are responsible for converting the upper and lower components, respectively, into the large and small components for positive E, and vice versa for negative E. The expansion coefficients are given in Table(14). Since the ratio of the magnitudes of the polynomial factors is generally of the

Table 14: Coefficients for the Polynomial in the Hydrogen atom Radial Wavefunctions.

State	c_0	c_1	d_0	d_1
$\kappa = -1$ $1S_{\frac{1}{2}}$	1	0	1	0
$\kappa = -1$ $2S_{\frac{3}{2}}$	$2\left(\frac{E}{m\,c^2}\right)$	$\left[\frac{(\frac{2\,E}{m\,c^2})+1}{2\,s\,+\,1}\right]$	$2\left[\left(\frac{E}{m\,c^2}\right)\,+\,1\right]$	$\left[\frac{(\frac{2\,E}{m\,c^2})+1}{2\,s\,+\,1}\right]$
$\kappa = 1$ $2P_{\frac{1}{2}}$	$2\left[\left(\frac{E}{m\,c^2}\right)\,-\,1\right]$	$\left[\frac{(\frac{2\,E}{m\,c^2})-1}{2\,s\,+\,1}\right]$	$2\left(\frac{E}{m\,c^2}\right)$	$\left[\frac{(\frac{2\,E}{m\,c^2})-1}{2\,s\,+\,1}\right]$
$\kappa = -2$ $2P_{\frac{3}{2}}$	1	0	1	0

order of unity, the ratio of the magnitudes of the small to large components is found to be of the order of γ.

12.10.3   The Relativistic Corrections for Hydrogen

The Dirac equation for Hydrogen will be examined in the non-relativistic limit, and the lowest-order relativistic corrections will be retained. The resulting equation will be recast in the form of a Schrödinger equation, in which the Hamiltonian contains additional interaction terms. The resulting interactions, when treated by first-order perturbation theory, yield the fine structure. The physical interpretation of the interactions will be examined. Historically, the following type of analysis and the ensuing discussion of the Thomas precession played a

decisive role in compelling Pauli to reluctantly accept Dirac's theory.

The Dirac equation can be expressed as the set of coupled equations

$$\left(i \, \hbar \frac{\partial}{\partial t} \, - \, V \, - \, m \, c^2 \right) \phi^A \; = \; - \, i \, \hbar \, c \, (\, \underline{\sigma} \, . \, \underline{\nabla} \,) \, \phi^B$$

$$\left(i \, \hbar \frac{\partial}{\partial t} \, - \, V \, + \, m \, c^2 \right) \phi^B \; = \; - \, i \, \hbar \, c \, (\, \underline{\sigma} \, . \, \underline{\nabla} \,) \, \phi^A \qquad (1758)$$

where ϕ^A and ϕ^B are, respectively, the upper and lower two component spinors of the four-component Dirac spinor ψ. The energy eigenvalues of these equations are sought, so to this end the explicit time-dependence of the energy eigenstates will be separated out via

$$\psi \; = \; \left(\begin{array}{c} \phi^A \\ \phi^B \end{array} \right) \, \exp \left[\, - \, \frac{i}{\hbar} \, E \, t \, \right] \qquad (1759)$$

Also the non-relativistic energy ϵ will be defined as the energy referenced with respect to the rest-mass energy

$$E \; = \; m \, c^2 \, + \, \epsilon \qquad (1760)$$

The coupled equations reduce to

$$\left(\epsilon \, - \, V \right) \phi^A \; = \; - \, i \, \hbar \, c \, (\, \underline{\sigma} \, . \, \underline{\nabla} \,) \, \phi^B$$

$$\left(\epsilon \, - \, V \, + \, 2 \, m \, c^2 \right) \phi^B \; = \; - \, i \, \hbar \, c \, (\, \underline{\sigma} \, . \, \underline{\nabla} \,) \, \phi^A \qquad (1761)$$

The pair of equations will be expanded in powers of $(\frac{p}{m\,c})^2$ and only the first-order relativistic corrections will be retained. One can express ϕ^B as

$$\begin{aligned} \phi^B \; &= \; \frac{- \, i \, \hbar \, c \, (\, \underline{\sigma} \, . \, \underline{\nabla} \,) \, \phi^A}{\epsilon \, - \, V \, + \, 2 \, m \, c^2} \\ &= \; \frac{1}{2 \, m \, c} \left[1 \, + \, \frac{\epsilon \, - \, V}{2 \, m \, c^2} \right]^{-1} (\, \underline{\sigma} \, . \, \hat{\underline{p}} \,) \, \phi^A \\ &\approx \; \frac{1}{2 \, m \, c} \left(1 \, - \, \frac{\epsilon \, - \, V}{2 \, m \, c^2} \, + \, \ldots \right) (\, \underline{\sigma} \, . \, \hat{\underline{p}} \,) \, \phi^A \qquad (1762) \end{aligned}$$

to the required order of approximation. The above equation can be used to obtain a Schrödinger-like equation for the two-component spinor ϕ^A. Since a Schrödinger equation is sought for ψ_S, a correspondence must be established between the pair of spinors (ϕ^A, ϕ^B) and ψ_S. The probability density is the physical quantity which is directly associated with both types of wave functions. The probability density associated with the Schrödinger equation should be equivalent to the probability density associated with the Dirac equation. The

303

probability density associated with the four-component Dirac spinor depends on both ϕ^A and ϕ^B,

$$P(\underline{r}) \; = \; \left(\; \phi^{A\dagger} \, \phi^A \; + \; \phi^{B\dagger} \, \phi^B \; \right) \tag{1763}$$

The probability density associated with the two-component Schrödinger wave function depends on ψ_S

$$P(\underline{r}) \; = \; \psi_S^{\dagger} \, \psi_S \tag{1764}$$

The probability density is normalized to unity. On equating the two expressions for the normalization and substituting for ϕ^B, one obtains

$$\begin{aligned}
\int d^3\underline{r} \, \psi_S^{\dagger} \, \psi_S \; &= \; \int d^3\underline{r} \, \left(\; \phi^{A\dagger} \, \phi^A \; + \; \frac{1}{4 \, m^2 \, c^2} \, (\, \underline{\sigma} \cdot \underline{\hat{p}} \, \phi^A \,)^{\dagger} \, (\, \underline{\sigma} \cdot \underline{\hat{p}} \, \phi^A \,) \; \right) \\
&= \; \int d^3\underline{r} \, \left(\; \phi^{A\dagger} \, \phi^A \; + \; \frac{1}{4 \, m^2 \, c^2} \, \phi^{A\dagger} \, (\, \underline{\sigma} \cdot \underline{\hat{p}} \,) \, (\, \underline{\sigma} \cdot \underline{\hat{p}} \,) \, \phi^A \; \right) \\
&= \; \int d^3\underline{r} \, \phi^{A\dagger} \, \left(\; I \; + \; \frac{\hat{p}^2}{4 \, m^2 \, c^2} \; \right) \, \phi^A \tag{1765}
\end{aligned}$$

Therefore, the two-component Schrödinger wave function can be identified as

$$\psi_S \; = \; \left(\; I \; + \; \frac{\hat{p}^2}{8 \, m^2 \, c^2} \; + \; \dots \; \right) \, \phi^A \tag{1766}$$

or, on inverting the expansion

$$\phi^A \; \approx \; \left(\; I \; - \; \frac{\hat{p}^2}{8 \, m^2 \, c^2} \; + \; \dots \; \right) \, \psi_S \tag{1767}$$

Expressing ϕ_A in terms of ψ_S in the equation for ϕ^B yields the equation

$$\begin{aligned}
\phi^B \; &\approx \; \frac{1}{2 \, m \, c} \, \left(1 \; - \; \frac{\epsilon \, - \, V}{2 \, m \, c^2} \right) \, (\, \underline{\sigma} \cdot \underline{\hat{p}} \,) \, \left(I \; - \; \frac{\hat{p}^2}{8 \, m^2 \, c^2} \right) \, \psi_S \\
&\approx \; \frac{1}{2 \, m \, c} \, \left[\, (\, \underline{\sigma} \cdot \underline{\hat{p}} \,) \, \left(I \; - \; \frac{\hat{p}^2}{8 \, m^2 \, c^2} \right) \; - \; \left(\frac{\epsilon \, - \, V}{2 \, m \, c^2} \right) \, (\, \underline{\sigma} \cdot \underline{\hat{p}} \,) \, \right] \, \psi_S
\end{aligned} \tag{1768}$$

On substituting ϕ^B and ψ_S into the equation for ϕ^A, one finds the (two-component) energy eigenvalue equation

$$\begin{aligned}
\left(\; \epsilon \; - \; V \; \right) \, &\left(I \; - \; \frac{\hat{p}^2}{8 \, m^2 \, c^2} \right) \, \psi_S \\
&= \; \frac{(\, \underline{\sigma} \cdot \underline{\hat{p}} \,)}{2 \, m} \, \left[\, (\, \underline{\sigma} \cdot \underline{\hat{p}} \,) \, \left(I \; - \; \frac{\hat{p}^2}{8 \, m^2 \, c^2} \right) \; - \; \left(\frac{\epsilon \, - \, V}{2 \, m \, c^2} \right) \, (\, \underline{\sigma} \cdot \underline{\hat{p}} \,) \, \right] \, \psi_S
\end{aligned} \tag{1769}$$

or

$$\left(\epsilon \; - \; V \; - \; \epsilon \; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; + \; V \; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; \right) \psi_S$$
$$= \; \left[\; \frac{\hat{p}^2}{2 \; m} \left(\; I \; - \; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; \right) \; - \; (\; \underline{\sigma} \cdot \hat{\underline{p}} \;) \left(\frac{\epsilon \; - \; V}{4 \; m^2 \; c^2} \right) (\; \underline{\sigma} \cdot \hat{\underline{p}} \;) \; \right] \psi_S$$

$$(1770)$$

The above energy eigenvalue equation can be expressed as

$$\left(\epsilon \; - \; V \; - \; \frac{\hat{p}^2}{2 \; m} \; + \; \epsilon \; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; + \; V \; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; \right) \psi_S$$
$$= \; \left[\; - \; \frac{\hat{p}^4}{16 \; m^3 \; c^2} \; + \; (\; \underline{\sigma} \cdot \hat{\underline{p}} \;) \left(\frac{V}{4 \; m^2 \; c^2} \right) (\; \underline{\sigma} \cdot \hat{\underline{p}} \;) \; \right] \psi_S \quad (1771)$$

The term proportional to the product of the energy eigenvalue ϵ and the kinetic energy can be re-written as

$$\epsilon \; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; \psi_S \;\; = \;\; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; \epsilon \; \psi_S$$
$$\approx \;\; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; (\; V \; + \; \frac{\hat{p}^2}{2 \; m} \;) \; \psi_S \qquad (1772)$$

to the required order of approximation. On substituting the above expression into the energy eigenvalue equation (1771), one finds

$$\left(\epsilon \; - \; V \; - \; \frac{\hat{p}^2}{2 \; m} \; + \; \frac{\hat{p}^4}{8 \; m^3 \; c^2} \; + \; V \; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; + \; \frac{\hat{p}^2}{8 \; m^2 \; c^2} \; V \; \right) \psi_S$$
$$= \; \left[\; (\; \underline{\sigma} \cdot \hat{\underline{p}} \;) \left(\frac{V}{4 \; m^2 \; c^2} \right) (\; \underline{\sigma} \cdot \hat{\underline{p}} \;) \; \right] \psi_S \qquad (1773)$$

The above equation will be interpreted as the non-relativistic energy eigenvalue equation for the two-component wave function ψ_S, which contains relativistic corrections of order $(\frac{v}{c})^2$. The energy eigenvalue equation (1773) will be written in the form

$$\epsilon \; \psi_S \;\; = \;\; \left[\; \frac{\hat{p}^2}{2 \; m} \; + \; V \; \right] \psi_S \; - \; \frac{\hat{p}^4}{8 \; m^3 \; c^2} \; \psi_S$$
$$- \; \left(\frac{\hat{p}^2 \; V \; + \; V \; \hat{p}^2}{8 \; m^2 \; c^2} \right) \psi_S \; + \; \left[\; (\; \underline{\sigma} \cdot \hat{\underline{p}} \;) \left(\frac{V}{4 \; m^2 \; c^2} \right) (\; \underline{\sigma} \cdot \hat{\underline{p}} \;) \; \right] \psi_S$$

$$(1774)$$

where the relativistic corrections are symmetric in p^2 and V. This represents the energy eigenvalue equation for a two-component wave function ψ_S, similar to the Schrödinger wave function, but the above equation does include relativistic corrections to the Hamiltonian. The first correction term is

$$\hat{H}_{\text{Kin}} \;\; = \;\; - \; \frac{\hat{p}^4}{8 \; m^3 \; c^2} \qquad (1775)$$

305

which is recognized as the relativistic kinematic energy correction, that originates from the expansion of the kinetic energy

$$\begin{aligned} \epsilon &= \sqrt{m^2\,c^2\,+\,p^2\,c^2}\,-\,m\,c^2 \\ &\approx \frac{p^2}{2\,m}\,-\,\frac{p^4}{8\,m^3\,c^2}\,+\,\dots \end{aligned} \tag{1776}$$

The remaining two correction terms

$$\left[\,(\,\underline{\sigma}\cdot\hat{\underline{p}}\,)\left(\frac{V}{4\,m^2\,c^2}\right)(\,\underline{\sigma}\cdot\hat{\underline{p}}\,)\,\right]\,-\,\left(\frac{\hat{p}^2\,V\,+\,V\,\hat{p}^2}{8\,m^2\,c^2}\right) \tag{1777}$$

will be interpreted as the sum of the spin-orbit interaction and the Darwin term. It should be noted that the sum of these two terms would identically cancel in a purely classical theory. This cancellation can be shown to occur since, in the classical limit, V and p commute, and then the Pauli-identity can be used to show that the resulting pairs of terms cancel.

The factor

$$2\,(\,\underline{\sigma}\cdot\hat{\underline{p}}\,)\,V\,(\,\underline{\sigma}\cdot\hat{\underline{p}}\,)\,-\,\left(\hat{p}^2\,V\,+\,V\,\hat{p}^2\right) \tag{1778}$$

can be evaluated as

$$2\,\hat{\underline{p}}\cdot V\,\hat{\underline{p}}\,-\,\left(\hat{p}^2\,V\,+\,V\,\hat{p}^2\right)\,+\,2\,i\,\underline{\sigma}\cdot\left(\hat{\underline{p}}\,\wedge\,V\,\hat{\underline{p}}\right) \tag{1779}$$

The first two terms can be combined to form a double commutator, yielding

$$-\,[\,\hat{\underline{p}}\,,\,[\,\hat{\underline{p}}\,,\,V\,]\,]\,+\,2\,i\,\underline{\sigma}\cdot\left(\hat{\underline{p}}\,\wedge\,V\,\hat{\underline{p}}\right) \tag{1780}$$

or

$$+\,\hbar^2\,\nabla^2\,V\,+\,2\,i\,\underline{\sigma}\cdot\left(\hat{\underline{p}}\,\wedge\,V\,\hat{\underline{p}}\right) \tag{1781}$$

The last term can be evaluated, resulting in the expression

$$+\,\hbar^2\,\nabla^2\,V\,+\,2\,\hbar\,\underline{\sigma}\cdot\left(\underline{\nabla}\,V\,\wedge\,\hat{\underline{p}}\right) \tag{1782}$$

since

$$\hat{\underline{p}}\,\wedge\,\hat{\underline{p}}\,\equiv\,0 \tag{1783}$$

Using these substitutions, the remaining interactions can be expressed as the sum of the spin-orbit interaction and the Darwin interaction

$$\hat{H}_{SO}\,+\,\hat{H}_{\text{Darwin}}\,=\,+\,\frac{\hbar}{4\,m^2\,c^2}\,\underline{\sigma}\cdot\left(\underline{\nabla}\,V\,\wedge\,\hat{\underline{p}}\right)\,+\,\frac{\hbar^2}{8\,m^2\,c^2}\,\nabla^2\,V \tag{1784}$$

The first term is the spin-orbit interaction term, and the second term is the Darwin term. For central potentials, the Darwin term is only important for electrons with $l = 0$. The evaluation and the physical interpretation of the energy shifts due to the three fine-structure interactions will be discussed separately.

12.10.4 The Kinematic Correction

The kinematic interaction

$$\hat{H}_{\text{Kin}} = - \frac{\hat{p}^4}{8 \, m^3 \, c^2} \tag{1785}$$

originates from the expansion of the relativistic expression for the kinetic energy of a classical particle

$$\begin{aligned} \epsilon &= \sqrt{m^2 \, c^2 + p^2 \, c^2} - m \, c^2 \\ &\approx \frac{p^2}{2 \, m} - \frac{p^4}{8 \, m^3 \, c^2} + \ldots \end{aligned} \tag{1786}$$

The first-order energy shift due to the kinematic correction $\hat{H}_{\text{Kin}}$ can be evaluated by using the solution to the non-relativistic Schrödinger equation

$$\frac{\hat{p}^2}{2 \, m} \, \psi_S = \left[- \frac{m \, c^2}{2 \, n^2} \left(\frac{Z \, e^2}{\hbar \, c} \right)^2 + \frac{Z \, e^2}{r} \right] \psi_S \tag{1787}$$

which leads to

$$\begin{aligned} \Delta E_{\text{Kin}} &= - \int d^3\underline{r} \, \psi_S^\dagger(r) \, \frac{\hat{p}^4}{8 \, m^3 \, c^2} \, \psi_S(\underline{r}) \\ &= - \frac{1}{2 \, m \, c^2} \int d^3\underline{r} \, \psi_S^\dagger(r) \left[- \frac{m \, c^2}{2 \, n^2} \left(\frac{Z \, e^2}{\hbar \, c} \right)^2 + \frac{Z \, e^2}{r} \right]^2 \psi_S(\underline{r}) \\ &= m \, c^2 \left(\frac{Z \, e^2}{\hbar \, c} \right)^4 \frac{3}{8 \, n^4} - \frac{Z^2 \, e^4}{2 \, m \, c^2} \int d^3\underline{r} \, \psi_S^\dagger(\underline{r}) \, \frac{1}{r^2} \, \psi_S(\underline{r}) \end{aligned} \tag{1788}$$

Hence, the first-order energy shift due to the kinematic correction is evaluated as

$$\Delta E_{\text{Kin}} = m \, c^2 \left(\frac{Z \, e^2}{\hbar \, c} \right)^4 \left[\frac{3}{8 \, n^4} - \frac{1}{n^3 \, (2 \, l + 1)} \right] \tag{1789}$$

This term is found to lift the degeneracy between states with fixed principle quantum numbers n and values of the angular momenta l. The relativistic kinematic correction to the energy is found to be smaller than the non-relativistic energy by a factor of

$$\left(\frac{Z \, e^2}{\hbar \, c} \right)^2 \sim Z^2 \times 10^{-4} \tag{1790}$$

which can be identified with a factor of $(\frac{v}{c})^2$ as can be inferred from an analysis based on the Bohr model of the atom. One sees that the relativistic corrections become more important for atoms with larger Z, since the correction varies as Z^4. This occurs because for larger Z the electrons are drawn closer to the nucleus and, hence have higher kinetic energies, so the electron's velocities draw closer to the velocity of light.

12.10.5   Spin-Orbit Coupling

To elucidate the meaning of the spin-orbit interaction, the interaction will be re-derived starting from quasi-classical considerations of the anomalous Zeeman interaction of a spin with a magnetic field.

Consider a particle moving with a velocity $\underline{v}$ in a static electric field $\underline{E}$. In the particle's rest frame, it will experience a magnetic field $\underline{B}'$ which is given by

$$\underline{B}' \;=\; -\,\frac{1}{c}\,\frac{\underline{v}\,\wedge\,\underline{E}}{\sqrt{1\,-\,\frac{v}{c}^2}} \;\approx\; \frac{1}{c}\,\underline{E}\,\wedge\,\underline{v} \tag{1791}$$

for small velocities v. The magnetic field B' is a relativistic correction due to the motion of the source of the electric field. If an electron is moving in a central electrostatic potential $\phi(r)$ caused by a charged nucleus, the radial electric field is given by

$$\underline{E} \;=\; -\,\frac{\underline{r}}{r}\,\left(\frac{\partial \phi}{\partial r}\right) \tag{1792}$$

Hence, the magnetic field experienced by an electron in its rest frame is given by

$$
\begin{aligned}
\underline{B}' \;&=\; -\,\frac{1}{m\,c\,r}\,\left(\frac{\partial \phi}{\partial r}\right)\,\underline{r}\,\wedge\,\underline{p} \\
&=\; -\,\frac{1}{m\,c\,r}\,\left(\frac{\partial \phi}{\partial r}\right)\,\underline{L}
\end{aligned}
\tag{1793}
$$

which is caused by the apparent rotation of the charged nucleus. In the electron's rest frame, the electron's spin $\underline{S}$ should interact with the magnetic field through the Zeeman interaction

$$\hat{H}_{Int}^{\text{rest}} \;=\; -\,\frac{q}{2\,m\,c}\,g_S\,\underline{B}'\cdot\underline{S} \tag{1794}$$

where g_S is the gyromagnetic ratio for the electron's spin. Dirac's theory predicts that the spin is a relativistic phenomenon and also that $g_S = 2$ for an electron in its rest frame. This interaction with the magnetic field will cause the spin of the electron to precess. The spin precession rate found in the electrons rest frame is calculated as

$$\omega_{\text{rest}} \;=\; \frac{e}{2\,m\,c}\,g_S\,B' \tag{1795}$$

However, the electron is bound to the nucleus and is orbiting with angular momentum $\underline{L}$. Therefore, one has to consider the corrections to the precession rate (and the interaction) caused by the acceleration of the electron's rest frame.

Thomas Precession

Electrons exhibit two different gyromagnetic ratios. The gyromagnetic ratio of $g_S = 2$ couples a spin to an external magnetic field and there is a gyromagnetic ratio of unity for the lab frame. This gyromagnetic ratio of unity (in the lab frame) enters the coupling between the spin of an electron in a circular orbit to the magnetic field B' experienced in the electron's rest frame[113]. We shall find the gyromagnetic ratio in the lab frame, by calculating the rate of precession that is observed in the lab frame and then inferring the (lab frame) interaction which produces the same rate of precession.

In the electron's rest frame, the gyromagnetic ratio due to the orbital magnetic field B' (caused by the charged nucleus) is given by $g_s = 2$. This gyromagnetic ratio yields a spin precession rate in the electron's rest frame of

$$\omega_{\text{rest}} \;=\; \frac{e}{2\,m\,c}\; g_S\; B' \tag{1796}$$

The spin precession rate observed in the lab frame will be calculated later. The rate of precession as observed in the electron's rest frame has to be corrected by taking into account the motion of the electron. The correction is due to the non-additivity of velocities in successive Lorentz transformations. First, the transformation properties of Dirac spinors under infinitesimal rotations and boosts will be re-examined. Secondly, infinitesimal transformations will be successively applied to describe the particle's instantaneous rest frame and the Thomas precession.

A Lorentz transform of a spinor field ψ is achieved by the rotation operator $\hat{\mathcal{R}}$ via

$$\psi'(\underline{r}) \;=\; \hat{\mathcal{R}}\; \psi(R^{-1}\underline{r}) \tag{1797}$$

where $\hat{\mathcal{R}}$ shuffles the components of the spinor. For a passive rotation (of the coordinate system) through the infinitesimal angle $\delta\varphi$ in the i - j plane, the infinitesimal Lorentz transform has the non-zero elements

$$\epsilon_{i,j} \;=\; -\,\epsilon_{j,i} \;=\; -\,\delta\varphi \tag{1798}$$

Hence, the four-component spinor is transformed by a rotation operator of the form

$$\hat{\mathcal{R}}(\delta\varphi) \;=\; \exp\left[\; +\; i\,\frac{\delta\varphi}{2}\,\sigma^{i,j}\;\right] \tag{1799}$$

[113]L. H. Thomas, Nature **117**, 514 (1926).
L. H. Thomas, Phil. Mag. **3**, 1 (1927).

where

$$\sigma^{i,j} = \frac{i}{2} \left[\gamma^{(i)} , \gamma^{(j)} \right]$$

$$= \sum_k \xi^{i,j,k} \begin{pmatrix} \sigma^{(k)} & 0 \\ 0 & \sigma^{(k)} \end{pmatrix} \qquad (1800)$$

Hence, for a passive rotation through an infinitesimal angle $\delta\varphi$, the four-component Dirac spinor is rotated by

$$\hat{\mathcal{R}}(\delta\varphi) = \hat{I} + i \frac{\delta\varphi}{2} \sum_k \xi^{i,j,k} \begin{pmatrix} \sigma^{(k)} & 0 \\ 0 & \sigma^{(k)} \end{pmatrix} + \dots \qquad (1801)$$

which can be expressed in terms of the projection of the (block diagonal) spin operator $\underline{\hat{S}} = \frac{\hbar}{2} \, \underline{\hat{\sigma}}$ on the axis of rotation $\hat{e}$ as

$$\hat{\mathcal{R}}(\delta\varphi) = \hat{I} + \frac{i \, \delta\varphi}{2} \, (\hat{e} \cdot \underline{\hat{\sigma}}) + \dots$$

$$= \hat{I} + \frac{i \, \delta\varphi}{\hbar} \, (\hat{e} \cdot \underline{\hat{S}}) + \dots \qquad (1802)$$

which is in accord with the definition of spin $\underline{\hat{S}}$ as the generator of rotations.

If the primed frame of reference has a velocity v along the k-axis relative to the un-primed frame, the infinitesimal Lorentz transform has the non-zero elements

$$\epsilon_{0,k} = - \epsilon_{k,0} = - \frac{v}{c} \qquad (1803)$$

A Lorentz boost along the k-axis corresponds to a rotation in the 0 - k plane through an "angle" χ

$$\hat{\mathcal{R}}(\chi) = \exp\left[+ i \frac{\chi}{2} \sigma^{0,k} \right] \qquad (1804)$$

where the "angle" χ is governed by the boost velocity v through

$$\tanh\chi = \frac{v}{c} \qquad (1805)$$

However,

$$\sigma^{0,k} = \frac{i}{2} \left[\gamma^{(0)} , \gamma^{(k)} \right] = i \, \alpha^{(k)} \qquad (1806)$$

so

$$\hat{\mathcal{R}}(\chi) = \exp\left[- \frac{\chi}{2} \alpha^{(k)} \right] \qquad (1807)$$

Therefore, for a Lorentz boost with an infinitesimal velocity v along the k-th direction, one finds

$$\hat{\mathcal{R}}(\chi) = \hat{I} - \frac{v}{2 \, c} \alpha^{(k)} + \dots \qquad (1808)$$

310

The infinitesimal transformation is guaranteed to be consistent with the source free solution of the Dirac equation. For example, if the above transformation is applied to the solution of the Dirac equation describing a positive-energy particle at rest, the transformed solution describes a particle moving with momentum $p = -mv$ when viewed from the moving frame of reference.
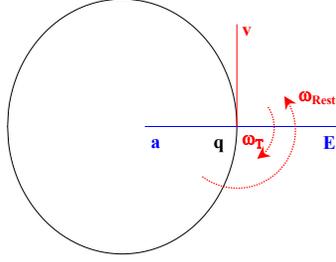

Figure 52: A cartoon depicting a rotating charged spin one-half particle, along with the precession of the spin due to the external field in the particle's rest frame and the Thomas precession.

Consider the rotation $\hat{\mathcal{R}}_1$ of a spinor due to an infinitesimal Lorentz transformation with "small" velocity $\underline{v}$, then

$$\hat{\mathcal{R}}_1 = \hat{I} - \frac{1}{2\,c}\,\underline{\alpha}\cdot\underline{v} + \dots \tag{1809}$$

At a time δt later, the electron has changed its velocity since it is accelerating. The new velocity of the electron's rest frame is given by

$$\underline{v}' = \underline{v} + \underline{a}\,\delta t \tag{1810}$$

On performing a second Lorentz transform with the boost $\underline{a}\,\delta t$, one finds the rotation

$$\hat{\mathcal{R}}_2 = \hat{I} - \frac{1}{2\,c}\,\underline{\alpha}\cdot\underline{a}\,\delta t + \dots \tag{1811}$$

The combined Lorentz transform is given by

$$\begin{aligned}
\hat{\mathcal{R}} &= \hat{\mathcal{R}}_2\,\hat{\mathcal{R}}_1 = \left(\hat{I} - \frac{1}{2\,c}\,\underline{\alpha}\cdot\underline{a}\,\delta t \right)\left(\hat{I} - \frac{1}{2\,c}\,\underline{\alpha}\cdot\underline{v} \right) + \dots \\
&= \hat{I} - \frac{1}{2\,c}\,\underline{\alpha}\cdot(\,\underline{v} + \underline{a}\,\delta t\,) + \frac{1}{4\,c^2}\,(\,\underline{\alpha}\cdot\underline{a}\,)\,(\,\underline{\alpha}\cdot\underline{v}\,)\,\delta t + \dots \tag{1812}
\end{aligned}$$

The Pauli identity can be used to evaluate the last term

$$\begin{aligned}
(\,\underline{\alpha}\cdot\underline{a}\,)\,(\,\underline{\alpha}\cdot\underline{v}\,) &= (\,\hat{\underline{\sigma}}\cdot\underline{a}\,)\,(\,\hat{\underline{\sigma}}\cdot\underline{v}\,) \\
&= \underline{a}\cdot\underline{v}\,\hat{I} + i\,\hat{\underline{\sigma}}\cdot(\,\underline{a}\wedge\underline{v}\,) \tag{1813}
\end{aligned}$$

where, since the product of the two α's yields a two by two block diagonal form which involves the four by four matrices $\hat{I}$ and $\hat{\sigma}$. Hence, the right-hand side acts equally on both the upper and lower two-component spinors. Furthermore, since the orbit is circular, the acceleration is perpendicular to the velocity, therefore

$$\underline{a} \cdot \underline{v} \; = \; 0 \tag{1814}$$

Thus, the combined boost corresponds to the transformation

$$\hat{\mathcal{R}} \; = \; \hat{I} \; - \; \frac{1}{2\,c} \; \underline{\alpha} \cdot (\, \underline{v} \; + \; \underline{a} \; \delta t \,) \; + \; \frac{i}{4\,c^2} \; \hat{\underline{\sigma}} \cdot (\, \underline{a} \, \wedge \, \underline{v} \,) \; \delta t \; + \; \ldots \tag{1815}$$

The combined boost is identified as producing an infinitesimal Lorentz boost through $\underline{v} \; + \; \underline{a} \; \delta t$ and a rotation around an axis $\hat{e}$ through the infinitesimal angle $\delta\varphi$ given by

$$\delta\varphi \; \hat{e} \; \approx \; \frac{1}{2\,c^2} \; (\, \underline{a} \, \wedge \, \underline{v} \,) \; \delta t \tag{1816}$$

The rotation part acts on both the upper and lower two-component spinors in the Dirac spinor. The rotation angle $\delta\varphi$ is linearly proportional to the time interval δt. This class of rotations due to the combination of Lorentz boosts are known as a Wigner rotations. Hence, it was shown that the spinor rotates with the angular velocity given by

$$\begin{aligned} \underline{\omega}_T \; &= \; \frac{1}{2\,c^2} \; (\, \underline{a} \, \wedge \, \underline{v} \,) \\ &= \; \frac{q}{2\,m\,c^2} \; (\, \underline{E} \, \wedge \, \underline{v} \,) \end{aligned} \tag{1817}$$

The magnitude of ω_T is calculated as

$$\omega_T \; = \; \frac{e}{2\,m\,c} \; B' \tag{1818}$$

and its direction is opposite to the precession of the spin in the electron's rest frame.

On combing the two precession frequencies, one finds that in the lab frame the spin's precession rate is given by

$$\begin{aligned} \omega_{\text{Lab}} \; &= \; \omega_{\text{rest}} \; - \; \omega_T \\ &= \; \frac{e}{2\,m\,c} \; (\, g_S \; - \; 1 \,) \; B' \end{aligned} \tag{1819}$$

It is clear that the moving spin experiences an effective interaction which is reduced by the factor

$$\left(\frac{g_S \; - \; 1}{g_S} \right) \tag{1820}$$

when compared to the interaction in the electron's rest frame. Hence, the gyromagnetic ratio that enters the spin-orbit coupling should not be g_S but should

be given by ($g_S - 1$).

The Spin-Orbit Interaction

In the lab frame, the interaction between the moving electron's spin $\underline{S}$ magnetic moment and its field is inferred to be

$$\hat{H}_{Int}^{\text{Lab}} = - \frac{q}{2\ m\ c}\ (\ g_S\ -\ 1\ )\ \underline{B}'\ .\ \underline{S} \tag{1821}$$

where g_S is the gyromagnetic ratio. Since the magnetic induction field is given by

$$\underline{B}' = - \frac{1}{m\ c\ r}\ \left(\frac{\partial \phi}{\partial r} \right)\ \underline{L} \tag{1822}$$

where the electrostatic potential is given by

$$\phi(r) = \frac{q'}{r} \tag{1823}$$

the spin-orbit interaction can be expressed as

$$\hat{H}_{\text{SO}} = - \frac{q}{2\ m\ c}\ (\ g_S\ -\ 1\ )\ \frac{q'}{m\ c\ r^3}\ \underline{L}\ .\ \underline{S} \tag{1824}$$

Hence, the spin-orbit interaction is found to be given by

$$\hat{H}_{\text{SO}} = \frac{Z\ e^2}{2\ m^2\ c^2\ r^3}\ (\ g_S\ -\ 1\ )\ \underline{L}\ .\ \underline{S} \tag{1825}$$

The spin-orbit coupling is a relativistic coupling which, apart from the Thomas precession factor, indicates that the electron's spin interacts with a magnetic field in its rest frame via the gyromagnetic ratio of 2. The magnitude of the interaction agrees precisely with the interaction found from the perturbative treatment of the Dirac equation.

To first-order in perturbation theory, the spin-orbit coupling interaction yields a shift of the energy levels. Since the total angular momentum $\underline{J}$ is a good quantum number, one can write

$$\underline{L}\ .\ \underline{S} = \frac{1}{2}\ \left(\ j\ (\ j\ +\ 1\ )\ -\ l\ (\ l\ +\ 1\ )\ -\ \frac{3}{4}\ \right) \tag{1826}$$

but j for a single electron can only take on the values $j\ =\ l\ \pm\ \frac{1}{2}$, so

$$\underline{L}\ .\ \underline{S} = \frac{1}{2}\ \left(\ \pm\ (\ l\ +\ \frac{1}{2}\ )\ -\ \frac{1}{2}\ \right) \tag{1827}$$

The expectation value of r^{-3} is evaluated as

$$\int d^3\underline{r}\ \psi_S^\dagger(\underline{r})\ \frac{1}{r^3}\ \psi_S(\underline{r}) = \frac{1}{l\ (\ l\ +\ \frac{1}{2}\ )\ (\ l\ +\ 1\ )}\ \left(\frac{Z}{n\ a} \right)^3 \tag{1828}$$

313

for $l \neq 0$. So the first-order energy-shift due to the spin-orbit coupling can be expressed as

$$\Delta E_{SO} = m c^2 \left(\frac{Z e^2}{\hbar c} \right)^4 \left[\frac{\pm (l + \frac{1}{2}) - \frac{1}{2}}{4 n^3 l (l + \frac{1}{2}) (l + 1)} \right] \tag{1829}$$

Therefore, the spin-orbit interaction lifts the degeneracy between states with different $j = l \pm \frac{1}{2}$ values. For $l = 0$, the numerator vanishes since the total angular momentum can only take the value

$$j = + \frac{1}{2} \tag{1830}$$

The energy shift produced by the spin-orbit coupling is about a factor of the square of the fine structure constant

$$\left(\frac{e^2}{\hbar c} \right)^2 \sim \left(\frac{1}{137} \right)^2 \sim 10^{-4} \tag{1831}$$

smaller than the energy levels of the hydrogen-like atom

$$E_n \approx - \frac{m c^2}{2 n^2} \left(\frac{Z e^2}{\hbar c} \right)^2 \tag{1832}$$

calculated using the non-relativistic Schrödinger equation. The spin-orbit split levels are labeled by the angular momentum values and the j values, and are denoted by nL_j. Hence, for $n = 2$ and $l = 1$, one has the two levels $2P_{\frac{1}{2}}$ and $2P_{\frac{3}{2}}$, while for $n = 3$ and $l = 2$ one has the levels $3D_{\frac{3}{2}}$ and $3D_{\frac{5}{2}}$, and so on. It is seen that the spin-orbit interaction is increasingly important for atoms with large Z values, as it varies like Z^4.

12.10.6   The Darwin Term

The Darwin term has no obvious classical interpretation. It only has physical consequences for states with zero orbital angular momentum. However, it does play an important role for the s electronic state of hydrogen, and is essential in describing why the Dirac's theory makes the $2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ states of hydrogen degenerate. This degeneracy was an essential ingredient in the discovery of the Lamb shift and the subsequent development of Quantum Electrodynamics.

The Darwin interaction is given by

$$\hat{H}_{\text{Darwin}} = \frac{\pi Z e^2 \hbar^2}{2 m^2 c^2} \delta^3(\underline{r}) \tag{1833}$$

which produces the first-order shift

$$\Delta E_{\text{Darwin}} = \frac{\pi Z e^2 \hbar^2}{2 m^2 c^2} \psi_S^{\dagger}(0) \psi_S(0) \tag{1834}$$

314

Hence, the shift only occurs for electrons with $l = 0$. Furthermore, since the probability density for finding the electron at the origin is given by

$$\psi_S^\dagger(0)\,\psi_S(0) \;=\; \frac{1}{\pi}\,\left(\frac{Z}{n\,a}\right)^3 \delta_{l,0} \tag{1835}$$

to first order, the Darwin term produces a shift

$$\Delta E_{\text{Darwin}} \;=\; \frac{m\,c^2\,Z^4}{2\,n^3}\,\left(\frac{e^2}{\hbar\,c}\right)^4 \delta_{l,0} \tag{1836}$$

which shifts the energies of s states upwards. The Darwin term reflects the fact that the relativistic corrections are important for small r since the inequality

$$m\,c^2 \;\gg\; \frac{Z\,e^2}{r} \tag{1837}$$

required for the non-relativistic treatment to be reasonable is violated in this region.

12.10.7   The Fine Structure of Hydrogen
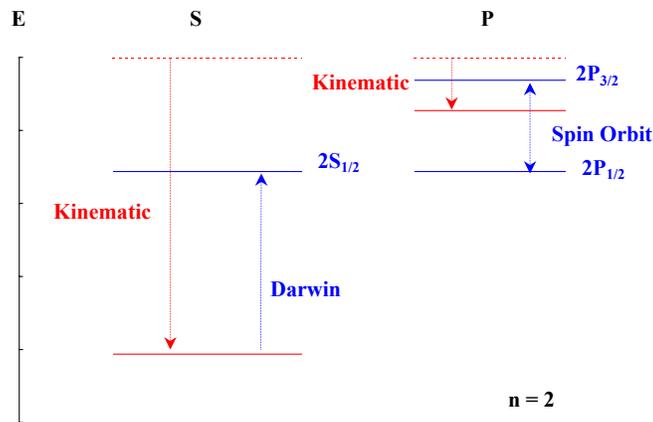
Figure 53: The Grotarian energy level diagram for the $n = 2$ shell of hydrogen (blue). The diagram shows the magnitude and sign of the various relativistic corrections. It should be noted that states with the same j are degenerate.

When the various relativistic corrections are combined, for $l = 0$, the Darwin term exactly compensates for the absence of the spin-orbit interaction. Therefore, the energy shifts combine to yield one formula in which l drops out. This implies that the energy levels only depend on the principle quantum number n

and the total angular momentum j. States with different orbital angular momenta are degenerate, even though the individual interactions appear to raise the degeneracy. The relativistic corrections inherent in Dirac's theory of hydrogen yields energy shifts and line-splittings which are described as fine structure. The energy levels are described by

$$ E \approx m c^2 \left[1 - \frac{1}{2} \frac{Z^2 \alpha^2}{n^2} - \frac{1}{2} \frac{Z^4 \alpha^4}{n^3} \left(\frac{1}{(j + \frac{1}{2})} - \frac{3}{4n} \right) + \dots \right] \tag{1838} $$

where

$$ \alpha = \left(\frac{e^2}{\hbar c} \right) \tag{1839} $$

is the fine structure constant. Generally, states with larger j values have higher energies. The fine structure splittings decrease with increasing n like n^{-3}, but increase with increasing Z like Z^4. The splitting of the lower energy levels are largest, for example

$$ E_{2P_{\frac{3}{2}}} - E_{2P_{\frac{1}{2}}} = - \frac{m c^2 \alpha^4}{16} \left(\frac{1}{2} - \frac{1}{1} \right) \approx 4.533 \times 10^{-5} \text{ eV} \tag{1840} $$

This splitting corresponds to a frequency of 10.96 GHz. The energy levels are predicted to be doubly degenerate (in addition to the degeneracy associated with j_3), the degeneracy is just the number of states with different l values that yield the same value of j. Since j is found by combining l with the electronic spin $s = \frac{1}{2}$, there are two possible l values for each energy level which are given by the solutions of either

$$ j = l + \frac{1}{2} \tag{1841} $$

or

$$ j = l - \frac{1}{2} \tag{1842} $$

The higher-order relativistic corrections do not alter the conclusion that the states labeled by (n, j) are degenerate, as the energy levels found from the exact solution of the Dirac equation only depend on n and j. For $j = \frac{1}{2}$ the energy levels, although predicted to be degenerate by Dirac's theory, are experimentally observed as being non-degenerate. The first experiments that revealed this splitting were performed by Lamb and Retherford[114]. These scientists found that the $2S_{\frac{1}{2}}$ was shifted by about 1057 MHz to higher energies relative to the $2P_{\frac{1}{2}}$. The relative shift of the $nS_{\frac{1}{2}}$ level of hydrogen with respect to the $nP_{\frac{1}{2}}$ level is known as the Lamb shift.

Lamb and Retherford's Experiment

[114]W. E. Lamb Jr. and R. E. Retherford, Phys. Rev. **72**, 241 (1947).
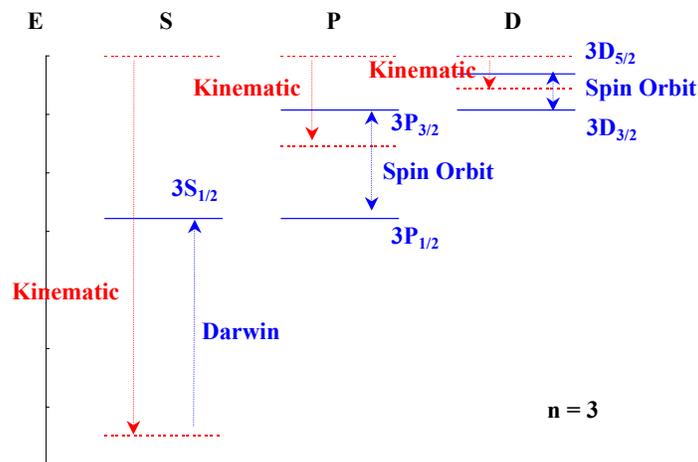
Figure 54: The Grotarian energy level diagram for the $n = 3$ shell of hydrogen (blue). The diagram shows the magnitude and sign of the various relativistic corrections. It should be noted that states with the same j are degenerate.

Lamb and Retherford designed an experiment to accurately measure the fine structure of the hydrogen atom. In the experiment, the time scales were such that the population of all excited states, other than the meta-stable $2S_{\frac{1}{2}}$ state of hydrogen, radiatively decayed to the ground state. Hence, the number of induced transitions from the $2S_{\frac{1}{2}}$ state could be monitored by simply observing of the population of hydrogen atoms not in the ground state.
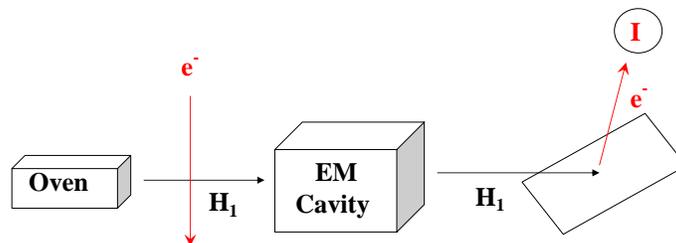
Figure 55: A schematic of the apparatus used in the Lamb-Retherford experiment. The beam of H molecules is produced in an oven, the beam is excited by cross-bombardment with an electron beam. The population of the $2S_{\frac{1}{2}}$ is altered in the microwave resonator, and the population is observed via the current emitted at the tungsten plate.

A beam of hydrogen atoms was produced by dissociating hydrogen molecules

in an oven. The thermal beam of hydrogen atoms was then cross-bombarded with electrons, which excited some of the hydrogen atoms out of the ground state. Since the electron-atom scattering doesn't obey the radiation selection rules, a finite population of atoms (about 1 in 10^8) were excited to the long-lived $2S_{\frac{1}{2}}$ state. Subsequently, the other excited electronic states rapidly decayed to the ground state by the emission of radiation. The beam of hydrogen atoms was then passed through a tuneable (microwave) electromagnetic resonator, which could cause the hydrogen atoms in the meta-stable level to make transitions to selected nearby energy levels. Again, any non-$2S$ excited state of hydrogen produced by the action of the resonator rapidly decayed to the ground state. The resulting beam of hydrogen atoms was incident on a Tungsten plate, and the collision could result in electron emission if the atoms were in an excited state, but no emission would take place if the hydrogen atom was in the ground state. Therefore, the current due to the emitted electrons was proportional to the number of meta-stable hydrogen atoms that survived the passage through the resonator. Hence, analysis of the experiment yielded the number of transitions undergone in the electromagnetic resonator.
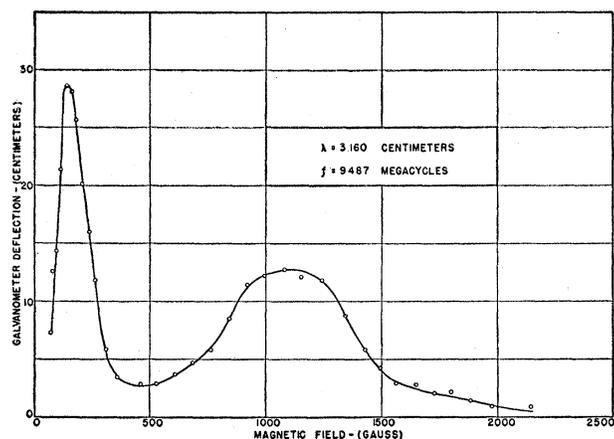
Figure 56: The dependence of the current emitted from the tungsten plate on the applied magnetic field. The resonance frequency was set to 9487 Megacycles. [W. E. Lamb Jr. and R. C. Retherford, Phys. Rev. **72**, 241 (1947).]

In the resonator, an applied magnetic field Zeeman split the excited levels of hydrogen and, when the oscillating field was on-resonance with the splitting of the energy levels, the hydrogen atom made transitions out from the meta-stable $2S_{\frac{1}{2}}$ state. At resonance, the frequency of the oscillating electromagnetic field is equal to the energy splitting. Therefore, for fixed frequency, knowledge of the resonance magnetic field allowed the splitting of the energy levels to be accu-

rately determined. The field dependence of the resonance frequency indicated
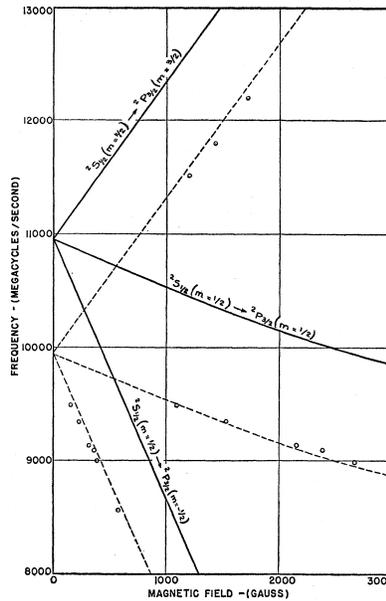
Figure 57: The observed dependence of the resonance frequencies on the applied magnetic field. The solid lines are the predictions of the Dirac theory and the dashed lines are the result of Dirac's theory if the energy of the $2S$ state is simply shifted. [W. E. Lamb Jr. and R. C. Retherford, Phys. Rev. **72**, 241 (1947).]

that at zero field the degeneracy between the $2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ states were lifted, with the $2S_{\frac{1}{2}}$ state having the higher energy.

12.10.8   A Particle in a Spherical Square Well

The radial equation for a relativistic spin one-half particle in a spherically symmetric "square well" potential is given by

$$(\, E \, - \, V(r) \, - \, m \, c^2 \,) \, f(r) \, + \, c \, \hbar \left(\, \frac{\partial}{\partial r} \, - \, \frac{\kappa}{r} \, \right) g(r) \; = \; 0$$

$$(\, E \, - \, V(r) \, + \, m \, c^2 \,) \, g(r) \, - \, c \, \hbar \left(\, \frac{\partial}{\partial r} \, + \, \frac{\kappa}{r} \, \right) f(r) \; = \; 0$$

$$(1843)$$

319

We shall examine the case of an attractive central square well potential $V(r)$ which is defined by

$$V(r) \;=\; \begin{cases} -\,V_0 & \text{for } r \;<\; a \\ 0 & \text{for } r \;>\; a \end{cases} \tag{1844}$$

In the region $r \;<\; a$ where the potential is finite, the Dirac radial equation
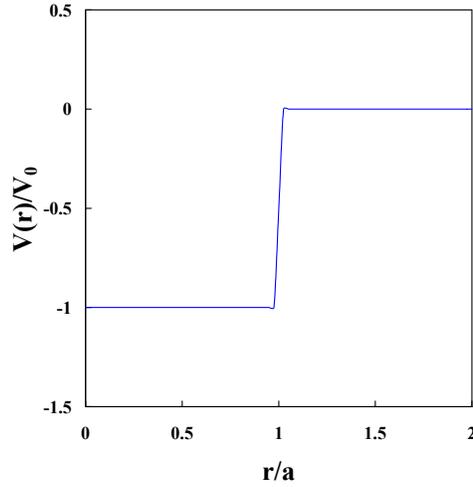
Figure 58: A spherically symmetric potential well, of depth V_0 and radius a.

becomes

$$(E \;+\; V_0 \;-\; m\,c^2)\,f(r) \;+\; c\,\hbar \left(\frac{\partial}{\partial r} \;-\; \frac{\kappa}{r} \right) g(r) \;=\; 0$$

$$(E \;+\; V_0 \;+\; m\,c^2)\,g(r) \;-\; c\,\hbar \left(\frac{\partial}{\partial r} \;+\; \frac{\kappa}{r} \right) f(r) \;=\; 0 \tag{1845}$$

The function $f(r)$ satisfies a second-order differential equation, which can be found by pre-multiplying the second equation by the operator

$$c\,\hbar \left(\frac{\partial}{\partial r} \;-\; \frac{\kappa}{r} \right) \tag{1846}$$

and then eliminating $g(r)$ by using the first equation. This process yields the equation

$$c^2\,\hbar^2 \left(\frac{\partial^2}{\partial r^2} \;-\; \frac{\kappa\,(\,\kappa \;+\; 1\,)}{r^2} \right) f(r) \;=\; -\left((\,E \;+\; V_0\,)^2 \;-\; m^2\,c^4 \right) f(r) \tag{1847}$$

By using a similar procedure, starting from the second equation, one can find the analogous equation for $g(r)$

$$c^2 \, \hbar^2 \left(\frac{\partial^2}{\partial r^2} - \frac{\kappa \, (\, \kappa \, - \, 1 \,)}{r^2} \right) g(r) = - \left((\, E \, + \, V_0 \,)^2 - m^2 \, c^4 \right) g(r)$$

(1848)

It should be recognized that the term proportional to $\kappa \, (\, \kappa \, + \, 1 \,)$ on the left-hand side of the eqn(1847) for the large component, when divided by $2 \, m \, c^2$, is equivalent to the centrifugal potential in the non-relativistic limit. The small component experiences a different centrifugal potential. Furthermore, the quantity

$$(\, E \, + \, V_0 \,)^2 - m^2 \, c^4$$

(1849)

plays a similar role to the kinetic energy in the non-relativistic Schrödinger equation.

Real Momenta

If the quantity $(\, E \, + \, V_0 \,)^2 - m^2 \, c^4$ is positive, it can be written as

$$(\, E \, + \, V_0 \,)^2 - m^2 \, c^4 = c^2 \, \hbar^2 \, k_0^2 > 0$$

(1850)

where k_0 is real. These equations can be expressed in dimensionless form by introducing the dimensionless variable variable $\rho = k_0 \, r$. The radial equations simplify to become

$$\rho^2 \, \frac{\partial^2 f}{\partial \rho^2} + \left(\rho^2 - \kappa \, (\, \kappa \, + \, 1 \,) \right) f = 0$$

$$\rho^2 \, \frac{\partial^2 g}{\partial \rho^2} + \left(\rho^2 - \kappa \, (\, \kappa \, - \, 1 \,) \right) g = 0$$

(1851)

Since (apart from the sign) κ is identified with a form of angular momentum, one sees that the upper and lower components experience different centrifugal potentials. These equations have forms which are closely related to Bessel's equation. If one sets

$$f = \rho^{\frac{1}{2}} \, X_{|\kappa + \frac{1}{2}|}$$

(1852)

and

$$g = \rho^{\frac{1}{2}} \, Y_{|\kappa - \frac{1}{2}|}$$

(1853)

the equations reduce to the pair of Bessel's equations

$$\rho^2 \, \frac{\partial^2 X_{|\kappa + \frac{1}{2}|}}{\partial \rho^2} + \rho \, \frac{\partial X_{|\kappa + \frac{1}{2}|}}{\partial \rho} + \left(\rho^2 - (\, \kappa \, + \, \frac{1}{2} \,)^2 \right) X_{|\kappa + \frac{1}{2}|} = 0$$

$$\rho^2 \, \frac{\partial^2 Y_{|\kappa - \frac{1}{2}|}}{\partial \rho^2} + \rho \, \frac{\partial Y_{|\kappa - \frac{1}{2}|}}{\partial \rho} + \left(\rho^2 - (\, \kappa \, - \, \frac{1}{2} \,)^2 \right) Y_{|\kappa - \frac{1}{2}|} = 0$$

(1854)

of half-integer order. The spherical Bessel functions and spherical Neumann functions of order n are defined in terms of the Bessel functions via

$$
\begin{aligned}
j_n(\rho) &= \sqrt{\frac{\pi}{2\,\rho}}\; J_{n+\frac{1}{2}}(\rho) \\
\eta_n(\rho) &= \sqrt{\frac{\pi}{2\,\rho}}\; N_{n+\frac{1}{2}}(\rho)
\end{aligned}
\tag{1855}
$$

Therefore, the general solutions of each of the radial equations can be expressed as

$$
\frac{f(r)}{r} \;=\; A_0\, j_{|\kappa+\frac{1}{2}|-\frac{1}{2}}(k_0 r) \;+\; A_1\, \eta_{|\kappa+\frac{1}{2}|-\frac{1}{2}}(k_0 r)
\tag{1856}
$$

and

$$
\frac{g(r)}{r} \;=\; B_0\, j_{|\kappa-\frac{1}{2}|-\frac{1}{2}}(k_0 r) \;+\; B_1\, \eta_{|\kappa-\frac{1}{2}|-\frac{1}{2}}(k_0 r)
\tag{1857}
$$

However, since the functions $f(r)$ and $g(r)$ in the upper and lower components are related by the differential equations

$$
\left(\frac{\partial}{\partial \rho} + \frac{1+\kappa}{\rho} \right) \left(\frac{f(r)}{r} \right) = \frac{(E + V_0 + m\,c^2)}{c\,\hbar\,k_0} \left(\frac{g(r)}{r} \right)
\tag{1858}
$$

and

$$
\left(\frac{\partial}{\partial \rho} + \frac{1-\kappa}{\rho} \right) \left(\frac{g(r)}{r} \right) = - \frac{(E + V_0 - m\,c^2)}{c\,\hbar\,k_0} \left(\frac{f(r)}{r} \right)
\tag{1859}
$$

the two sets of coefficients (A_0, A_1) and (B_0, B_1) must also be related. The explicit relations can be found by using the recurrence relations for the spherical Bessel functions $j_n(\rho)$

$$
\frac{\partial}{\partial \rho} \left(\rho^{n+1}\, j_n(\rho) \right) = \rho^{n+1}\, j_{n-1}(\rho)
\tag{1860}
$$

and

$$
\frac{\partial}{\partial \rho} \left(\rho^{-n}\, j_n(\rho) \right) = - \rho^{-n}\, j_{n+1}(\rho)
\tag{1861}
$$

The spherical Neumann functions $\eta_n(\rho)$ satisfy identical recurrence relations. This yields the relations

$$
\begin{aligned}
A_0 &= \mathrm{sign}\!\left(\kappa\right) \left(\frac{E + V_0 + m\,c^2}{c\,\hbar\,k_0} \right) B_0 \\
A_1 &= \mathrm{sign}\!\left(\kappa\right) \left(\frac{E + V_0 + m\,c^2}{c\,\hbar\,k_0} \right) B_1
\end{aligned}
\tag{1862}
$$

Hence, for positive-energy solutions. the upper components are the large components and the lower components are the small components. In the inner region, one must set $A_1 = B_1 = 0$, since the wave function are required to be

322

normalizable near the origin and the spherical Neumann functions $\eta_n(\rho)$ diverge as $\rho^{-(n+1)}$ as $\rho \to 0$.

Imaginary Momenta

If the quantity $(E + V_0)^2 - m^2 c^4$ is negative, it can be written as

$$(E + V_0)^2 - m^2 c^4 = - c^2 \hbar^2 \kappa_0^2 < 0 \tag{1863}$$

where κ_0 is real. This corresponds to the case of negative kinetic energies. In this case, one can express the solution in terms of the modified spherical Bessel functions

$$\frac{f(r)}{r} = A_0' \, i_{|\kappa+\frac{1}{2}|}(\kappa_0 r) + A_1' \, k_{|\kappa+\frac{1}{2}|}(\kappa_0 r) \tag{1864}$$

and

$$\frac{g(r)}{r} = B_0' \, i_{|\kappa-\frac{1}{2}|}(\kappa_0 r) + B_1' \, k_{|\kappa-\frac{1}{2}|}(\kappa_0 r) \tag{1865}$$

Because to the factors of i in the definitions of the modified spherical Bessel functions, the amplitudes of the upper and lower components are related via

$$A_0' = - \left(\frac{E + V_0 + m c^2}{c \, \hbar \, \kappa_0} \right) B_0'$$

$$A_1' = \left(\frac{E + V_0 + m c^2}{c \, \hbar \, \kappa_0} \right) B_1' \tag{1866}$$

where a minus sign has appeared in the first equation. Again, we see that for positive energies, for $r < a$, the upper components are the larger components and the lower components are the smaller components.

Bound States

The bound state energy E must occur in the energy interval

$$m c^2 > E > - m c^2 \tag{1867}$$

so that the wave function in the region $r < a$ where the potential is zero is exponentially decaying. Since $E^2 - m^2 c^4 < 0$, the wave functions in the outer region should also be expressed in terms of the modified spherical Bessel functions. The quantity κ_1 can be defined as

$$E^2 - m^2 c^4 = - \hbar^2 c^2 \kappa_1^2 \tag{1868}$$

and the equations can be expressed in terms of the dimensionless variable

$$\rho = i \kappa_1 r \tag{1869}$$

In this case, it is more useful to express the solution of the radial Dirac equation in terms of the spherical Hankel functions $h_n^\pm(\rho)$. The spherical Hankel functions are defined via

$$h_n^\pm(\rho) = j_n(\rho) \pm i \, \eta_n(\rho) \tag{1870}$$

For asymptotically large ρ, these functions are complex conjugates and represent out-going or incoming spherical waves

$$\lim_{\rho \to \infty} h_n^{\pm}(\rho) \;\to\; \frac{1}{\rho}\;\exp\left[\;\pm\;i\left(\;\rho\;-\;(\,n\;+\;\frac{1}{2}\,)\;\frac{\pi}{2}\;\right)\right] \qquad (1871)$$

The factor of ρ^{-1} reflects the fact that the intensity of an outgoing wave-packet decreases in proportion to ρ^{-2} in order to conserve energy and probability. From the asymptotic variation, it is seen that the spherical Hankel functions $h_n^{\pm}(i\rho)$ with imaginary arguments, respectively, represent exponentially attenuating or growing spherical waves. In the exterior region, the solutions are represented by

$$\frac{f(r)}{r} \;=\; C_0'\;h_{|\kappa+\frac{1}{2}|-\frac{1}{2}}^{+}(i\kappa_1 r) \;+\; C_1'\;h_{|\kappa+\frac{1}{2}|-\frac{1}{2}}^{-}(i\kappa_1 r) \qquad (1872)$$

and

$$\frac{g(r)}{r} \;=\; D_0'\;h_{|\kappa-\frac{1}{2}|-\frac{1}{2}}^{+}(i\kappa_1 r) \;+\; D_1'\;h_{|\kappa-\frac{1}{2}|-\frac{1}{2}}^{-}(i\kappa_1 r) \qquad (1873)$$

The coefficients of the upper and lower components are related via

$$C_0' \;=\; -\left(\frac{E\;+\;m\;c^2}{c\;\hbar\;\kappa_1}\right)D_0'$$

$$C_1' \;=\; \left(\frac{E\;+\;m\;c^2}{c\;\hbar\;\kappa_1}\right)D_1' \qquad (1874)$$

as can be seen by substituting the asymptotic form of the Hankel functions given by eqn(1871) in the asymptotic form of the differential equations relating $f(r)$ and $g(r)$ with $V_0 \;=\; 0$. If this wave function is to be normalizable at $\rho \to \infty$, one must set $C_1' \;=\; D_1' \;=\; 0$.

The solutions for the wave functions have been found in the inner and outer regions of the potential. The solution must also hold at $r = a$. This is achieved by demanding that the upper and lower components of the wave function are continuous at $r = a$. These conditions are demanded due to charge conservation $\partial_\mu\,j^\mu \;=\; 0$, since the current j^μ only depends on the components of ψ and does not (explicitly) depend on their derivatives.

Since the wave function at the origin must be normalizable, and since the wave function must be exponentially decaying, when $r \;\to\; \infty$, the matching condition for the upper component becomes

$$A_0\;j_{|\kappa+\frac{1}{2}|-\frac{1}{2}}(k_0 a) \;=\; C_0'\;h_{|\kappa+\frac{1}{2}|-\frac{1}{2}}^{+}(i\kappa_1 a) \qquad (1875)$$

and the matching condition for the lower components becomes

$$B_0\;j_{|\kappa-\frac{1}{2}|-\frac{1}{2}}(k_0 a) \;=\; D_0'\;h_{|\kappa-\frac{1}{2}|-\frac{1}{2}}^{+}(i\kappa_1 a) \qquad (1876)$$

324

By eliminating the amplitudes from the two matching conditions by using eqn(1874), one can arrive at the equation

$$\text{sign}(\kappa) \left(\frac{E \ + \ V_0 \ + \ m \ c^2}{c \ \hbar \ k_0} \right) \left(\frac{j_{|\kappa + \frac{1}{2}| - \frac{1}{2}}(k_0 a)}{j_{|\kappa - \frac{1}{2}| - \frac{1}{2}}(k_0 a)} \right) = - \left(\frac{E \ + \ m \ c^2}{c \ \hbar \ \kappa_1} \right) \left(\frac{h^+_{|\kappa + \frac{1}{2}| - \frac{1}{2}}(i\kappa_1 a)}{h^+_{|\kappa - \frac{1}{2}| - \frac{1}{2}}(i\kappa_1 a)} \right)$$

$$(1877)$$

In the above expression, the quantities k_0 and κ_1 are defined by

$$\hbar^2 \ c^2 \ k_0^2 \ = \ (\ E \ + \ V_0 \ )^2 \ - \ m^2 \ c^4 \tag{1878}$$

and

$$\hbar^2 \ c^2 \ \kappa_1^2 \ = \ m^2 \ c^4 \ - \ E^2 \tag{1879}$$

These equations determine the allowed values for the energy. The above set of equations have to be solved numerically to find the energy eigenvalues. We note that for the Dirac particle, the spin effectively results in the formation of a centrifugal barrier (either for the upper or the lower component) even for electrons in s states. As a result, the potential V_0 must exceed a critical strength if it is to yield a bound state.

12.10.9   The MIT Bag Model

From the point of view of symmetry, a baryon, such as a neutron or proton, are thought of as being composed of three (valence) quarks. For example, the proton is considered to be made of two up quarks and a down quark ($p = (uud)$), while the neutron is considered to be made of one up quark and two down quarks ($n = (udd)$). These valence quarks are assumed to be surrounded by a sea of gluons which bind the quarks together and a sea of virtual quark/anti-quark pairs that are produced by the gluon field. Likewise, mesons are considered to be made of a quark and an anti-quark, but these valence quarks are also surrounded by a sea of gluons and quark/anti-quark pairs. The gluon force has the property that the energy of interaction increases as the separation between the quarks increases. It is this property of the gluon force that results in the quarks being confined, so that no single quark can be found in nature.

The MIT bag model[115] is a simple purely phenomenological model for the structure of strongly interacting particles (hadrons). The model is based on the spherically symmetric potential of radius a, but it will be assumed that the quark mass can have one or the other of two values. The quark is assumed to have a small mass (approximately zero) if it is located within a sphere of radius a, and the mass is assumed to be very large (or infinite) if $r > a$. To be sure, the quark mass is assumed to be a function of r such that

$$\begin{aligned} m \ &= \ 0 \quad \text{if} \ r \ < \ a \\ m \ &\rightarrow \ \infty \quad \text{if} \ r \ > \ a \end{aligned} \tag{1880}$$

[115]A. Chodos, R. L. Jaffe, K. Johnson, C. B. Thorn, and V. F. Weisskopf, Phys. Rev. D **9**, 3471, (1974).

It is the infinite mass of the quark for $r > a$ that results in the confinement of the quark to within the hadron. That is, in the exterior region, the infinite rest mass energy exceeds the bound state energy so the exterior region is classically forbidden, therefore, the particle is confined to the interior.

Inside the hadron, where both the potential energy and the mass m are zero, the kinetic energy parameter k_0 can be expressed entirely in terms of the energy via $E = \hbar c k_0$ since the potential is assumed to be zero. Therefore, the radial components of the Dirac wave function can be expressed as

$$
\begin{aligned}
\frac{f(r)}{r} &= A_0 \, j_{|\kappa+\frac{1}{2}|-\frac{1}{2}}(k_0 r) \\
\frac{g(r)}{r} &= \text{sign}(\kappa) \, A_0 \, j_{|\kappa-\frac{1}{2}|-\frac{1}{2}}(k_0 r)
\end{aligned}
\tag{1881}
$$

where the amplitudes of the upper and lower components are the same, since the potential and mass are zero for $r < a$.

Outside the hadron, where $r > a$, the energy E is assumed to be much less than the rest mass energy, $m\,c^2 \gg E$, therefore, the momentum parameter is imaginary and one can set $\hbar\,c\,\kappa_1 \approx m\,c^2$. In the exterior region, the radial functions can be expressed as

$$
\begin{aligned}
\frac{f(r)}{r} &= C_0' \, h^+_{|\kappa+\frac{1}{2}|-\frac{1}{2}}(i\kappa_1 r) \\
\frac{g(r)}{r} &= -\,C_0' \, h^+_{|\kappa-\frac{1}{2}|-\frac{1}{2}}(i\kappa_1 r)
\end{aligned}
\tag{1882}
$$

since the imaginary momentum parameter has a magnitude which is governed by the large mass m. Due to the large magnitude of κ_1, the wave function decays very rapidly in the exterior region.

The bound state energy is determined from the matching condition

$$
\text{sign}(\kappa)\left(\frac{j_{|\kappa+\frac{1}{2}|-\frac{1}{2}}(k_0 a)}{j_{|\kappa-\frac{1}{2}|-\frac{1}{2}}(k_0 a)}\right) = -\left(\frac{h^+_{|\kappa+\frac{1}{2}|-\frac{1}{2}}(i\kappa_1 a)}{h^+_{|\kappa-\frac{1}{2}|-\frac{1}{2}}(i\kappa_1 a)}\right)
\tag{1883}
$$

Due to the asymptotic properties of the spherical Hankel functions, their ratio is unity for large κ_1. This leads to the energies of the quarks being governed by the simplified matching condition

$$
j_{|\kappa+\frac{1}{2}|-\frac{1}{2}}(k_0 a) = -\,\text{sign}(\kappa)\, j_{|\kappa-\frac{1}{2}|-\frac{1}{2}}(k_0 a)
\tag{1884}
$$

where

$$
E = c\,\hbar\,k_0
\tag{1885}
$$

The above equation governs the ground state and excited state energies of the individual quarks inside the hadron. Since the spherical Bessel functions oscillate in sign, the above equations will result in a set of solutions for k_0 with fixed

326

κ. From the structure of the equations, it is seen that the solutions k_0 will only depend on the integer number κ and the value of a. Since another boundary condition should also be imposed at the bag's surface, only states with angular momentum $j = \frac{1}{2}$ should be retained. This extra condition restricts the interest to states with $\kappa = -1$.

We shall examine the lowest-energy bound state which corresponds to the case $\kappa = -1$. The bound state energies are given by the matching condition

$$j_0(k_0 a) = j_1(k_0 a) \tag{1886}$$

Since

$$j_0(\rho) = \frac{\sin \rho}{\rho} \tag{1887}$$

and

$$j_1(\rho) = \left(\frac{\sin \rho - \rho \cos \rho}{\rho^2} \right) \tag{1888}$$

the energy eigenvalues are determined by the solutions of

$$\rho = \frac{1}{1 + \cot \rho} \tag{1889}$$

which has an infinite number of solutions which, asymptotically, are spaced by π. The smallest solution corresponds to $k_0 \, a = 2.04$. Hence, the energy of the
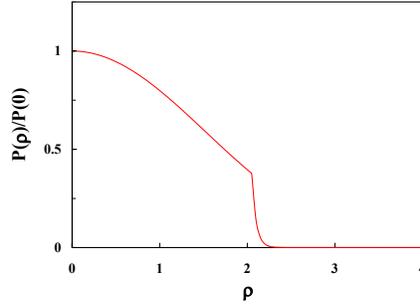
Figure 59: The radial dependence of the quark-distribution in the ground state of the MIT bag.

lowest-energy quark is give by the formula

$$E_{\kappa=-1,n_r=0} = \frac{2.04 \, c \, \hbar}{a} \tag{1890}$$

The solutions with larger values of k_0, corresponding to excited states with $\kappa = -1$ are given by analogous expressions. Therefore, if one knows the value of a, one could find the energies required to excite a single-quark between the

Table 15: The lowest single-particle energies (in units of $E_{\kappa, n_r} a/c\,\hbar$) of the MIT Bag Model.

n_r	$\kappa = -1$	$\kappa = +1$	$\kappa = -2$	$\kappa = +2$
$n_r = 0$	2.04	3.81	3.21	5.12
$n_r = 1$	5.40	7.00	6.76	8.41
$n_r = 2$	8.58	10.17	10.01	11.61
$n_r = 3$	11.73	13.31	13.20	14.79

single particle levels. This could allow one to calculate the excitation energies required to change the hadron's internal structure.

In conclusion, the MIT bag model, when interpreted as being a strictly single-particle picture, predicts that the set of excitation energies (for the internal structure) of each of the basic hadrons can be put into a one-to-one correspondence with each other. That is, the family of excitation energies for each hadron should fall on-top of each other, if one scales the energies by multiplying them with the hadron's characteristic length scale a. The bag radius is determined by the use of further phenomenological considerations. However, although the model can be used to fit the right size for a nucleon (~ 1 fm), the model predicts that a meson (such as the pions which are composed of a quark and anti-quark in the combinations of either $(u, \bar{d})$, $(d, \bar{u})$ or $\frac{1}{\sqrt{2}}(d\bar{d} - u\bar{u})$) should have almost the same radii[116]

$$\frac{a_n}{a_\pi} = \left(\frac{3}{2}\right)^{\frac{1}{4}} \tag{1891}$$

Hence, the ratio of the nucleon mass M_n to the pion mass M_π is expected to be given by the formula

$$\frac{M_n}{M_\pi} = \frac{3 \times 2.04/a_n}{2 \times 2.04/a_\pi} = \frac{3}{2}\left(\frac{2}{3}\right)^{\frac{1}{4}} \tag{1892}$$

---

[116]It is assumed that the bag energy is given by the sum of a volume term $B\,a^3$ and the sum of the quark energies $\frac{c\,\hbar}{a}\sum_n \alpha_n$. Minimizing the energy w.r.t a results in the bag radius a being determined by

$$a^4 = \frac{c\,\hbar}{3\,B}\sum_n \alpha_n$$

Table 16: The Observed Energy Levels for the charmonium system ($c\bar{c}$) in units of MeV/c^2.

$^1S_0$	$^3S_1$	$^3P_0$	$^3P_1$	$^3P_2$
2981	3097	3415	3510	3556
3686	3770	-	-	-
-	4040	-	-	-
-	4160	-	-	-

which yields a ratio of 1.36. This ratio is far too small for the triplet of π mesons since $M_\pi \sim 139$ MeV/c^2, and $M_n \sim 938$ MeV/c^2. Although it is in adequate for the pseudo-scalar mesons, the MIT bag model is more appropriate for the ω vector meson which is composed of $\frac{1}{\sqrt{2}}(u\bar{u} + d\bar{d})$ and has a mass of $M_\omega \sim 783$ MeV/c^2, or the ρ vector meson $\frac{1}{\sqrt{2}}(u\bar{u} + d\bar{d})$ with a mass $M_\rho \sim 776$ MeV/c^2. Hence, at best, the MIT Bag model produces mixed results. The MIT Bag model is also quite unappealing, since the basic assumptions of the bag model do not follow from Quantum Chromodynamics, and the model is neither re-normalizable nor is it Lorentz invariant.

12.10.10 The Temple Meson Model

A quark and anti-quark pair form bound states. Thus, for example a charmed quark/anti-quark pair $(c,\bar{c})$ can form states with different internal quantum numbers[117]. The experimentally determined energies for the J/Ψ system[118] are given in Table(16). Similarly, the Upsilon particle[119] ($b\bar{b}$) has a similar set of energy levels. The energy levels of the Upsilon system[120] are tabulated in Table(17). For positronium[121], like the hydrogen atom, it is the electromagnetic force mediated by vector photons which binds the electron and positron into a bound state. For a quark/anti-quark bound state, it is the color force mediated by massless vector gluons that bind the quark/anti-quark pair together. The color force has the property that it increases with increasing separation of the quark/anti-quark pair, which has the consequence that the quarks are confined. Furthermore, high-energy inelastic scattering experiments on hadrons indicate

---

[117]J. E. Augustin *et al.* Phys. Rev. Lett. **33**, 1406 (1974).
J. J. Aubert Phys. Rev. Lett. **33**, 1404 (1974).
[118]The data are taken from the Particle Data Group: http://pdg.lbl.gov
[119]S. W. Herb, *et al.* Phys. Rev. Lett. **39**, 252 (1977).
W. R. Innes *et al.* Phys. Rev. Lett. **39**, 1240 (1977).
[120]The data are taken from the Particle Data Group: http://pdg.lbl.gov
[121]M. Deutsch, Phys. Rev. **82**, 455 (1951).

Table 17: The Observed Energy Levels for the Upsilon system ($b\bar{b}$) in units of MeV/c^2.

$^1S_0$	$^3P_0$	$^3P_1$	$^3P_2$
9460	9860	9893	9913
10025	10232	10255	10268
10355	-	-	-
10580	-	-	-

that at small separations the quarks only interact weakly. This property is called asymptotic freedom. It was the realization by 't Hooft[122], Gross and Wilczek[123] and Politzer[124] that non-Abelian gauge theories possessed the properties of asymptotic freedom that led to the acceptance of the theory of Quantum Chromodynamics. The screening of the color force between the quarks at large distances (due to virtual quark/anti-quark pairs) is more than compensated by an anti-screening due to virtual gluon pairs. However, at small distances the color force vanishes.

The rest-mass energy of the quarks and anti-quarks will be modeled by

$$m(r) \, c \; = \; m_0 \left(\, c \, - \, i \, \omega \, \underline{\alpha} \, . \, \underline{r} \, \right) \tag{1893}$$

which describes an energy similar to that of an elastic string which couples to the spin[125]. The model has two undetermined parameters, the quark mass m_0 and the string tension $m_0 \, c \, \omega$. The mass $m(r)$, and the Dirac equation, can be used to determine the energy levels of quarkonium.

Exercise:

Show that the positive energy eigenvalues of the Dirac equation with the mass $m(r)$ given by

$$m(r) \, c \; = \; m_0 \left(\, c \, - \, i \, \omega \, \underline{\alpha} \, . \, \underline{r} \, \right) \tag{1894}$$

are determined as

$$E_{n,j,l} \; = \; m_0 \, c^2 \sqrt{ \, t \, A \, + \, 1} \tag{1895}$$

[122]G. t' Hooft, unpublished (1972).

[123]D. J. Gross and F. A. Wilczek, Phys. Rev. Lett. **30**, 1343 (1973).

[124]H. D. Politzer, Phys. Rev. Lett. **30**, 1346 (1973).

[125]D. Ito, K. Mori and E. Carriere, Nuovo Cimento, **51** A, 1119, (1967).
P. A. Cook, Nuovo Cimento Lett. **1**, 419 (1971).

where the dimensionless parameter t corresponding to the string tension is given by

$$t = \left(\frac{\hbar \, \omega}{m_0 \, c^2} \right) \tag{1896}$$

and A is given in terms of the quantum numbers as

$$
\begin{aligned}
A &= 2 \, (n - j) + 1 \; \text{ if } \; j = l + \frac{1}{2} \\
A &= 2 \, (n + j) + 3 \; \text{ if } \; j = l - \frac{1}{2}
\end{aligned}
\tag{1897}
$$

Hence, find the best fit to the excitation spectra of quarkonium.

12.11   Scattering by a Spherically Symmetric Potential

First, the polarization dependence of scattering of an electron from a Coulomb potential will be examined in terms of the scattering amplitudes, and second, by using a partial wave analysis, the scattering amplitudes will be expressed in terms of phase shifts.

12.11.1   Polarization in Coulomb Scattering.

The scattering of a relativistic electron by a Coulomb force field results in spin-flip scattering since the electron has a magnetic moment which interacts with the magnetic field produced in the electron's rest frame. Since the Coulomb potential is spherically symmetric, the angular momentum $\hat{J}^2$ and $\hat{J}^{(3)}$ commute with the Hamiltonian, hence, (j, j_3) are constants of motion. However, the orbital angular momentum $\underline{\hat{L}}$ does not commute with $\hat{H}$.

The Dirac wave function $\psi(\underline{r})$ can be expressed in terms of two two-component spinors

$$\psi(\underline{r}) = \left(\begin{array}{c} \phi^A(\underline{r}) \\ \phi^B(\underline{r}) \end{array} \right) \tag{1898}$$

One only need specify the upper component $\phi^A(\underline{r})$, since once $\phi^A(\underline{r})$ has been specified $\phi^B(\underline{r})$ is completely determined. For example, for the in and out asymptotes, the Dirac equation reduces to

$$\left(\begin{array}{cc} E_p \, - \, m \, c^2 & - \, c \, \hat{\underline{p}} \cdot \underline{\sigma} \\ - \, c \, \hat{\underline{p}} \cdot \underline{\sigma} & E_p \, + \, m \, c^2 \end{array} \right) \left(\begin{array}{c} \phi^A(\underline{r}) \\ \phi^B(\underline{r}) \end{array} \right) = 0 \tag{1899}$$

Hence, the lower two-component spinor is completely determined in terms of the upper two-component spinor

$$\phi^B(\underline{r}) = \frac{c \, \hat{\underline{p}} \cdot \underline{\sigma}}{E_p \, + \, m \, c^2} \, \phi^A(\underline{r}) \tag{1900}$$

In the scattering experiment, a plane-wave with momentum $\underline{p}$ parallel to the $\hat{e}_3$-axis falls incident on the target. The in-asymptote can be described by a state which is in a superposition of eigenstates of $\hat{S}^{(3)}$ given by

$$\psi_\pm^{in}(\underline{r}) \; = \; \mathcal{N}_{E_p} \; \left(\; \pm \; \begin{matrix} \chi_\pm \\ \frac{c\ p}{E_p\ +\ m\ c^2}\ \chi_\pm \end{matrix} \; \right) \; \exp\left[\; i \; \frac{p\ r}{\hbar} \; \cos\theta \; \right] \tag{1901}$$

From the Rayleigh expansion, one observes that the in-asymptotes are not eigenstates of $(\hat{\underline{J}})^2 = (\hat{\underline{L}} + \hat{\underline{S}})^2$ since they are formed of linear superpositions of many states with differing eigenvalues of $\hat{\underline{L}}^2$ but have a fixed eigenvalue of $\hat{\underline{S}}^2$. However, the in-asymptote are eigenstates of $\hat{J}^{(3)} \; = \; \hat{L}^{(3)} \; + \; \hat{S}^{(3)}$ with eigenvalues $\pm \frac{\hbar}{2}$.
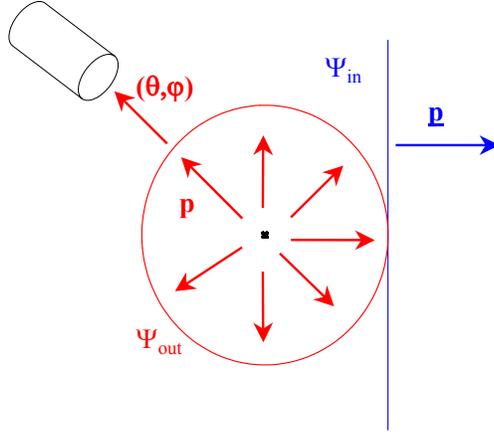
Figure 60: The geometry of the asymptotic final state of Mott scattering. At large r, the beam separated into an unscattered beam ψ_{in} and a spherical outgoing wave ψ_{out}.

The corresponding out-asymptotes can be described as spherical outgoing waves. Even though the in-asymptote may have a definite eigenvalue of $\hat{S}^{(3)}$, the spherically symmetric out-asymptote waves may contain a component with flipped spin, due to the action of the spin-orbit coupling

$$\left(\; \hat{\underline{S}} \cdot \hat{\underline{L}} \; \right) \; = \; \hat{S}^{(3)} \; \hat{L}^{(3)} \; + \; \frac{1}{2} \left(\; \hat{S}^+ \; \hat{L}^- \; + \; \hat{S}^- \; \hat{L}^+ \; \right) \tag{1902}$$

active in the vicinity of the target. In spherical polar coordinates, the orbital

angular momentum raising and lowering operators are given by

$$\hat{L}^{\pm} \;=\; \pm\,\hbar\,\exp[\pm i\varphi]\left(\frac{\partial}{\partial\theta}\;\pm\;i\,\cot\theta\,\frac{\partial}{\partial\varphi}\right) \tag{1903}$$

Hence, on noting that

$$\hat{S}^{\pm}\,\chi_{\pm}\;\equiv\;0 \tag{1904}$$

one finds that the out-asymptotes can be expressed as

$$\psi_{\pm}^{out}(\underline{r}) \;=\; \mathcal{N}_{E_p}\left(\begin{array}{cc}(\,f(\theta)\,\pm\,g(\theta)\,\exp[\,\pm\,i\varphi\,]\,\hat{S}^{\mp}\,) & \chi_{\pm} \\ \frac{c\,p\,(\,\hat{e}_r\,\cdot\,\underline{\sigma}\,)}{E_p\,+\,m\,c^2}\,(\,f(\theta)\,\pm\,g(\theta)\,\exp[\,\pm\,i\varphi\,]\,\hat{S}^{\mp}\,) & \chi_{\pm}\end{array}\right)\frac{1}{r}\,\exp\left[\,i\,\frac{p\,r}{\hbar}\,\right] \tag{1905}$$

where $\hat{e}_r$ is a unit vector in the radial direction. It should be noted that the out-asymptote describes an outgoing spherical wave when $r\to\infty$. Therefore, the operator $(\,\underline{\sigma}\,\cdot\,\hat{\underline{p}}\,)$ appearing in the asymptote has simplified since

$$\begin{aligned}\lim_{r\to\infty}(\,\underline{\sigma}\,\cdot\,\hat{\underline{p}}\,) \;&=\; \lim_{r\to\infty}\left(\frac{\underline{r}\,\cdot\,\underline{\sigma}}{r}\right)\left[\,-\,i\,\hbar\,\frac{\partial}{\partial r}\,+\,\frac{2\,i}{\hbar}\left(\frac{\hat{\underline{S}}\,\cdot\,\hat{\underline{L}}}{r}\right)\right] \\ &\to\; \left(\frac{\underline{r}\,\cdot\,\underline{\sigma}}{r}\right)\left(\,-\,i\,\hbar\,\frac{\partial}{\partial r}\,\right)\end{aligned} \tag{1906}$$

which reflects that the spin-orbit coupling term is ineffective at $r\to\infty$. Similarly, the effect of the differential operator can be evaluated as

$$\lim_{r\to\infty}-\,i\,\hbar\,\frac{\partial}{\partial r}\left(\frac{1}{r}\,\exp\left[\,i\,\frac{p\,r}{\hbar}\,\right]\right)\;\to\;\frac{p}{r}\,\exp\left[\,i\,\frac{p\,r}{\hbar}\,\right] \tag{1907}$$

In light of the comment about the upper two-component spinor, one sees that the scattered wave is determined by

$$\left(\begin{array}{c}f(\theta) \\ g(\theta)\,\exp[+i\varphi]\end{array}\right) \tag{1908}$$

for an incident beam with positive helicity, and by

$$\left(\begin{array}{c}-g(\theta)\,\exp[-i\varphi] \\ f(\theta)\end{array}\right) \tag{1909}$$

if the initial beam has a negative helicity. The quantities $f(\theta)$ and $g(\theta)$ are generalized scattering amplitudes that have the dimensions of length, and depend on θ but do not depend on φ as both the in and out asymptotes are eigenstates of $\hat{J}^{(3)}$ with eigenvalues $\pm\frac{\hbar}{2}$. A partial wave analysis can be performed on the Dirac equation to yield expressions for the scattering amplitudes $f(\theta)$ and $g(\theta)$ in terms of phase shifts. A detailed knowledge of the scattering amplitudes is not required for the following analysis.

If the in-asymptote has the spin quantized along the direction given by $(\sin \theta_s \cos \varphi_s, \sin \theta_s \sin \varphi_s, \cos \theta_s)$, the upper component of the Dirac wave spinor is determined by the two-component spinor

$$\chi_s = \begin{pmatrix} \cos \frac{\theta_s}{2} \exp[-i\frac{\varphi_s}{2}] \\ \sin \frac{\theta_s}{2} \exp[+i\frac{\varphi_s}{2}] \end{pmatrix} \tag{1910}$$

The out-asymptote is then determined by the two-component spinor

$$\phi^A(\underline{r}) = \begin{pmatrix} f(\theta) \cos \frac{\theta_s}{2} \exp[-i\frac{\varphi_s}{2}] - g(\theta) \sin \frac{\theta_s}{2} \exp[+i\frac{\varphi_s}{2}] \exp[-i\varphi] \\ g(\theta) \cos \frac{\theta_s}{2} \exp[-i\frac{\varphi_s}{2}] \exp[+i\varphi] + f(\theta) \sin \frac{\theta_s}{2} \exp[+i\frac{\varphi_s}{2}] \end{pmatrix} \frac{1}{r} \exp\left[i\frac{p\,r}{\hbar} \right] \tag{1911}$$

The probability for scattering is proportional to

$$I(\theta, \varphi) \propto \left| f(\theta) \cos \frac{\theta_s}{2} \exp[-i\frac{\varphi_s}{2}] - g(\theta) \sin \frac{\theta_s}{2} \exp[+i\frac{\varphi_s}{2}] \exp[-i\varphi] \right|^2$$

$$+ \left| g(\theta) \cos \frac{\theta_s}{2} \exp[-i\frac{\varphi_s}{2}] \exp[+i\varphi] + f(\theta) \sin \frac{\theta_s}{2} \exp[+i\frac{\varphi_s}{2}] \right|^2$$

$$= \left(|f(\theta)|^2 + |g(\theta)|^2 \right) + \sin \theta_s \sin(\varphi - \varphi_s)\, i\, \left(f^*(\theta)\, g(\theta) - f(\theta)\, g^*(\theta) \right) \tag{1912}$$

which clearly depends on the azimuthal angle φ.

If the initial beam is unpolarized, the direction of the initial spin (θ_s, φ_s) should be averaged over by integrating over the solid angle $d\Omega_s = d\varphi_s\, d\theta_s\, \sin \theta_s$. This process yields the scattering probability for the unpolarized beam

$$\int \frac{d\Omega_s}{4\,\pi}\, I(\theta, \varphi) = \left(|f(\theta)|^2 + |g(\theta)|^2 \right) \tag{1913}$$

which is independent of the azimuthal angle φ. It should be noted that the unpolarized cross-section differs from the polarized cross-section.

Even if the initial beam is unpolarized, the final beam will be partially polarized. The direction of the net polarization is determined by evaluating the matrix elements of $\hat{\underline{S}}$ and averaging over the direction of the initial spin, θ_s and φ_s. The result is proportional to

$$\overline{\hat{\underline{S}}} = \frac{\hbar}{2}\, i\, \left[\frac{f^*(\theta)\, g(\theta) - f(\theta)\, g^*(\theta)}{|f(\theta)|^2 + |g(\theta)|^2} \right] (\sin \varphi, -\cos \varphi, 0) \tag{1914}$$

Hence, the polarization is perpendicular to the scattering plane. It should also be noted that the net polarization of the scattered wave is determined by the relative deviation of the scattering cross-section for polarized electrons from the unpolarized scattering cross-section.

12.11.2 Partial Wave Analysis

The Dirac equation with a spherically symmetric potential $V(r)$ has solutions of the form

$$\psi(\underline{r}) \;=\; \left(\begin{array}{c} \frac{f(r)}{r}\, \Omega_{j,j_z}^{j\pm\frac{1}{2}}(\theta,\varphi) \\ i\, \frac{g(r)}{r}\, \Omega_{j,j_z}^{j\mp\frac{1}{2}}(\theta,\varphi) \end{array} \right) \tag{1915}$$

where the two-component spinor spherical harmonics $\Omega_{j,j_z}^{j\pm\frac{1}{2}}(\theta,\varphi)$ are given by

$$\Omega_{j,j_z}^{j\pm\frac{1}{2}}(\theta,\varphi) \;=\; \left(\begin{array}{c} \mp\, \sqrt{\frac{j+\frac{1}{2}\pm\frac{1}{2}\mp j_z}{2j+1\pm 1}}\; Y_{j_z-\frac{1}{2}}^{j\pm\frac{1}{2}}(\theta,\varphi) \\ \sqrt{\frac{j+\frac{1}{2}\pm\frac{1}{2}\pm j_z}{2j+1\pm 1}}\; Y_{j_z+\frac{1}{2}}^{j\pm\frac{1}{2}}(\theta,\varphi) \end{array} \right) \tag{1916}$$

and the radial functions $f_\kappa(r)$ and $g_\kappa(r)$ satisfy

$$\left(E \;-\; V(r) \;-\; m\,c^2 \right) f_\kappa(r) \;=\; -\,c\,\hbar \left(\frac{\partial}{\partial r} \;-\; \frac{\kappa}{r} \right) g_\kappa(r)$$

$$\left(E \;-\; V(r) \;+\; m\,c^2 \right) g_\kappa(r) \;=\; c\,\hbar \left(\frac{\partial}{\partial r} \;+\; \frac{\kappa}{r} \right) f_\kappa(r) \tag{1917}$$

where $\kappa \;=\; \pm\,(\,j\,+\,\frac{1}{2}\,)$. If the momentum $\hbar\,k$ is defined via

$$c^2\,\hbar^2\,k^2 \;=\; E^2 \;-\; m^2\,c^4 \tag{1918}$$

the asymptotic $r \to \infty$ form of the solutions of these coupled equations with positive values of κ are of the form of a linear superposition

$$\frac{f_\kappa(r)}{r} \;=\; A_\kappa\, j_\kappa(kr) \;+\; B_\kappa\, \eta_\kappa(kr) \tag{1919}$$

where $j_\kappa(kr)$ and $\eta_\kappa(kr)$ are the spherical Bessel and the spherical Neumann functions. For negative values of κ, the solutions are given by

$$\frac{f_\kappa(r)}{r} \;=\; A_\kappa\, j_{-\kappa-1}(kr) \;+\; B_\kappa\, \eta_{-\kappa-1}(kr) \tag{1920}$$

The spherical Bessel and spherical Neumann functions have the asymptotic forms

$$j_\kappa(kr) \;\to\; \frac{\cos(kr - (\kappa+1)\frac{\pi}{2})}{kr}$$

$$\eta_\kappa(kr) \;\to\; \frac{\sin(kr - (\kappa+1)\frac{\pi}{2})}{kr} \tag{1921}$$

The solutions for a free particle do not involve the spherical Neumann functions, since they are not normalizable at the origin. The amplitudes of the asymptotic solution in the presence of a finite potential $V(r)$ are usually written as

$$\frac{B_\kappa}{A_\kappa} \;=\; -\,\tan\delta_\kappa(k) \tag{1922}$$

where $\delta_\kappa(k)$ are the phase shifts that characterize the potential. The phase shifts depend directly on κ (and the energy) and only depend indirectly on j and l through κ. The phase shifts are defined so that the asymptotic variation of the radial functions is given by

$$\frac{f_\kappa(r)}{r} \sim e^{i\delta_\kappa(k)} \frac{\cos(kr - (\kappa+1)\frac{\pi}{2} + \delta_\kappa(k))}{r} \tag{1923}$$

and only differs from the asymptotic variation of the free particle solutions through the phase shifts. Furthermore, if this is decomposed in terms of incoming and outgoing spherical waves,

$$\frac{f_\kappa(r)}{r} \sim \frac{\exp\left[i \left(k\,r - (\kappa+1)\frac{\pi}{2} + 2\,\delta_\kappa(k) \right) \right]}{2\,r}$$
$$+ \frac{\exp\left[-i \left(k\,r - (\kappa+1)\frac{\pi}{2} \right) \right]}{2\,r} \tag{1924}$$

their fluxes are equal due to conservation of particles and, as written, the incoming spherical waves are not modified by the phase-shifts.

The general asymptotic $r \to \infty$ form of the wave function for the scattering is composed of the un-scattered wave and a spherical outgoing wave. The polar-axis is chosen to be parallel to direction of the incident beam which is also chosen to be the quantization axis for the spin. If the incident beam is polarized with spin-up, the upper two-component spinor has the form

$$\phi_\uparrow^A(\underline{r}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \exp\left[i\,k\,r\cos\theta \right] + \begin{pmatrix} f(\theta) \\ g(\theta)\,\exp[\,i\,\varphi\,] \end{pmatrix} \frac{\exp\left[i\,k\,r \right]}{r} \tag{1925}$$

whereas for a down-spin polarized incident beam

$$\phi_\downarrow^A(\underline{r}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \exp\left[i\,k\,r\cos\theta \right] + \begin{pmatrix} -g(\theta)\,\exp[\,-i\,\varphi\,] \\ f(\theta) \end{pmatrix} \frac{\exp\left[i\,k\,r \right]}{r} \tag{1926}$$

On recalling the Rayleigh expansion

$$\exp\left[i\,k\,r\cos\theta \right] = \sum_l i^l\,(\,2\,l + 1\,)\,j_l(kr)\,P_l(\cos\theta) \tag{1927}$$

one can find the scattered spherical outgoing wave by subtracting the un-scattered beam from the total wave function. On using the asymptotic large r variation, one obtains the asymptotic form

$$\exp\left[i\,k\,r\cos\theta \right] \to \sum_l i^l\,(\,2\,l + 1\,)\,\frac{\cos(kr - (l+1)\frac{\pi}{2})}{kr}\,P_l(\cos\theta) \tag{1928}$$

which has a similar form to the asymptotic form of the total wave function. In particular, the spin and orbital angular momentum eigenstates can be decomposed in terms of the spinor spherical harmonics. Thus, for the up-spin polarized incident beam one has the upper two-component spinor

$$
\begin{aligned}
P_l(\cos\theta)\,\chi_+ \;&=\; \sqrt{\frac{4\,\pi}{2\,l\,+\,1}}\;Y_0^l(\theta,\varphi)\,\chi_+ \\
&=\; \frac{\sqrt{4\,\pi}}{2\,l\,+\,1}\left(\sqrt{l+1}\;\Omega^l_{l+\frac{1}{2},\frac{1}{2}}\;-\;\sqrt{l}\;\Omega^l_{l-\frac{1}{2},\frac{1}{2}}\right) \quad (1929)
\end{aligned}
$$

and for the down-spin beam

$$
\begin{aligned}
P_l(\cos\theta)\,\chi_- \;&=\; \sqrt{\frac{4\,\pi}{2\,l\,+\,1}}\;Y_0^l(\theta,\varphi)\,\chi_- \\
&=\; \frac{\sqrt{4\,\pi}}{2\,l\,+\,1}\left(\sqrt{l+1}\;\Omega^l_{l+\frac{1}{2},-\frac{1}{2}}\;+\;\sqrt{l}\;\Omega^l_{l-\frac{1}{2},-\frac{1}{2}}\right) \quad (1930)
\end{aligned}
$$

Therefore, when expressed in terms of a superposition of continuum energy eigenstates corresponding to different values of j and κ, the asymptotic form of the Rayleigh expansion becomes

$$
\exp\left[i\,k\,r\cos\theta\right]\chi_+ \;\to\; \sqrt{4\,\pi}\sum_l i^l\,\frac{\cos(kr-(l+1)\frac{\pi}{2})}{kr}\left(\sqrt{l+1}\;\Omega^l_{l+\frac{1}{2},\frac{1}{2}}\;-\;\sqrt{l}\;\Omega^l_{l-\frac{1}{2},\frac{1}{2}}\right)
$$

$$(1931)$$

and

$$
\exp\left[i\,k\,r\cos\theta\right]\chi_- \;\to\; \sqrt{4\,\pi}\sum_l i^l\,\frac{\cos(kr-(l+1)\frac{\pi}{2})}{kr}\left(\sqrt{l+1}\;\Omega^l_{l+\frac{1}{2},-\frac{1}{2}}\;+\;\sqrt{l}\;\Omega^l_{l-\frac{1}{2},-\frac{1}{2}}\right)
$$

$$(1932)$$

Although the coefficients A_κ of the exact wave function are as yet unknown, they can be determined by requiring the scattered spherical wave does not contain terms proportional to

$$
\frac{\exp\left[-\,i\,k\,r\right]}{r} \qquad (1933)
$$

which would represent an incoming spherical wave. This requirement leads to the outgoing spherical wave having a spin-up component given by

$$
\begin{aligned}
\frac{\sqrt{4\,\pi}}{2\,i\,k}\sum_l\Bigg[&\frac{(\,l\,+\,1\,)}{\sqrt{2\,l\,+\,1}}\left(\exp[\,2\,i\,\delta_{-l-1}(k)\,]\;-\;1\right) \\
&+\;\frac{l}{\sqrt{2\,l+1}}\left(\exp[\,2\,i\,\delta_l(k)\,]\;-\;1\right)\Bigg]\,Y_0^l(\theta,\varphi)\,\frac{\exp\left[i\,k\,r\right]}{r}
\end{aligned}
$$

$$(1934)$$

and the down-spin component is given by

$$
\frac{\sqrt{4\,\pi}}{2\,i\,k} \sum_l \sqrt{\frac{l\,(\,l\,+\,1\,)}{2\,l\,+\,1}} \left[\left(\exp[\,2\,i\,\delta_{-l-1}(k)\,]\ -\ 1 \right) \right.
$$
$$
\left. -\ \left(\exp[\,2\,i\,\delta_l(k)\,]\ -\ 1 \right) \right]\,Y_1^l(\theta,\varphi)\ \frac{\exp\left[\,i\,k\,r\,\right]}{r} \qquad (1935)
$$

In the above expressions, the index on the phase-shifts $\delta_\kappa(k)$ refer to the value of κ. Hence, for a spin-up polarized incident beam, the scattering amplitudes are given in terms of the phase-shifts via

$$
f(\theta)\ =\ \frac{\sqrt{4\,\pi}}{2\,i\,k} \sum_l \left[\frac{(\,l\,+\,1\,)}{\sqrt{2\,l\,+\,1}} \left(\exp[\,2\,i\,\delta_{-l-1}(k)\,]\ -\ 1 \right) \right.
$$
$$
\left. +\ \frac{l}{\sqrt{2\,l\,+\,1}} \left(\exp[\,2\,i\,\delta_l(k)\,]\ -\ 1 \right) \right]\,Y_0^l(\theta,\varphi) \quad (1936)
$$

and

$$
g(\theta)\ \exp[\,i\,\varphi\,]\ =\ \frac{\sqrt{4\,\pi}}{2\,i\,k} \sum_l \sqrt{\frac{l\,(\,l\,+\,1\,)}{2\,l\,+\,1}} \left[\left(\exp[\,2\,i\,\delta_{-l-1}(k)\,]\ -\ 1 \right) \right.
$$
$$
\left. -\ \left(\exp[\,2\,i\,\delta_l(k)\,]\ -\ 1 \right) \right]\,Y_1^l(\theta,\varphi) \qquad (1937)
$$

A similar analysis can be applied to the scattering of an incident beam which is down-spin polarized, giving similar results.

If the incident beam is un-polarized, the elastic scattering cross-section is given in terms of the scattering amplitudes by

$$
\left(\frac{d\sigma}{d\Omega} \right)\ =\ \left(\mid f(\theta)\mid^2\ +\ \mid g(\theta)\mid^2 \right) \qquad (1938)
$$

where the polar angle θ is the scattering angle.

12.12   An Electron in a Uniform Magnetic Field

We shall consider a Dirac electron in a constant magnetic field $\underline{B}\ =\ B^{(z)}\,\hat{e}_z$ aligned parallel to the z direction. The vector potential can be chosen such that

$$
\underline{A}\ =\ B\,x\,\hat{e}_y \qquad (1939)
$$

We shall search for stationary states with energy E, where

$$
\psi\ =\ \left(\begin{array}{c} \phi^A \\ \phi^B \end{array} \right) \exp\left[\ -\ \frac{i}{\hbar}\,E\,t \right] \qquad (1940)
$$

In the standard representation, the energy eigenvalue equation is represented by the set of coupled equations

$$(E - m c^2) \phi^A(\underline{r}) = c \underline{\sigma} \cdot (\hat{\underline{p}} - \frac{q}{c} \underline{A}) \phi^B(\underline{r})$$

$$(E + m c^2) \phi^B(\underline{r}) = c \underline{\sigma} \cdot (\hat{\underline{p}} - \frac{q}{c} \underline{A}) \phi^A(\underline{r})$$

(1941)

Substituting the expression for ϕ^B from the second equation into the first, one obtains the second-order differential equation for ϕ^A

$$(E^2 - m^2 c^4) \phi^A = c^2 \left(\underline{\sigma} \cdot (\hat{\underline{p}} - \frac{q}{c} \underline{A}) \right)^2 \phi^A(\underline{r})$$

$$= c^2 \left((\hat{\underline{p}} - \frac{q}{c} \underline{A})^2 - \frac{q \hbar}{c} \underline{\sigma} \cdot \underline{B} \right) \phi^A(\underline{r})$$

$$= \left(\hat{p}^2 c^2 + q^2 B^2 x^2 - 2 q \hat{p}_y c B x - q c \hbar \sigma^{(z)} B \right) \phi^A(\underline{r})$$

(1942)

Since $\hat{p}_y$ and $\hat{p}_z$ commute with x, one can find simultaneous eigenstates of $\hat{H}$, $\hat{p}_y$ and $\hat{p}_z$. Hence, the two-component spinor ϕ^A can be expressed as

$$\phi^A(\underline{r}) = \exp\left[i k_y y + i k_z z \right] \Phi^A(x)$$

(1943)

in which $\Phi^A(x)$ is a two-component spinor which only depends on the variable x. In this case, the exponential term can be factored out of the eigenvalue equation. The resulting equation has the form

$$\left[- \hbar^2 c^2 \frac{\partial^2}{\partial x^2} + (c \hbar k_y - q B x)^2 - q c \hbar B \sigma^{(z)} \right] \Phi^A(x) = (E^2 - m^2 c^4 - c^2 \hbar^2 k_z^2) \Phi^A(x)$$

(1944)

The equations decouple if the two-component spinor $\Phi^A(x)$ can be taken to be an eigenstate of the z-component of the spin operator

$$\Phi^A(x) = f(x) \chi_\sigma$$

(1945)

where

$$\sigma^{(z)} \chi_\sigma = \sigma \chi_\sigma$$

(1946)

in which the eigenvalues of $\sigma^{(z)}$ are denoted by σ. Therefore, the eigenvalue equation can be reduced to

$$\left[- \hbar^2 c^2 \frac{\partial^2}{\partial x^2} + (q B)^2 \left(x - \frac{c \hbar k_y}{q B} \right)^2 \right] f(x) = (E^2 - m^2 c^4 - c^2 \hbar^2 k_z^2 + q c \hbar B \sigma) f(x)$$

(1947)

which (apart from an overall scale factor) is formally equivalent[126] to the (non-relativistic) energy eigenvalue equation for a shifted harmonic oscillator, with frequency $2\,c\,|\,q\,|\,B$. The modulus sign was inserted to ensure that the frequency ω_{HO} is positive. The energy eigenvalues are determined from

$$(E^2 \; - \; m^2\,c^4 \; - \; c^2\,\hbar^2\,k_z^2 \; + \; q\,c\,\hbar\,B\,\sigma) \; = \; 2\,|\,q\,|\,c\,\hbar\,B\,(n \; + \; \frac{1}{2}) \quad (1948)$$

Hence, for an electron with negative charge $q \; = \; -\,e$ one finds that the positive-energy eigenvalue is given by the solution

$$E \; = \; c\,\sqrt{ m^2\,c^2 \; + \; \hbar^2\,k_z^2 \; + \; (2\,n \; + \; 1 \; + \; \sigma)\,\frac{|\,e\,|\,\hbar}{c}\,B } \qquad (1949)$$

This expression has an infinite degeneracy as it is independent of the continuous variable k_y. It also has a discrete (two-fold) degeneracy between the levels with quantum numbers $(n, \sigma = 1)$ and $(n+1, \sigma = -1)$. The two-fold degeneracy can be understood as a consequence of the generalized helicity $\underline{\sigma}\,.\,(\,\hat{p} \; - \; \frac{q}{c}\,\underline{A}\,)$ commuting with the Hamiltonian $\hat{H}$. This results in the spin's alignment with the electron's velocity being preserved, as the spin's precession is precisely balanced by the electron's orbital precession. It should be noted that if the g factor deviates from 2, and such an anomaly in the g factor is expected from Quantum Electrodynamics and has been found in experiment, then this degeneracy will be lifted. The calculated $(\,g \; - \; 2\,)$ anomaly for an electron is given by

$$\left(\frac{g\,-\,2}{2}\right)_{\mathrm{Theor}} = \; \frac{1}{2}\left(\frac{\alpha}{\pi}\right) - 0.3284986\left(\frac{\alpha}{\pi}\right)^2 + 1.17611\left(\frac{\alpha}{\pi}\right)^3 - 1.434\left(\frac{\alpha}{\pi}\right)^4 + \dots$$
$$(1950)$$

where

$$\alpha \; = \; \left(\frac{e^2}{\hbar\,c}\right) \qquad (1951)$$

is the fine structure constant. The experimentally determined value of the g anomaly is found as

$$\left(\frac{g\,-\,2}{2}\right)_{\mathrm{Expt}} = \; 0.0011659208 \qquad (1952)$$

---

[126]The explicit (but dimensionally incorrect) analogy is obtained by setting the Harmonic Oscillator mass, m_{HO}, as

$$m_{\mathrm{HO}} \; = \; \frac{1}{2\,c^2}$$

and then determine the frequency from

$$m_{\mathrm{HO}}^2\,\omega_{\mathrm{HO}}^2 \; = \; \left(\frac{q\,B}{c}\right)^2$$

.

and differs from the theoretical value in the last two decimal places[127]. In the non-relativistic limit, the expression for the relativistic energy eigenvalue reproduces the expression for energies of the well-known Landau levels

$$E \approx m\,c^2 + \frac{\hbar^2\,k_z^2}{2\,m} + (\,n + \frac{1+\sigma}{2}\,)\left(\frac{|\,e\,|\,\hbar\,B}{m\,c}\right) \tag{1953}$$

which are doubly-degenerate.

12.13 Motion of an Electron in a Classical Electromagnetic Field

Consider an electron in a classical electromagnetic field represented by the real vector potential A^μ. For simplicity, electromagnetic field will be represented by a plane wave defined over Minkowski space that depends on the phase ϕ defined by

$$\phi = k_\mu\,x^\mu \tag{1954}$$

Hence, the vector potential is written as

$$A^\mu = A^\mu(\phi) \tag{1955}$$

The vector potential satisfies the Lorentz Gauge condition

$$\partial_\mu\,A^\mu = k_\mu\,A^\mu(\phi)' = 0 \tag{1956}$$

where the prime indicates differentiation with respect to ϕ. The classical vector potential must satisfy the source-free wave equation

$$\partial_\nu\,\partial^\nu\,A^\mu = k_\nu\,k^\nu\,A^\mu(\phi)'' = 0 \tag{1957}$$

which results in the condition

$$k_\nu\,k^\nu = 0 \tag{1958}$$

which is the dispersion relation for a free electromagnetic field.

The Dirac equation for a spin one-half particle with charge q can be used to obtain the second-order differential equation

$$\left[-\hbar^2\,\partial_\mu\,\partial^\mu - 2\,i\,\hbar\,\frac{q}{c}\,A^\mu\,\partial_\mu + \frac{q^2}{c^2}\,A_\mu\,A^\mu - m^2\,c^2 - i\,\hbar\,\frac{q}{c}\,\gamma^\mu\,k_\mu\,\gamma^\nu\,A_\nu(\phi)'\right]\psi = 0 \tag{1959}$$

---

[127]This discrepancy could indicate the importance of virtual processes in which heavy particle/antiparticle pairs are created. The $(g-2)$ anomalies for the muon and its anti-particle have also been measured [G. W. Bennett *et al.*, Phys. Rev. Lett. **92**, 1618102 (2004).]. These experiments show that particles and anti-particles precess at the same rate. However, the value of the $(g-2)$ anomaly is inconsistent with the theoretical prediction based on the standard model of particle physics.

where ψ is the four-component Dirac spinor. In deriving this, the Lorenz gauge condition has been used to re-write

$$\gamma^\mu \, \gamma^\nu \, \partial_\mu \left(A_\nu \, \psi \right) \;\; = \;\; \gamma^\mu \, \gamma^\nu \, \partial_\mu \left(A_\nu \, \psi \right) \; - \; g^{\mu,\nu} \left(\partial_\mu \, A_\nu \right) \psi$$

$$(1960)$$

in the diagonal terms.

Following Volkow[128], the solution of the second-order differential equation can be found in the form

$$\psi \;\; = \;\; \exp \left[\; - \; i \; \frac{p_\mu \, x^\mu}{\hbar} \; \right] F(\phi) \tag{1961}$$

where p_μ is a four-vector and $F(\phi)$ is a four-component spinor. This form reduces to the form of a free particle solution when $A^\mu \equiv 0$ in which case p_μ becomes the momentum of the free particle. The exponential form is unaltered when the vector potential is non-zero since arbitrary multiples of the electromagnetic wave vector $\underline{k}$ can be added to the momentum of the free particle, in which case p_μ has a different interpretation. For a transverse polarized vector potential describing an electromagnetic wave travelling in the $\hat{e}_3$ direction, the operators

$$\begin{aligned} \hat{p}_1 &= i \, \hbar \, \partial_1 \\ \hat{p}_2 &= i \, \hbar \, \partial_2 \end{aligned} \tag{1962}$$

commute with the time-dependent Dirac Hamiltonian and are constants of motion. Although the particle's energy and momentum operators do not commute with the Hamiltonian, as these quantities are not conserved due to the interaction with the field, the quantity

$$\hat{p}_3 \; - \; \hat{p}_0 \; = \; i \, \hbar \left(\partial_3 \; - \; \partial_0 \right) \tag{1963}$$

does commutes with the Hamiltonian and, therefore, is conserved. The conservation of this quantity can be interpreted in terms of the energy absorbed or emitted by the electron due to interaction with the classical electromagnetic field being accompanied by the absorption or emission of similar amount of momentum[129]. Despite the different interpretation of p_μ in the presence of the classical field, the four-vector p_μ shall be chosen to satisfy the condition

$$p^\mu \, p_\mu \; = \; m^2 \, c^2 \tag{1964}$$

which is the dispersion relation for a free electron[130].

---

[128] D. M. Volkow, Zeit. für Physik, **94**, 25 (1935).

[129] For the quantized electromagnetic field, the absorption of a photon involves the absorption of the energy and momentum given by the four-vector $\hbar \, k^\mu$, where $k^\mu = (k, 0, 0, k)$.

[130] If the condition on p^μ is dropped, the function $F(\phi)$ will acquire an overall phase factor that depends linearly on ϕ and on the constant value of $p^\mu \, p_\mu - m^2 \, c^2$.

The form of the wave function of eqn(1961) is to be substituted into the second-order differential eqn(1959). It shall be noted that

$$A^\mu \, \partial_\mu \, F(\phi) \;=\; k_\mu \, A^\mu \, F(\phi)' \;=\; 0$$
$$\partial^\mu \, \partial_\mu \, F(\phi) \;=\; k^\mu \, k_\mu \, F(\phi)'' \;=\; 0 \qquad (1965)$$

since A^μ satisfies the Lorenz gauge condition and k^μ satisfies the dispersion relation for electromagnetic waves in vacuum. On substituting the ansatz into the second-order equation, using the above two equations and the choice of p_μ satisfying the free-electron dispersion relation, one finds that the second-order equation reduces to a first-order differential equation for the spinor $F(\phi)$

$$2 \, i \, \hbar \, p_\mu \, k^\mu \, F(\phi)' \;=\; \left[\, 2 \, \frac{q}{c} \, A^\mu \, p_\mu \;-\; \frac{q^2}{c^2} \, A^\mu \, A_\mu \;+\; i \, \hbar \, \frac{q}{c} \, \gamma^\mu \, k_\mu \, \gamma^\nu \, A_\nu(\phi)' \,\right] F(\phi)$$
$$(1966)$$

which only depends on ϕ since the exponential phase-factor which depends on $p_\mu \, x^\mu$ has been factored out. The first-order equation can be integrated w.r.t. ϕ to yield

$$F(\phi) \;=\; \exp\left[\, -\, \frac{i \, q}{\hbar \, c \, p_\lambda \, k^\lambda} \int_0^\phi \left(\, p^\mu \, A_\mu(\phi') - \frac{1}{2} \frac{q}{c} \, A^\mu(\phi') \, A_\mu(\phi') \,\right) d\phi' + \frac{q}{c} \, \frac{\gamma^\mu \, k_\mu \, \gamma^\nu \, A_\nu}{2 \, p_\lambda \, k^\lambda} \,\right] F(0)$$
$$(1967)$$

where $F(0)$ is an arbitrary constant four-component spinor. The exponential of the matrix is defined in terms of its series expansion.

$$F(\phi) \;=\; \exp\left[\, -\, \frac{i \, q}{\hbar \, c \, p_\lambda \, k^\lambda} \int_0^\phi \left(\, p^\mu \, A_\mu(\phi') - \frac{1}{2} \frac{q}{c} \, A^\mu(\phi') \, A_\mu(\phi') \,\right) d\phi' \,\right]$$
$$\times \; \exp\left[\, \frac{q}{c} \, \frac{\gamma^\mu \, k_\mu \, \gamma^\nu \, A_\nu}{2 \, p_\lambda \, k^\lambda} \,\right] F(0) \qquad (1968)$$

The above form can be simplified by expanding the last exponential factor due to the identity

$$\left(\, \gamma^\mu \, k_\mu \, \gamma^\nu \, A_\nu \,\right)^n \;=\; 0 \qquad (1969)$$

for all integers n such that $n \, > \, 1$. The identity can be proved by

$$\gamma^\mu \, k_\mu \, \gamma^\nu \, A_\nu \, \gamma^\tau \, k_\tau \, \gamma^\rho \, A_\rho \;=\; -\, \gamma^\mu \, k_\mu \, \gamma^\tau \, k_\tau \, \gamma^\nu \, A_\nu \, \gamma^\rho \, A_\rho \;+\; 2 \, g^{\nu,\tau} \, A_\nu \, k_\tau \, \gamma^\mu \, k_\mu \, \gamma^\rho \, A_\rho$$
$$=\; -\, \gamma^\mu \, k_\mu \, \gamma^\tau \, k_\tau \, \gamma^\nu \, A_\nu \, \gamma^\rho \, A_\rho \qquad (1970)$$

where the first line follows by using the anti-commutation relations for the γ matrices and the second line follows from applying the Lorenz gauge condition. The expression can be further simplified by noting that on anticommuting the first pair of γ matrices, one has

$$=\; -\, \gamma^\mu \, k_\mu \, \gamma^\tau \, k_\tau \, \gamma^\nu \, A_\nu \, \gamma^\rho \, A_\rho$$
$$=\; \gamma^\tau \, k_\tau \, \gamma^\mu \, k_\mu \, \gamma^\nu \, A_\nu \, \gamma^\rho \, A_\rho \;+\; 2 \, g^{\mu,\tau} \, k_\mu \, k_\tau$$
$$=\; \gamma^\tau \, k_\tau \, \gamma^\mu \, k_\mu \, \gamma^\nu \, A_\nu \, \gamma^\rho \, A_\rho$$
$$=\; \gamma^\mu \, k_\mu \, \gamma^\tau \, k_\tau \, \gamma^\nu \, A_\nu \, \gamma^\rho \, A_\rho \qquad (1971)$$

the third line follows from the condition $k^\mu\, k_\mu\, =\, 0$ and the last line follows from interchanging the first two pairs of summation indices. On comparing the first and last lines, one notes that the right-hand side is zero. Therefore, one has proved the identity

$$\gamma^\mu\, k_\mu\, \gamma^\nu\, A_\nu\, \gamma^\tau\, k_\tau\, \gamma^\rho\, A_\rho\, =\, 0 \tag{1972}$$

Using, the above identity, the spinor $F(\phi)$ can be expanded as

$$F(\phi)\ =\ \exp\left[\ -\ \frac{i\, q}{\hbar\, c\, p_\lambda\, k^\lambda}\ \int_0^\phi\ \left(\ p^\mu\, A_\mu(\phi')\ -\ \frac{1}{2}\, \frac{q}{c}\, A^\mu(\phi')\, A_\mu(\phi')\ \right)\, d\phi'\ \right]$$
$$\times\ \left(\ \hat{I}\ +\ \frac{q}{c}\, \frac{\gamma^\mu\, k_\mu\, \gamma^\nu\, A_\nu}{2\, p_\lambda\, k^\lambda}\ \right)\, F(0) \tag{1973}$$

Hence, the spinor solution of the second-order differential equation can be expressed as

$$\psi(x)\ =\ \exp\left[\ i\, \frac{S}{\hbar}\ \right]\left(\ \hat{I}\ +\ \frac{q}{c}\, \frac{\gamma^\mu\, k_\mu\, \gamma^\nu\, A_\nu}{2\, p_\lambda\, k^\lambda}\ \right)\, F(0) \tag{1974}$$

where S given by

$$S\ =\ -\, p_\mu\, x^\mu\ -\ \frac{q}{c\, p_\lambda\, k^\lambda}\ \int_0^\phi\ \left(\ p^\mu\, A_\mu(\phi')\ -\ \frac{1}{2}\, \frac{q}{c}\, A^\mu(\phi')\, A_\mu(\phi')\ \right)\, d\phi' \tag{1975}$$

is the classical action of a particle moving in an electromagnetic field.

If the above equation is to be a solution of the Dirac equation, one needs to exclude redundant solutions of the second-order equation. This can be achieved by demanding that as $r\ \to\ \infty$ one has $A^\mu\ \to\ 0$. In this limit, the above solution reduces to

$$\psi\ \to\ \exp\left[\ -\ i\, \frac{p_\mu\, x^\mu}{\hbar}\ \right]\, F(0) \tag{1976}$$

which satisfies the Dirac equation if

$$\left(\ \gamma^\mu\, p_\mu\ -\ m\, c\ \right)\, F(0)\ =\ 0 \tag{1977}$$

Therefore, one demands that $F(0)$ satisfies the above supplementary condition which is the same as for a free particle. Hence, one can set

$$F(0)\ =\ \mathcal{N}_F\ \begin{pmatrix} \chi_\sigma \\ \frac{\underline{p}\, \cdot\, \underline{\sigma}}{p^{(0)}\, +\, m\, c}\, \chi_\sigma \end{pmatrix} \tag{1978}$$

where the normalization constant is given by

$$\mathcal{N}_F\ =\ \sqrt{\frac{p^{(0)}\ +\ m\, c}{2\, p^{(0)}\, V}} \tag{1979}$$

The spectrum of eigenvalues of the electron's energy can be found by Fourier transforming the above solution with respect to time, which shows that the electron absorbs and emits radiation in multiples of $\hbar \, \omega$. The Volkov solutions have been used to describe the Compton scattering of electrons by intense coherent laser beams, and is also the basis of the strong-field approximation sometimes found useful in atomic physics[131].

The current density is derived from the expression

$$j^\mu \; = \; c \, \overline{\psi}^\dagger \, \gamma^\mu \, \psi \tag{1980}$$

Since the Dirac adjoint spinor is given by

$$\overline{\psi}^\dagger \; = \; \overline{F}^\dagger(0) \, \left(\, \hat{I} \; + \; \frac{q}{c} \, \frac{\gamma^\nu \, A_\nu \, \gamma^\mu \, k_\mu}{2 \, p_\lambda \, k^\lambda} \, \right) \, \exp\left[\, - \, i \, \frac{S}{\hbar} \, \right] \tag{1981}$$

the current density is evaluated as

$$j^\mu \; = \; \frac{c}{p^{(0)} \, V} \, \left[\, p^\mu \; - \; \frac{q}{c} \, A^\mu \; + \; k^\mu \left(\, \frac{q}{c} \, \frac{p_\nu \, A^\nu}{k^\lambda \, p_\lambda} \; - \; \frac{q^2}{c^2} \, \frac{A_\nu \, A^\nu}{2 \, k^\lambda \, p_\lambda} \, \right) \right] \tag{1982}$$

Hence, the current is composed of a constant component p^μ and an oscillatory component form the vector potential, and an oscillatory component which is second order in the vector potential. This implies that the electromagnetic field has measurable consequences. For a vector potential A^μ which is a periodic function with a time-averaged value of zero, the time-averaged current density is given by

$$\overline{j^\mu} \; = \; \frac{c}{p^{(0)} \, V} \, \left[\, p^\mu \; - \; k^\mu \left(\frac{q^2}{c^2} \, \frac{\overline{A_\nu \, A^\nu}}{2 \, k^\lambda \, p_\lambda} \, \right) \right] \tag{1983}$$

which shows that the electromagnetic wave does not drop out from time-averaged quantities.

12.14 The Limit of Zero Mass

The Dirac equation has the form

$$\gamma^\mu \, \hat{p}_\mu \, \psi \; = \; m \, c \, \psi \tag{1984}$$

where the γ matrices are any set of matrices which satisfy the anti-commutation relations

$$\gamma^\mu \, \gamma^\nu \; + \; \gamma^\nu \, \gamma^\mu \; = \; 2 \, g^{\mu,\nu} \, \hat{I} \tag{1985}$$

[131]L. V. Keldysh, Zh. Eksp. Teor. Fiz. **47**, 1945 (1964). [Sov. Phys. J.E.T.P. **20**, 1307 (1965).]
F. H. M. Faisal, J. Phys. B **6**, L89 (1973).
H. R. Reiss, Phys. Rev. A **22**, 1786 (1980).

The Dirac equation is independent of the specific representation of the γ matrices. We have chosen the representation

$$\gamma^{(0)} = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{1986}$$

and

$$\gamma^{(i)} = \begin{pmatrix} 0 & \sigma^{(i)} \\ -\sigma^{(i)} & 0 \end{pmatrix} \tag{1987}$$

where $\sigma^{(i)}$ are the Pauli-matrices. This is the standard representation.

We can find other representations which differ through unitary transformations

$$\psi' = \hat{U} \psi \tag{1988}$$

where the explicit form of the γ matrices transform via

$$\gamma^{\mu\prime} = \hat{U} \gamma^\mu \hat{U}^\dagger \tag{1989}$$

and the Dirac adjoint is transformed via

$$\overline{\psi}^{\dagger\prime} = \psi'^\dagger \gamma^{(0)\prime} \tag{1990}$$

These unitary transformations of the gamma operators keep matrix elements of the form

$$\int d^3\underline{r}\, \psi^\dagger\, \hat{A}\, \psi \tag{1991}$$

invariant.

The chiral representation is found by performing the unitary transform

$$\hat{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} \tag{1992}$$

starting with the standard representation. In the chiral representation, the γ matrices have the form

$$\gamma^{(0)\prime} = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{1993}$$

and

$$\gamma^{(i)\prime} = \begin{pmatrix} 0 & \sigma^{(i)} \\ -\sigma^{(i)} & 0 \end{pmatrix} \tag{1994}$$

The components of the wave function in the chiral representation ψ' are denoted as

$$\psi' = \begin{pmatrix} \phi^{L\prime} \\ \phi^{R\prime} \end{pmatrix} \tag{1995}$$

The components ϕ^L and ϕ^R are related to the components of ψ in the standard representation via

$$\begin{pmatrix} \phi^{L\prime} \\ \phi^{R\prime} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^A - \phi^B \\ \phi^A + \phi^B \end{pmatrix} \tag{1996}$$

The chiral representation is particularly useful for the description of massless spin one-half particles, such as might be the case for the neutrino. The neutrino masses are extremely small. The masses have evaded direct experimental measurement. However, direct measurements have set upper limits on the masses which decrease with time[132]. In this case, with the limit $m \to 0$, the Dirac equation takes the form

$$\begin{pmatrix} 0 & \left(\frac{\partial}{\partial t} + c\,\underline{\sigma} \cdot \underline{\nabla} \right) \\ \left(\frac{\partial}{\partial t} - c\,\underline{\sigma} \cdot \underline{\nabla} \right) & 0 \end{pmatrix} \begin{pmatrix} \phi^{L\prime} \\ \phi^{R\prime} \end{pmatrix} = 0 \tag{1997}$$

Hence, the Dirac equation for a massless free particle reduces to two uncoupled equations, each of which are equations proposed by Weyl[133]

$$\left(\frac{\partial}{\partial t} + c\,\underline{\sigma} \cdot \underline{\nabla} \right) \phi^{R\prime} = 0 \tag{1998}$$

and

$$\left(\frac{\partial}{\partial t} - c\,\underline{\sigma} \cdot \underline{\nabla} \right) \phi^{L\prime} = 0 \tag{1999}$$

The Weyl equation describes a spin one-half massless particle by a two component spinor wave function. The Weyl equation violates parity invariance. The Weyl equation was considered to be un-physical until the discovery of the (anti-)neutrino[134] and the associated violation of parity invariance[135]. After the parity violation of the weak interaction was established, the Weyl equation was adopted to describe the neutrino[136].

Inexplicably nature seems to have selected the Weyl equation for ϕ_L, but not ϕ_R to describing neutrinos. The solutions of the Weyl equation for free particles

$$\left(\frac{\partial}{\partial t} - c\,\underline{\sigma} \cdot \underline{\nabla} \right) \phi^L = 0 \tag{2000}$$

[132]L. Langer and R. Moffat, Phys. Rev. **88**, 689 (1952).
V. A. Lyubimov, F. G. Novikov, V. Z. Nozik, F. F. Tretyakov, and V. S. Kosik, Phys. Lett. **94B**, 266 (1980).
A. I. Belesev *et al.*, Phys. Lett. **350**, 263 (1995).
[133]H. Weyl, Zeit. für Physik, **56**, 330 (1929).
[134]C. L. Cowan Jr., F. Reines, F. B. Harrison, H. W. Kruse and A. D. McGuire, Science **124**, 103 (1956).
F. Reines and C. L. Cowan Jr., Phys. Rev. **113**, 273 (1959).
[135]C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. F. Hudson, Phys. Rev. **108**, 1413 (1957).
[136]T. D. Lee and C. N. Yang, Phys. Rev. **105**, 1671 (1957).
A. Salam, Nuovo Cimento **5**, 299 (1957).
L. Landau, Nuclear Phys. **3**, 127 (1957).

can be written as

$$\phi^L \;=\; \left(\begin{array}{c} u^{(0)} \\ u^{(1)} \end{array} \right) \; \frac{1}{\sqrt{V}} \; \exp\left[\; - \; \frac{i}{\hbar} \; (\; E \, t \; - \; \underline{p} \cdot \underline{r} \;) \right] \qquad (2001)$$

Since helicity is conserved, one can choose the direction of $\underline{p}$ as the axis of quantization. The positive-energy solution is given by

$$\phi_-^L \;=\; \left(\begin{array}{c} 0 \\ 1 \end{array} \right) \; \frac{1}{\sqrt{V}} \; \exp\left[\; - \; \frac{i}{\hbar} \; (\; E \, t \; - \; p \, z \;) \right] \qquad (2002)$$

which has negative helicity and has energy given by

$$E_- \;=\; c \, p \qquad (2003)$$

The negative-energy solution is given by

$$\phi_+^L \;=\; \left(\begin{array}{c} 1 \\ 0 \end{array} \right) \; \frac{1}{\sqrt{V}} \; \exp\left[\; - \; \frac{i}{\hbar} \; (\; E \, t \; - \; p \, z \;) \right] \qquad (2004)$$

which has positive helicity and the energy is given by

$$E_+ \;=\; - \, c \, p \qquad (2005)$$

This negative-energy solution will describe anti-particles. The Weyl equation for ϕ^R has a positive-energy solution with positive helicity, and a negative-energy solution with negative helicity. Since only neutrinos with negative helicity are observed in nature, only ϕ^L is needed. The anti-neutrinos have positive helicity and are represented by ϕ^R.
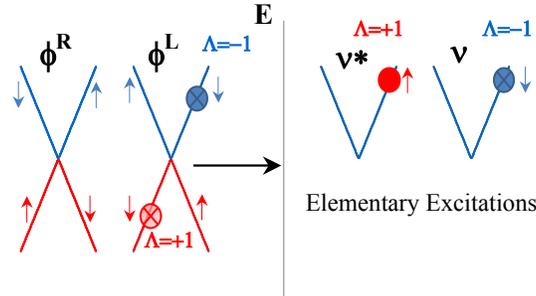

Figure 61: The dispersion relations for ϕ^L and ϕ^R. The elementary excitations are the negative-helicity neutrino ν and a positive-helicity anti-neutrino $\bar{\nu}$.

The Neutrino

The neutrino was postulated by Pauli to balance energy and momentum conservation in beta decay. In beta decay, it had been observed that neutron

decay products included a proton and an electron. However, it was observed that the emitted electron had a continuous range of kinetic energies. Therefore, another neutral particle must have been emitted in the decay. This particle was termed the anti-neutrino, and the reaction can be written as

$$n \; \rightarrow \; p \; + \; e^- \; + \; \overline{\nu}_e \tag{2006}$$

Conservation of angular momentum requires that the neutrino has a spin of $\frac{\hbar}{2}$. Furthermore, since an energy of 1.2934 MeV is released in the transformation of a neutron to a proton, and since sometime the decay processes produce electrons which seem to take up all the released energy, the neutrino was suggested as having zero mass. An upper limit on the neutrino's mass of a few eV follows from the Fermi-Kurie plot[137]. The Fermi-Kurie plot of the electron energy
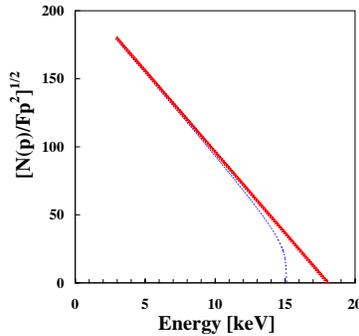
Figure 62: The Fermi-Kurie plot of the energy distribution of the electrons emitted in the beta decay of tritium, $^3H \rightarrow \; ^3He + e^- + \overline{\nu}_e$. The decay releases 18.1 keV. It is seen that the electrons produced in the decay process have a non-zero probability for carrying off most of the released energy. Hence, one concludes that the anti-neutrinos are almost massless. The dashed blue curve is the curve expected if the neutrino had a mass of 3 keV.

distribution is based on the phase space available for the emission of the electron and anti-neutrino[138]. The joint phase-space available for the electron of four-momentum $(E_e/c, p)$ and the anti-neutrino of four-momentum is $(E_\nu/c, q)$ is proportional to the factor

$$
\begin{aligned}
d\Gamma \;\; &= \;\; dp \; p^2 \; \int dq \; q^2 \; \delta(E - E_e(p) - E_\nu(q)) \\
&= \;\; dE_e(p) \; \frac{p \; E_e(p)}{c^2} \; \int dE_\nu(q) \; \frac{q \; E_\nu(q)}{c^2} \; \delta(E - E_e(p) - E_\nu(q))
\end{aligned}
$$

[137]L. Langer and R. Moffat, Phys. Rev. **88**, 689 (1952).
V. A. Lyubimov, F. G. Novikov, V. Z. Nozik, F. F. Tretyakov, and V. S. Kosik, Phys. Lett. **94B**, 266 (1980).
A. I. Belesev *et al.*, Phys. Lett. **350**, 263 (1995).
[138]E. Fermi, Zeit. für Physik, **88**, 161 (1934).
F. N. D. Kurie, J. R. Richardson and H. C. Paxton, Phys. Rev. **48**, 167 (1935).

$$= \frac{1}{c^5} \ dE_e(p) \ p \ E_e(p) \ \sqrt{E_\nu(q)^2 \ - \ m_\nu^2 \ c^4} \ E_\nu(q) \Big|_{E_\nu(q)=E-E_e(p)}$$

$$= \frac{1}{c^5} \ dE_e(p) \ p \ E_e(p) \ \sqrt{(\ E \ - \ E_e(p) \ )^2 \ - \ m_\nu^2 \ c^4} \ (\ E \ - \ E_e(p) \ )$$

$$(2007)$$

where, since the anti-neutrino's trajectory is unobservable, its momentum is integrated over. This phase-space factor partially governs the energy distribution of the emitted electrons. The second to last factor in the accessible volume of phase-space contains the dependence on the anti-neutrino's mass m_ν and it is this factor which is high-lighted by the Fermi-Kurie plot. The plot is designed to exhibit a linear energy variation until the line cuts the E-axis, if the anti-neutrino is massless. On the other hand, if the anti-neutrino has a finite mass, the line should curve over and cut the E-axis vertically. In this case, the anti-neutrino mass would be determined by the difference between the linearly extrapolated intercept and the actual intercept.

The process of beta decay does not conserve parity. The non-conservation of parity was discovered in the experiments of C. S. Wu *et al.*[139]. In these experiments, the spin of a ^{60}Co nucleus was aligned with a magnetic field. The spin $S = 5\hbar$ ^{60}Co nucleus decayed into a spin $S = 4\hbar$ ^{60}Ni nucleus by emitting an electron and an anti-neutrino.

$$^{60}Co \ \rightarrow \ ^{60}Ni \ + \ e^- \ + \ \bar{\nu}_e \qquad (2008)$$

Since angular momentum is conserved, the spin of the electron and the anti-neutrino initially must both be aligned with the field. In the experiment, the angular distribution of the emitted electrons was observed. Because the helicity of the electrons is conserved, the angular distribution of the electrons can be used to prove that the electrons all have negative helicity, and hence it is inferred that the anti-neutrinos should have positive helicity. Since helicity should be reversed under the parity operation, and since only negative helicity electrons are observed, the process is not invariant under parity. Hence, parity is not conserved.

The electrons that are emitted in beta decay have negative helicities. If the momentum of an emitted electron is given by (p, θ_p, φ_p), then its helicity operator is

$$\Lambda_p \ = \ \begin{pmatrix} \cos\theta_p & \sin\theta_p \exp[-i\varphi_p] \\ \sin\theta_p \exp[+i\varphi_p] & -\cos\theta_p \end{pmatrix} \qquad (2009)$$

The helicity operator has eigenstates χ given by

$$\Lambda_p \ \chi_{\theta_p}^{\pm} \ = \ \pm \ \chi_{\theta_p}^{\pm} \qquad (2010)$$

[139]C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. F. Hudson, Phys. Rev. **108**, 1413 (1957).

which are determined as

$$\chi_{\theta_p}^+ = \begin{pmatrix} \cos\frac{\theta_p}{2}\exp[-i\frac{\varphi_p}{2}] \\ \sin\frac{\theta_p}{2}\exp[+i\frac{\varphi_p}{2}] \end{pmatrix}$$

$$\chi_{\theta_p}^- = \begin{pmatrix} -\sin\frac{\theta_p}{2}\exp[-i\frac{\varphi_p}{2}] \\ \cos\frac{\theta_p}{2}\exp[+i\frac{\varphi_p}{2}] \end{pmatrix} \qquad (2011)$$

Since angular momentum is conserved and the emitted electrons only have negative helicity, the angular distribution of the emitted electrons is proportional to the square of the overlap of the initial electron spin-up spinor with the negative helicity spinors

$$\mid \chi_{\theta=0}^{+\dagger} \chi_{\theta_p}^- \mid^2 = \sin^2\frac{\theta_p}{2}$$

$$= \frac{1}{2}\,(\,1\,-\,\cos\theta_p\,) \qquad (2012)$$

which is in exact agreement with the experimentally observed distribution. From the distribution of emitted electrons one is led to expect that the anti-neutrino has positive helicity.

The helicity of the neutrino was measured in an experiment performed by Maurice Goldhaber *et al.*[140]. In the experiment, a ^{152}Eu nucleus with $J = 0$

---

[140]M. Goldhaber, I. Grodzins, A. W. Sunyar, Phys. Rev. **109**, 1015 (1958).
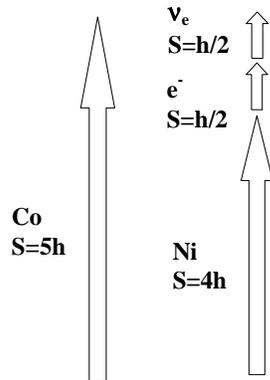
Figure 63: The spin $S = 5\hbar$ of the Co nucleus is aligned with the magnetic field. The Co undergoes beta decay to Ni which has $S = 4\hbar$ by emitting an electron e^- and an anti-neutrino $\bar{\nu}_e$. The spin of the electron and the anti-neutrino produced by the decay must initially be aligned with the magnetic field, due to conservation of angular momentum.
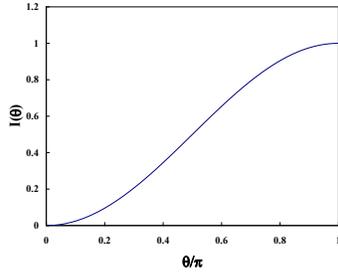
Figure 64: The angular distribution of the emitted electron in the beta decay experiment of Wu *et al.*

captures an electron from the K-shell and decays to the excited state of a ^{152}Sm nucleus with angular momentum $J = \hbar$ and emits a neutrino.

$$^{152}Eu + e^- \rightarrow {}^{152}Sm^* + \nu_e \tag{2013}$$

The $J = \hbar$ excited state of Sm^* subsequently decays into the $J = 0$ ground state of Sm by emitting a photon.

$$^{152}Sm^* \rightarrow {}^{152}Sm + \gamma \tag{2014}$$

Goldhaber *et al.* measured the photons with the full Doppler shift, from which they were able to infer the direction of the recoil of the nucleus. The photons
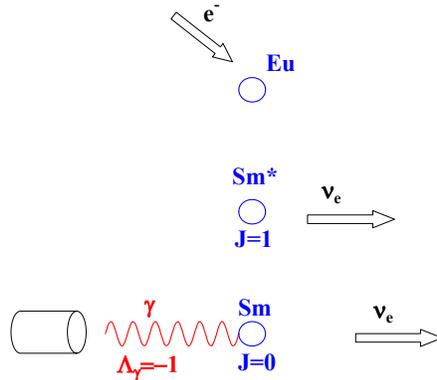
Figure 65: A schematic depiction of the experiments of Goldhaber *et al.* which determined the helicity of the neutrino.

were observed to be right-circularly polarized, which corresponds to having a negative helicity. Therefore, the photon's spin was parallel to the momentum of the emitted neutrino. Since the ground state of Sm has zero angular momentum, the excited state of the Sm^* nucleus must have had its angular momentum

oriented along the direction of motion of the emitted neutrino. Since the sum of the angular momentum of the excited state $(J = \hbar)$ and the emitted neutrino must equal the spin of the captured electron $\frac{\hbar}{2}$, the neutrino must have its spin oriented anti-parallel to the angular momentum of the Sm^* nucleus. Hence, the neutrino has negative helicity.

12.15   Classical Dirac Field Theory

The Dirac Lagrangian density is given by

$$\mathcal{L} = \overline{\psi}^\dagger \left(i \, \hbar \, c \, \gamma^\mu \, (\, \partial_\mu \, + \, i \, \frac{q}{c \, \hbar} \, A_\mu \,) \, - \, m \, c^2 \right) \psi \qquad (2015)$$

which, since $\psi^\dagger$ and ψ are independent, the momentum conjugate to ψ is

$$\begin{aligned}
\Pi \; &= \; \frac{1}{c} \, \frac{\partial \mathcal{L}}{\partial(\partial_0 \psi)} \\
&= \; i \, \hbar \, \overline{\psi}^\dagger \, \gamma^{(0)} \\
&= \; i \, \hbar \, \psi^\dagger \qquad (2016)
\end{aligned}$$

The momentum conjugate to $\psi^\dagger$ vanishes

$$\begin{aligned}
\Pi^\dagger \; &= \; \frac{1}{c} \, \frac{\partial \mathcal{L}}{\partial(\partial_0 \psi^\dagger)} \\
&= \; 0 \qquad (2017)
\end{aligned}$$

The Lagrangian equation of motion is found from the variational principle which states that the action is extremal with respect to ψ and $\psi^\dagger$. The condition that the action is extremal with respect to variations in $\psi^\dagger$ leads to the Dirac equation

$$i \, \hbar \, \gamma^\mu \, (\, \partial_\mu \, + \, i \, \frac{q}{c \, \hbar} \, A_\mu \,) \, \psi \; = \; m \, c \, \psi \qquad (2018)$$

after the resulting equation has been multiplied by a factor of $\gamma^{(0)}$. On making a variation of the action with respect to ψ, one finds the Hermitean conjugate equation

$$- \, i \, \hbar \, c \, (\, \partial_\mu \, - \, i \, \frac{q}{c \, \hbar} \, A_\mu \,) \, \overline{\psi}^\dagger \, \gamma^\mu \, - \, m \, c^2 \, \overline{\psi}^\dagger \; = \; 0 \qquad (2019)$$

That this is the Hermitean conjugate of the Dirac equation can be shown by taking its Hermitean conjugate, which results in

$$i \, \hbar \, c \, \gamma^{\mu\dagger} \, \gamma^{(0)} \, (\, \partial_\mu \, + \, i \, \frac{q}{c \, \hbar} \, A_\mu \,) \, \psi \, - \, m \, c^2 \, \gamma^{(0)} \, \psi \; = \; 0 \qquad (2020)$$

The above equation can be reduced to the conventional form by multiplying by $\gamma^{(0)}$ and by using the identities

$$\begin{aligned}
\gamma^{(0)} \, \gamma^{(0)} \; &= \; \hat{I} \\
\gamma^{(0)} \, \gamma^{\mu\dagger} \, \gamma^{(0)} \; &= \; \gamma^\mu \qquad (2021)
\end{aligned}$$

Hence, the equation found by varying ψ is just the Hermitean conjugate of the Dirac equation

$$i \, \hbar \, \gamma^\mu \, (\, \partial_\mu \, + \, i \, \frac{q}{c \, \hbar} \, A_\mu \,) \, \psi \; = \; m \, c \, \psi \qquad (2022)$$

Furthermore, it is surmised that the starting Lagrangian is appropriate to describe the Dirac field theory.

The Hamiltonian density $\mathcal{H}$ is determined from the Lagrangian by the usual Legendre transformation process

$$
\begin{aligned}
\mathcal{H} \; &= \; c \, \Pi \, \partial_0 \, \psi \, + \, c \, \Pi^\dagger \, \partial_0 \, \psi^\dagger \, - \, \mathcal{L} \\
&= \; i \, \hbar \, c \, \psi^\dagger \, \partial_0 \, \psi \, - \, \mathcal{L} \\
&= \; i \, \hbar \, c \, \overline{\psi}^\dagger \, \gamma^{(0)} \, \partial_0 \, \psi \, - \, \mathcal{L} \\
&= \; - \, i \, \hbar \, c \, \overline{\psi}^\dagger \, \underline{\gamma} \, . \, (\, \underline{\nabla} \, - \, i \, \frac{q}{c \, \hbar} \, \underline{A} \,) \, \psi \, + \, \overline{\psi}^\dagger \, (\, m \, c^2 \, + \, q \, \gamma^{(0)} \, A^{(0)} \,) \, \psi
\end{aligned}
$$
$$(2023)$$

where the relation between the covariant components of the vector potential to the contravariant components $A_{(i)} \; = \; - \, A^{(i)}$ has been used in the last line. The result is identifiable with the Hamiltonian density that appears in the usual expression for the quantum mechanical expectation value for the energy for the Dirac electron.

The set of conserved quantities can be obtained from Noether's theorem. The momentum-energy tensor $T^\mu{}_\nu$ is given by

$$T^\mu{}_\nu \; = \; \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \, \partial_\nu \psi \, + \, \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^\dagger)} \, \partial_\nu \psi^\dagger \, - \, \delta^\mu{}_\nu \, \mathcal{L} \qquad (2024)$$

which is evaluated as

$$
\begin{aligned}
T^\mu{}_\nu \; &= \; i \, \hbar \, c \, \overline{\psi}^\dagger \, \gamma^\mu \, \partial_\nu \psi \, - \, \delta^\mu{}_\nu \, \mathcal{L} \\
&= \; i \, \hbar \, c \, \overline{\psi}^\dagger \, \gamma^\mu \, \partial_\nu \psi \, + \, \delta^\mu{}_\nu \, \left[\, - \, i \, \hbar \, c \, \overline{\psi}^\dagger \, \gamma^\rho \, (\, \partial_\rho \, + \, i \, \frac{q}{c \, \hbar} \, A_\rho \,) \, \psi \, + \, m \, c^2 \, \overline{\psi}^\dagger \, \psi \, \right]
\end{aligned}
$$
$$(2025)$$

Hence, one finds the energy density $T^0{}_0$ is given by

$$
\begin{aligned}
T^0{}_0 \; &= \; - \, i \, \hbar \, c \, \overline{\psi}^\dagger \, \underline{\gamma} \, . \, (\, \underline{\nabla} \, - \, i \, \frac{q}{c \, \hbar} \, \underline{A} \,) \, \psi \, + \, \overline{\psi}^\dagger \, (\, m \, c^2 \, + \, q \, \gamma^{(0)} \, A^{(0)} \,) \, \psi \\
&= \; - \, i \, \hbar \, c \, \psi^\dagger \, \underline{\alpha} \, . \, (\, \underline{\nabla} \, - \, i \, \frac{q}{c \, \hbar} \, \underline{A} \,) \, \psi \, + \, \psi^\dagger \, (\, \beta \, m \, c^2 \, + \, q \, A^{(0)} \,) \, \psi
\end{aligned}
$$
$$(2026)$$

which is the Hamiltonian density $\mathcal{H}$. On integrating over all space, one sees that the energy of the Dirac Field is equal to the expectation value of the Dirac Hamiltonian operator

$$\int \, d^3\underline{r} \, T^0{}_0 \; = \; \int \, d^3\underline{r} \, \psi^\dagger \, \hat{H} \, \psi \qquad (2027)$$

354

Likewise, (c times) the momentum density $T^0{}_j$ is found from

$$
\begin{aligned}
T^0{}_j &= i \, \hbar \, c \, \overline{\psi}^\dagger \, \gamma^{(0)} \, \partial_j \psi \\
&= i \, \hbar \, c \, \psi^\dagger \, \partial_j \psi \\
&= c \, \psi^\dagger \, \hat{p}_j \, \psi
\end{aligned}
\tag{2028}
$$

where the partial derivative has been identified with the covariant momentum operator. Hence, the contravariant component of the momentum is given by

$$
\begin{aligned}
T^{0,j} &= - \, i \, \hbar \, c \, \psi^\dagger \, \frac{\partial}{\partial x^j} \, \psi \\
&= c \, \psi^\dagger \, \hat{p}^{(j)} \, \psi
\end{aligned}
\tag{2029}
$$

where the usual (contravariant) momentum operator is defined as

$$
\hat{p}^{(j)} = - \, i \, \hbar \, \frac{\partial}{\partial x^j}
\tag{2030}
$$

Therefore, the j-th component of the momentum is given by

$$
P^{(j)} = \frac{1}{c} \int d^3\underline{r} \, T^{0,j} = \int d^3\underline{r} \, \psi^\dagger \, \hat{p}^{(j)} \, \psi
\tag{2031}
$$

which is equal to the expectation value of the momentum operator.

One can also determine the conserved Noether charges by noting that the Lagrangian is invariant under a global gauge transformation

$$
\begin{aligned}
\psi \rightarrow \psi' &= \exp\left[\, + \, i \, \varphi \, \right] \psi \\
\psi^* \rightarrow \psi'^* &= \exp\left[\, - \, i \, \varphi \, \right] \psi^*
\end{aligned}
\tag{2032}
$$

where φ is a constant real number. The infinitesimal global gauge transformation produces a variation in the (independent) fields

$$
\begin{aligned}
\delta\psi &= + \, i \, \delta\varphi \, \psi \\
\delta\psi^* &= - \, i \, \delta\varphi \, \psi^*
\end{aligned}
\tag{2033}
$$

Since the Lagrangian is invariant under the transformation, then

$$
\delta\mathcal{L} = 0
\tag{2034}
$$

so we have

$$
\begin{aligned}
0 &= \delta \, \mathcal{L} \\
&= \left(\frac{\partial \mathcal{L}}{\partial \psi} \right) \delta\psi \, + \, \left(\frac{\partial \mathcal{L}}{\partial \psi^*} \right) \delta\psi^* \, + \, \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) \delta(\partial_\mu \psi) \, + \, \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} \right) \delta(\partial_\mu \psi^*)
\end{aligned}
\tag{2035}
$$

After substituting the Euler-Lagrange equations for the derivatives w.r.t. the fields ψ and ψ^*, the variation is expressed as

$$0 \;=\; \partial_\mu \left[\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) \delta\psi \;+\; \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} \right) \delta\psi^* \right] \tag{2036}$$

For an arbitrary gauge transformation through the fixed infinitesimal angle $\delta\varphi$, this condition becomes

$$0 \;=\; i\, \delta\varphi \, \partial_\mu \left[\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) \psi \;-\; \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} \right) \psi^* \right] \tag{2037}$$

Hence, one finds that there is a current j^μ which satisfies the continuity equation

$$\partial_\mu \, j^\mu \;=\; 0 \tag{2038}$$

where (apart from the infinitesimal constant of proportionality) the current is given by

$$j^\mu \;\propto\; i\, \delta\varphi \left[\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) \psi \;-\; \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} \right) \psi^* \right] \tag{2039}$$

For the Dirac Lagrangian, the second term is identically zero and the first term is non-zero. Hence, on adopting a conventional normalization, the conserved current is identified as

$$j^\mu \;=\; c\, \overline{\psi}^\dagger \, \gamma^\mu \, \psi \tag{2040}$$

This is the same expression for the conserved current that was previously derived for the one-electron Dirac equation. Hence, the one-particle Dirac equation yields the same expectation values and obeys the same conservation laws as the (classical) Dirac field theory.

12.15.1 Chiral Gauge Symmetry

In the limit of zero mass, the Dirac Lagrangian takes the form

$$\mathcal{L} \;=\; i\, \hbar\, c\, \overline{\psi}^\dagger \, \gamma^\mu \, \partial_\mu \, \psi \tag{2041}$$

Starting with the standard representation and making the unitary transform

$$\hat{U} \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} \tag{2042}$$

one finds that in the chiral representation the Dirac Lagrangian reduces to

$$\mathcal{L} \;=\; i\, \hbar\, c\, \left(\phi^{\dagger L} \, \sigma_L^\mu \, \partial_\mu \, \phi^L \;+\; \phi^{\dagger R} \, \sigma_R^\mu \, \partial_\mu \, \phi^R \right) \tag{2043}$$

where ϕ^L and ϕ^R are two-component Dirac spinors and the two sets of quantities σ^μ and $\tilde{\sigma}^\mu$ are expressed in terms of the Pauli matrices as

$$\begin{aligned} \sigma_L^\mu &= (\, \sigma^0 \,,\, -\, \underline{\sigma} \,) \\ \sigma_R^\mu &= (\, \sigma^0 \,,\, \underline{\sigma} \,) \end{aligned} \tag{2044}$$

The difference between σ_L^μ and σ_R^μ reflect the different chirality of ϕ^L and ϕ^R. In the absence of the mass term, the Dirac Lagrangian possesses two independent scalar gauge transformations. These transformations corresponds to the global gauge transformations

$$\phi^L \rightarrow \phi^{L\prime} = \phi^L \exp\left[i \, \theta_L \right]$$

$$\phi^R \rightarrow \phi^{R\prime} = \phi^R \exp\left[i \, \theta_R \right] \qquad (2045)$$

where θ_L and θ_R are independent angles. The Lagrangian has a $U(1) \times U(1)$ gauge symmetry. The presence of a mass term would couple the two fields and reduce the gauge transformation to one in which $\theta_R = \theta_L$.

In the chiral representation, the Hermitean matrix defined by

$$\gamma^{(4)} = i \, \gamma^{(0)} \, \gamma^{(1)} \, \gamma^{(2)} \, \gamma^{(3)} \qquad (2046)$$

takes the form

$$\gamma^{(4)} = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix} \qquad (2047)$$

The general gauge transformations for the massless fermion can be expressed as the product of two independent transformations

$$\psi \rightarrow \psi' = \exp\left[i \left(\frac{\theta_L + \theta_R}{2} \right) \hat{I} \right] \, \exp\left[i \left(\frac{\theta_R - \theta_L}{2} \right) \gamma^{(4)} \right] \psi \qquad (2048)$$

where ψ is a four-components spinor

$$\psi = \begin{pmatrix} \phi^L \\ \phi^R \end{pmatrix} \qquad (2049)$$

The first factor represents the usual global gauge transformation for the Dirac Lagrangian with finite mass. This transformation yields the usual conserved four-vector current j_V^μ defined by

$$j_V^\mu = c \, \overline{\psi}^\dagger \, \gamma^\mu \, \psi \qquad (2050)$$

The second factor is specific to the Dirac Lagrangian with zero mass. It is called the chiral transformation or axial $U(1)$ transformation. Using the anti-commutation relation

$$\{ \, \gamma^{(4)} \, , \, \gamma^\mu \, \}_+ = 0 \qquad (2051)$$

one can show that the exponential factor in the chiral gauge transformation has the property that

$$\gamma^\mu \, \exp\left[i \left(\frac{\theta_R - \theta_L}{2} \right) \gamma^{(4)} \right] = \exp\left[- \, i \left(\frac{\theta_R - \theta_L}{2} \right) \gamma^{(4)} \right] \gamma^\mu \qquad (2052)$$

This property can be used to show that the Lagrangian is invariant under the chiral transformation because

$$
\begin{aligned}
\overline{\psi}^{\dagger\,\prime}\,\gamma^\mu\,\partial_\mu\,\psi' \;&=\; \psi^\dagger\,\exp\left[\,-\,i\left(\frac{\theta_R\,-\theta_L}{2}\right)\gamma^{(4)}\right]\gamma^{(0)}\,\gamma^\mu\,\partial_\mu\,\exp\left[\,i\left(\frac{\theta_R\,-\theta_L}{2}\right)\gamma^{(4)}\right]\psi \\
&=\; \psi^\dagger\,\gamma^{(0)}\,\gamma^\mu\,\partial_\mu\,\psi \\
&=\; \overline{\psi}^{\dagger}\,\gamma^\mu\,\partial_\mu\,\psi
\end{aligned}
\tag{2053}
$$

which involves two commutations. Since the massless Dirac Lagrangian is invariant under the chiral transformation, Noether's theorem shows that the current

$$
j_A^\mu \;=\; c\,\overline{\psi}^{\dagger}\,\gamma^\mu\,\gamma^{(4)}\,\psi
\tag{2054}
$$

is conserved. This conserved current transforms like a vector under proper orthochronous Lorentz transformations but does not transform as a vector under improper orthochronous transformations. Therefore, the current is an axial current. The conserved axial density $j_A^{(0)}$ is given by

$$
\begin{aligned}
j_A^{(0)} \;&=\; \overline{\psi}^{\dagger}\,\gamma^{(0)}\,\gamma^{(4)}\,\psi \\
&=\; \psi^\dagger\,\gamma^{(4)}\,\psi \\
&=\; -\,\phi^{\dagger L}\,\phi^L \;+\; \phi^{\dagger R}\,\phi^R
\end{aligned}
\tag{2055}
$$

which is the difference between the number of particles with positive helicity and the number of particles with negative helicity.

In the presence of a mass m, the Dirac Lagrangian in the chiral representation is

$$
\mathcal{L} \;=\; i\,\hbar\,c\left(\phi^{\dagger L}\,\sigma_L^\mu\,\partial_\mu\,\phi^L \;+\; \phi^{\dagger R}\,\sigma_R^\mu\,\partial_\mu\,\phi^R\right) \;-\; m\,c^2\left(\phi^{\dagger L}\,\phi^R \;+\; \phi^{\dagger R}\,\phi^L\right)
\tag{2056}
$$

and one finds that the axial current is not conserved because the mass term is not invariant and acts like a current source

$$
\partial_\mu\,j_A^\mu \;=\; i\,\frac{2\,m\,c}{\hbar}\,\overline{\psi}^{\dagger}\,\gamma^{(4)}\,\psi
\tag{2057}
$$

To summarize, just like the Proca equation yields a zero mass for the photon if one imposes $U(1)$ gauge invariance on the electromagnetic field[141], the neutrino must have zero mass if one imposes a global $U(1)$ chiral gauge invariance. Furthermore, the existence of conservation of chirality for the massless neutrino implies that the weak interaction must involve a coupling proportional to a factor of either $(\hat{I}+\gamma^{(4)})$ or $(\hat{I}-\gamma^{(4)})$.

---

[141]Schwinger has noted that the condition of gauge invariance does not necessarily result in the photon being massless. He argued that if the electromagnetic coupling strength were larger, the photon could have finite mass. [J. Schwinger, "Gauge Invariance and Mass", Physical Review, 125, 397 (1962).]

Exercise:

By considering an infinitesimal chiral gauge transformation on the Lagrangian for massive Dirac particles, determine $\delta\mathcal{L}$ and show that this leads to the axial current j_A^μ not being conserved.

12.16   Hole Theory

The negative-energy solutions of the Dirac equation lead to the conclusion that one-particle quantum mechanics is an inadequate description of nature. In classical mechanics, the dispersion relation for a free particle is found to be given by

$$E = \pm \sqrt{m^2 \, c^4 + p^2 \, c^2} \qquad (2058)$$

The negative-energy states found in classical mechanics can be safely ignored. The rational for ignoring the negative-energy states in classical mechanics is that, the dynamics is governed by a set of differential equations which result in the classical variables changing in a continuous fashion. Since the particle's energy can only change in a continuous fashion, there is no mechanism which allows it to connect with the negative branch of the dispersion relation. However, in quantum mechanics, particles can make discontinuous transitions between different energy levels, by emitting photons. Hence, if one has a single electron in a positive-energy state where $E > m \, c^2$, this state would be unstable to the electron making a transition to a negative-energy state which occurs with the simultaneous emission of photons which carry away an energy greater than $2 \, m \, c^2$. The transition rate for such process is quite large, therefore, one might conclude that positive-energy particles should not exist in nature. Furthermore, if one does have particles in the negative-energy branch, they might be able to further lower their energies by multiple photon emission processes. Hence, the states of negative-energy particles with finite momenta could be unstable to states in which the momentum has an infinite value.

Dirac noted that if the negative-energy states were all filled, then the Pauli exclusion principle would prevent the decay of positive-energy particles into the negative-energy states. Furthermore, in the absence of any positive-energy particles, the Pauli exclusion principle would cause the set of particles in the negative-energy state to be completely inert. In this picture, the filled sea of negative-energy states would represent the physical vacuum, and would be unobservable in experiments. For example, if charge is measured, it is the non-uniform part of the charge distribution that is measured, but the infinite number of particles in the negative-energy states do produce a uniform charge density. Likewise, when energies are measured, the energy is usually measured with respect to some reference level. For the case of a vacuum in which all the negative-energy states are filled with electrons, the measured energies correspond to energy differences and so the infinite negative energy of the vacuum
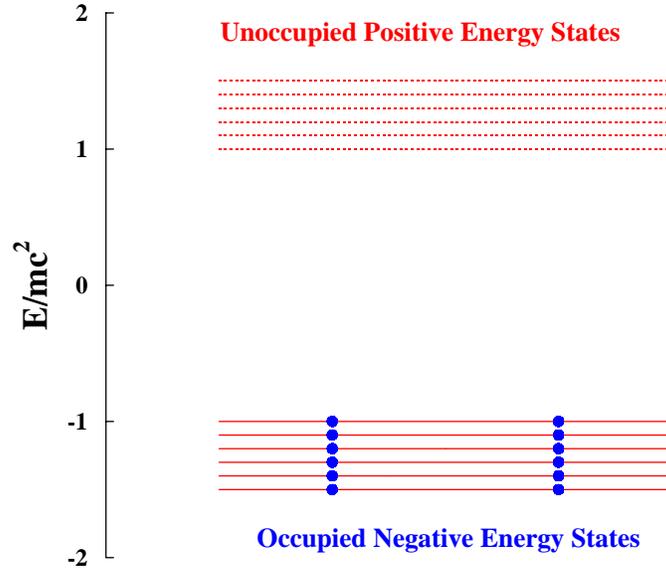
Figure 66: A cartoon depicting the vacuum for Dirac's Hole Theory, in which the negative-energy states are filled and the positive-energy states are empty.

should cancel. Therefore, Dirac postulated that the vacuum consists of the state in which all the negative-energy states are all filled with electrons[142]. Furthermore, physical states correspond to the states were a relatively few of the positive-energy states are filled with electrons and a few negative states are un-occupied. In this case, the electrons in the positive-energy states are identified with observable electrons, and the unfilled states or holes in the distribution of negative-energy states are also observable. These holes are known as positrons and are the anti-particles of the electrons. The properties of a positron are found by computing the difference between the property for a state with an absent negative-energy electron and the property of the vacuum state.

We shall assume that the vacuum contains of N electrons which completely fill all the N negative states and, for simplicity of discussion, the effect of coupling to the electromagnetic field can be ignored. Then the charge of a positron q_p is the difference between the charge of the vacuum with one missing electron, and the charge of the vacuum

$$q_p = (N - 1) q_e - N q_e \qquad (2059)$$

Therefore, one finds that the positron has the opposite charge to that of an electron

$$q_p = - q_e \qquad (2060)$$

---

[142]P. A. M. Dirac, Proc. Roy. Soc. **A 126**, 360 (1930).

360

Hence, the positron has a positive charge. Likewise, the energy of the vacuum in which all the electrons occupy all the negative-energy states is denoted by E_0. The positron energy will be denoted as $E_p(\underline{p}_e)$. The positron corresponds all states with negative energy being filled except for the state with the energy

$$E_e(\underline{p}_e) \; = \; - \; \sqrt{m^2 \; c^4 \; + \; p_e^2 \; c^4} \tag{2061}$$

which is unfilled. The positron energy is defined as the energy difference

$$
\begin{aligned}
E_p(\underline{p}_p) \;&=\; \left(\; E_0 \;-\; E_e(\underline{p}_e) \; \right) \;-\; E_0 \\
&=\; -\; E_e(\underline{p}_e) \\
&=\; \sqrt{m^2 \; c^4 \; + \; p_e^2 \; c^4}
\end{aligned}
\tag{2062}
$$

Therefore, the positron corresponds to a particle with a positive energy. From this it is seen that the rest mass energy of the positron is identical to the rest mass energy of the electron. If the vacuum corresponds to a state with momentum $\underline{P}_0$ and if the negative-energy state with momentum $\underline{p}_e$ is unfilled, then the momentum of the positron would be given by $\underline{p}_p$ where

$$
\begin{aligned}
\underline{p}_p \;&=\; \left(\; \underline{P}_0 \;-\; \underline{p}_e \; \right) \;-\; \underline{P}_0 \\
&=\; -\; \underline{p}_e
\end{aligned}
\tag{2063}
$$

Hence, the momentum of the positron is the negative of the momentum of the missing electron

$$\underline{p}_p \; = \; - \; \underline{p}_e \tag{2064}$$

Likewise, the spin of the positron is opposite to the spin of the missing electron, etc. The velocity of an electron is defined as the group velocity of a wave packet of momentum $\underline{p}_e$. Hence, one finds the velocity of the negative energy-electron from

$$
\begin{aligned}
\underline{v}_e \;&=\; \frac{\partial}{\partial \underline{p}_e} \; E_e(\underline{p}_e) \\
&=\; -\; \frac{\underline{p}_e \; c^2}{\sqrt{m^2 \; c^4 \;+\; \underline{p}_e^2 \; c^2}}
\end{aligned}
\tag{2065}
$$

while the velocity of the positron is given by

$$
\begin{aligned}
\underline{v}_p \;&=\; \frac{\partial}{\partial \underline{p}_p} \; E_p(\underline{p}_p) \\
&=\; \frac{\underline{p}_p \; c^2}{\sqrt{m^2 \; c^4 \;+\; \underline{p}_p^2 \; c^2}}
\end{aligned}
$$

Table 18: The relation between properties of Negative Energy Electron and Positron States.

Particle	Charge	Energy	Momentum	Spin	Helicity	Velocity
Electron	$-\mid e \mid$	$-\mid E \mid$	$+\,\underline{p}$	$+\,\frac{\hbar}{2}\,\sigma$	$\underline{\sigma}\cdot\underline{p}$	$\underline{v}$
Positron	$+\mid e \mid$	$+\mid E \mid$	$-\,\underline{p}$	$-\,\frac{\hbar}{2}\,\sigma$	$\underline{\sigma}\cdot\underline{p}$	$\underline{v}$

$$
\begin{aligned}
&= \quad -\,\frac{\underline{p}_e\,c^2}{\sqrt{m^2\,c^4\,+\,\underline{p}_e^2\,c^2}} \\
&= \quad \underline{v}_e
\end{aligned}
\tag{2066}
$$

Therefore, the positron and the negative-energy electron states have the same velocities.

Hole theory provides a simple description of the relation between a negative-energy state and anti-particle states. Mathematically, this relation is expressed in terms of the charge conjugation transformation. A unique signature of the hole theory is that a positive-energy electron can make a transition to an unfilled negative-energy state emitting radiation, which corresponds to the process in which a electron-positron pair annihilates[143]

$$
e\,+\,\bar{e}\,\rightarrow\,2\,\gamma
\tag{2067}
$$

In this process, it is necessary that the excess energy be carried off by two photons if the energy-momentum conservation laws are to be satisfied. Likewise, by supplying an energy greater than a threshold energy of $2\,m\,c^2$, it should be possible to promote an electron from a negative-energy state, thereby creating an electron-positron pair. Since it is unlikely that more than one photon can be absorbed simultaneously, electron-positron pair creation only occurs in the vicinity of a charged nucleus which can carry off any excess momentum.

$$
\gamma\,\rightarrow\,e\,+\,\bar{e}
\tag{2068}
$$

The positively charged electron, predicted by Dirac, was found experimentally

---

[143]P. A. M. Dirac, Proc. Camb. Phil. Soc. **26**, 361 (1930).

Unoccupied Positive Energy States

2

1

0

E/mc²

(k,α)

-1

-2

**Occupied Negative Energy States**
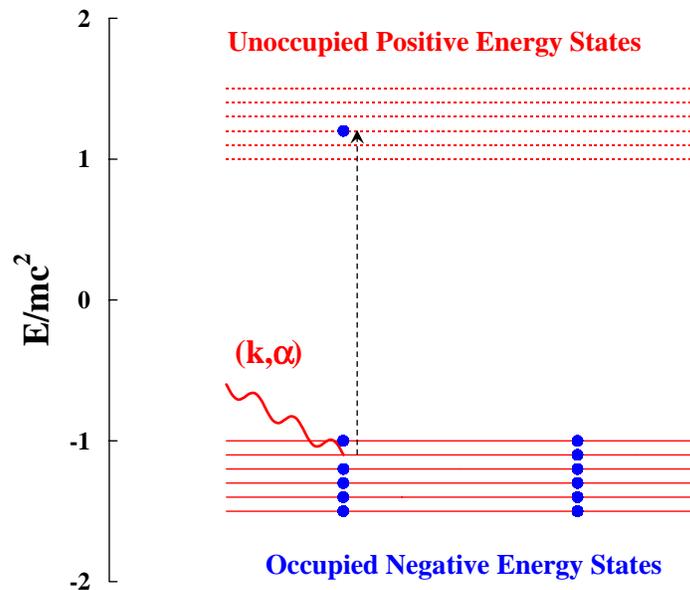
Figure 67: A cartoon depicting electron-positron production in Dirac's Hole Theory. In this case, an incident γ-ray produces an electron-hole pair. The process is restricted to occur in the vicinity of heavy particles that can act as a momentum sinks.

by Anderson[144] and the electron-positron creation[145] and annihilation processes[146] were observed shortly afterwards.

Dirac commented[147] that in scattering processes involving low-energy electrons, such as Thomson scattering, it is essential that negative-energy states appear as virtual states, if one is to recover the correct scattering cross-section in the non-relativistic limit. The involvement of negative-energy states in the scattering of light is a consequence that, in the standard representation, the lower two-component spinor in the Dirac wave function for a free (positive-energy) electron vanishes in the low energy limit, and also because the coupling

[144]C. D. Anderson, Phys. Rev. **43**, 491 (1933). Anderson observed the curved trajectories of the charged particles in a cloud chamber in the presence of a magnetic field. Anderson inferred the charge of the particles from their direction of motion. The insertion of a lead plate in a cloud chamber caused the particles to lose energy on one side of the plate which was observed as a change in the radius of curvature of the particle's track. Therefore, the examination of the radius of curvature of the track on both sides of the plate allowed the direction of motion to be established.

[145]P. M. S. Blackett and G. P. S. Occhialini, Proc. Roy. Soc. **A 139**, 688 (1933). These authors were the first who correctly identified the positively charged particle as the anti-particle of the electron, in full accord with the predictions of Dirac's hole theory.

[146]J. Thibaud, Phys. Rev, **35**, 78 (1934).

[147]P. A. M. Dirac, Proc. Roy. Soc. **A126**, 360 (1930).

to the radiation field is produced by $\gamma^{(0)} \underline{\gamma} \cdot \underline{A}$. The interaction operator can be expressed as

$$\hat{H}_{Int} \; = \; - \, q \, \underline{\alpha} \cdot \underline{A} \; = \; - \, q \, \begin{pmatrix} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{pmatrix} \cdot \underline{A} \qquad (2069)$$

which only connects the upper and lower two-component spinors of the initial and final states ψ_n and $\psi_{n'}$. Hence, as light scattering processes are at least of second-order in $\underline{A}$, the intermediate state $\psi_{n''}$ must involve a negative-energy electron state. Since the Pauli exclusion principle forbids the occupation of the filled negative-energy states, hole theory ascribes the intermediate states as involving virtual electron-positron creation and annihilation processes. This shows that, even for processes which appear to involve a single electron in the initial and final states, one must abandon single-particle quantum mechanics and adopt a multi-particle description. Therefore, a purely single-particle description is inadequate and one must consider a many-particle description such as quantum field theory.

12.16.1 Compton Scattering

We shall consider Thomson scattering of light by free electrons. In this process, light is scattered from the initial state $(\underline{k}, \alpha)$ to the final state $(\underline{k}, \alpha)$ and the (positive-energy) electron makes a transition from the initial state $(\underline{q}, \sigma)$ to its final (positive-energy) state $(\underline{q}', \sigma')$. The Thomson scattering cross-section of light is given by the expression

$$\left(\frac{d\sigma}{d\Omega_{\underline{k}'}} \right) \; = \; \left(\frac{V \, \omega_{\underline{k}'}}{2 \, \pi \, \hbar \, c^2} \right)^2 \, |\, M \,|^2 \qquad (2070)$$

where the matrix element M are determined from

$$M \;\; = \;\; \sum_{\underline{q}''} \left[\frac{< \, \underline{q}', \underline{k}', \alpha' \mid \hat{H}_{Int} \mid \underline{q}'' \, > < \, \underline{q}'' \mid \hat{H}_{Int} \mid \underline{q}, \underline{k}, \alpha \, >}{(\, E_{\underline{q}} \, + \, \hbar \, \omega_k \, - \, E_{\underline{q}''} \,)} \right.$$

$$\left. + \; \frac{< \, \underline{q}', \underline{k}', \alpha' \mid \hat{H}_{Int} \mid \underline{q}'', \underline{k}, \alpha, \underline{k}', \alpha' \, > < \, \underline{q}'', \underline{k}, \alpha, \underline{k}', \alpha' \mid \hat{H}_{Int} \mid \underline{q}, \underline{k}, \alpha \, >}{(\, E_{\underline{q}} \, - \, E_{\underline{q}''} \, - \, \hbar \, \omega_{\underline{k}'} \,)} \right]$$

$$(2071)$$

and where $\underline{q}$ indicates all the quantum numbers of a positive-energy free electron state. The sum over $\underline{q}''$ represents a sum over all possible intermediate states of the electron, no matter whether they are positive or negative-energy states. The matrix element M is composed of a coherent superposition of matrix elements for virtual processes which represent the absorption of a photon $(\underline{k}, \alpha)$ followed by the subsequent emission of a photon $(\underline{k}', \alpha')$ and the process where the emission of light precedes the absorption process.
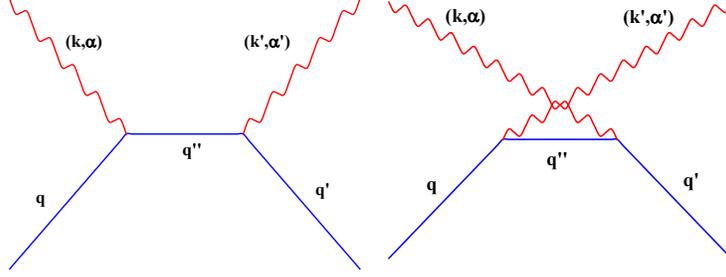
Figure 68: Processes involving negative electron states $\underline{q}''$ which contribute to Compton scattering.

Since the basis set is composed of momentum eigenstates, the evaluation of the spatial integration in the matrix elements of the interaction results in the condition of conservation of momentum. Hence, for the process where the photon $(\underline{k}, \alpha)$ is absorbed before the emission of the photon $(\underline{k}', \alpha')$, the momenta are restricted by

$$
\begin{aligned}
\underline{k} + \underline{q} &= \underline{q}'' \\
\underline{q}'' &= \underline{k}' + \underline{q}'
\end{aligned}
\tag{2072}
$$

which leads to the identification of the momentum of the intermediate and final states as

$$
\begin{aligned}
\underline{q}'' &= \underline{q} + \underline{k} \\
\underline{q}' &= \underline{q} + \underline{k} - \underline{k}'
\end{aligned}
\tag{2073}
$$

In the second process, where the emission process precedes the absorption, conservation of momentum yields

$$
\begin{aligned}
\underline{k} + \underline{q} &= \underline{k} + \underline{k}' + \underline{q}'' \\
\underline{k} + \underline{k}' + \underline{q}'' &= \underline{k}' + \underline{q}'
\end{aligned}
\tag{2074}
$$

which yields

$$
\begin{aligned}
\underline{q}'' &= \underline{q} - \underline{k}' \\
\underline{q}' &= \underline{q} + \underline{k} - \underline{k}'
\end{aligned}
\tag{2075}
$$

The limit in which the initial electron is at rest $\underline{q} = 0$ shall be considered. The momenta of the incident and scattered photon will be assumed sufficiently low so that the momentum of the electron in the intermediate state can be neglected since $\underline{q}'' \approx 0$. That is, the Compton scattering process will be consider in the limit $\overline{k} \to 0$ and $k' \to 0$.

If the initial (positive-energy) electron is stationary and has spin σ, its wave function can be represented by the Dirac spinor

$$\psi_{\sigma,\underline{q}}(\underline{r}) \; = \; \frac{1}{\sqrt{V}} \; \left(\begin{array}{c} \chi_\sigma \\ 0 \end{array} \right) \tag{2076}$$

Because the interaction Hamiltonian has the form of an off-diagonal 2×2 block matrix

$$\hat{H}_{Int} \; = \; - \; \frac{q}{c} \; \left(\begin{array}{cc} 0 & \underline{\sigma} \\ \underline{\sigma} & 0 \end{array} \right) \; \cdot \; \underline{A} \tag{2077}$$

the only non-zero matrix elements are those which connect the upper two-component spinor to the lower two-component spinor of the virtual state. Also, momentum conservation requires that the virtual state also be one of almost zero momentum. Hence, the electron in the virtual state must have the form of a negative-energy eigenstate

$$\psi_{\sigma'',\underline{q}''}(\underline{r}) \; \approx \; \frac{1}{\sqrt{V}} \; \left(\begin{array}{c} 0 \\ \chi_{\sigma''} \end{array} \right) \tag{2078}$$

since the contribution from a positive-energy state with small momentum is negligibly small. Therefore, the electronic part of the matrix elements involving the initial electron simply reduce to the expression

$$< \; \psi_{\sigma'',\underline{q}''} \; | \; \hat{H}_{Int} \; | \; \psi_{\sigma,\underline{q}} \; > \; = \; | \; e \; | \; \chi^\dagger_{\sigma''} \; \underline{\sigma} \; \chi_\sigma \; \cdot \; \underline{A} \tag{2079}$$

Likewise, the matrix elements which involve the final (positive energy) electron are evaluated as

$$< \; \psi_{\sigma',\underline{q}'} \; | \; \hat{H}_{Int} \; | \; \psi_{\sigma'',\underline{q}''} \; > \; = \; | \; e \; | \; \chi^\dagger_{\sigma'} \; \underline{\sigma} \; \chi_{\sigma''} \; \cdot \; \underline{A} \tag{2080}$$

From these one finds that, to second-order, the matrix elements that appear in the transition rate are given by

$$\begin{aligned} M \; &= \; e^2 \; \left(\frac{2 \, \pi \, \hbar \, c^2}{V \, \sqrt{\omega_{\underline{k}} \, \omega_{\underline{k}'}}} \right) \; \sum_{\sigma''} \; \left[\; \frac{(\, \chi^\dagger_{\sigma'} \; \underline{\sigma} \; \cdot \; \epsilon_{\alpha'}(\underline{k}') \; \chi_{\sigma''} \,) \; (\, \chi^\dagger_{\sigma''} \; \underline{\sigma} \; \cdot \; \epsilon_\alpha(\underline{k}) \; \chi_\sigma \,)}{E_{\underline{q}} \; - \; E_{\underline{q}''} \; + \; \hbar \, \omega_{\underline{k}}} \right. \\ &\quad + \; \frac{(\, \chi^\dagger_{\sigma'} \; \underline{\sigma} \; \cdot \; \epsilon_\alpha(\underline{k}) \; \chi_{\sigma''} \,) \; (\, \chi^\dagger_{\sigma''} \; \underline{\sigma} \; \cdot \; \epsilon_{\alpha'}(\underline{k}') \; \chi_\sigma \,)}{E_{\underline{q}} \; - \; E_{\underline{q}''} \; - \; \hbar \, \omega_{\underline{k}'}} \left. \right] \\ &\approx \; \left(\frac{e^2}{2 \, m \, c^2} \right) \; \left(\frac{2 \, \pi \, \hbar \, c^2}{V \, \sqrt{\omega_{\underline{k}} \, \omega_{\underline{k}'}}} \right) \; \sum_{\sigma''} \; \left[\; (\, \chi^\dagger_{\sigma'} \; \underline{\sigma} \; \cdot \; \epsilon_{\alpha'}(\underline{k}') \; \chi_{\sigma''} \,) \; (\, \chi^\dagger_{\sigma''} \; \underline{\sigma} \; \cdot \; \epsilon_\alpha(\underline{k}) \; \chi_\sigma \,) \right. \\ &\quad + \; (\, \chi^\dagger_{\sigma'} \; \underline{\sigma} \; \cdot \; \epsilon_\alpha(\underline{k}) \; \chi_{\sigma''} \,) \; (\, \chi^\dagger_{\sigma''} \; \underline{\sigma} \; \cdot \; \epsilon_{\alpha'}(\underline{k}') \; \chi_\sigma \,) \left. \right] \end{aligned} \tag{2081}$$

where one has set

$$E_{\underline{q}} \; - \; E_{\underline{q}''} \; \approx \; 2 \, m \, c^2 \tag{2082}$$

On using the completeness relation for the two-component Dirac spinors

$$\sum_{\sigma''} \; \chi_{\sigma''} \; \chi^\dagger_{\sigma''} \; = \; I \tag{2083}$$

the matrix elements are evaluated as

$$
M \approx \left(\frac{e^2}{2\,m\,c^2} \right) \left(\frac{2\,\pi\,\hbar\,c^2}{V\,\sqrt{\omega_{\underline{k}}\,\omega_{\underline{k}'}}} \right) \chi_{\sigma'}^{\dagger} \left[\, (\,\underline{\sigma}\cdot\hat{\epsilon}_\alpha(\underline{k})\,)\,(\,\underline{\sigma}\cdot\hat{\epsilon}_{\alpha'}(\underline{k}')\,) + (\,\underline{\sigma}\cdot\hat{\epsilon}_{\alpha'}(\underline{k}')\,)\,(\,\underline{\sigma}\cdot\hat{\epsilon}_\alpha(\underline{k})\,) \, \right] \chi_\sigma
$$

(2084)

The products in the above expression can be evaluated with the aid of the Pauli identity. The result is

$$
(\,\underline{\sigma}\cdot\hat{\epsilon}_\alpha(\underline{k})\,)\,(\,\underline{\sigma}\cdot\hat{\epsilon}_{\alpha'}(\underline{k}')\,) = (\,\hat{\epsilon}_\alpha(\underline{k})\cdot\hat{\epsilon}_{\alpha'}(\underline{k}')\,) + i\,\underline{\sigma}\cdot(\,\hat{\epsilon}_\alpha(\underline{k})\wedge\hat{\epsilon}_{\alpha'}(\underline{k}')\,) \quad (2085)
$$

Therefore, after combining both terms and noting that the pair of vector product terms cancel since the vector product is antisymmetric under the interchange $\underline{k} \leftrightarrow \underline{k}'$, one finds that the matrix elements reduce to

$$
\begin{aligned}
M &\approx \left(\frac{e^2}{2\,m\,c^2} \right) \left(\frac{2\,\pi\,\hbar\,c^2}{V\,\sqrt{\omega_{\underline{k}}\,\omega_{\underline{k}'}}} \right) \chi_{\sigma'}^{\dagger} \left[\, 2\,\hat{\epsilon}_\alpha(\underline{k})\cdot\hat{\epsilon}_{\alpha'}(\underline{k}') \, \right] \chi_\sigma \\
M &\approx \left(\frac{e^2}{2\,m\,c^2} \right) \left(\frac{2\,\pi\,\hbar\,c^2}{V\,\sqrt{\omega_{\underline{k}}\,\omega_{\underline{k}'}}} \right) \delta_{\sigma,\sigma'}\,2\,\hat{\epsilon}_\alpha(\underline{k})\cdot\hat{\epsilon}_{\alpha'}(\underline{k}')
\end{aligned}
$$

(2086)

Hence the matrix elements are diagonal in the spin indices. The above matrix elements are identical to the matrix elements that occur in the non-relativistic quantum theory of Thomson scattering. On substituting this result into eqn(2070), one recovers the non-relativistic expression for the differential scattering cross-section

$$
\begin{aligned}
\left(\frac{d\sigma}{d\Omega_{\underline{k}'}} \right) &\approx \delta_{\sigma,\sigma'} \left(\frac{\omega_{\underline{k}'}}{\omega_{\underline{k}}} \right) \left(\frac{e^2}{m\,c^2} \right)^2 |\,\hat{\epsilon}_\alpha(\underline{k})\cdot\hat{\epsilon}_{\alpha'}(\underline{k}')\,|^2 \\
&\approx \delta_{\sigma,\sigma'} \left(\frac{\omega_{\underline{k}'}}{\omega_{\underline{k}}} \right) \left(\frac{e^2}{m\,c^2} \right)^2 \cos^2\Theta
\end{aligned}
$$

(2087)

where

$$
\cos\Theta = \hat{\epsilon}_\alpha(\underline{k})\cdot\hat{\epsilon}_{\alpha'}(\underline{k}') \quad (2088)
$$

is the angle subtended by the initial and final polarization vectors. Hence, one concludes that the negative-energy states do play an important role in light scattering processes which involve low-energy electrons. The result, although correct, does need re-interpretation, since the states of negative energy are assumed to be filled with electrons in the vacuum and, therefore, the electron is forbidden to occupy these levels in the intermediate states.

Electron-Positron Interpretation

The first contribution to the matrix elements, which was described above, has to be re-interpreted as representing a process in which an electron that initially occupies the negative-energy state $\underline{q}''$ makes a transition to the positive-energy state $\underline{q}'$ while emitting the photon $(\underline{k}',\alpha')$. This transition is subsequently followed by the positive-energy electron $\underline{q}$ absorbing the photon $(\underline{k},\alpha)$ and falling
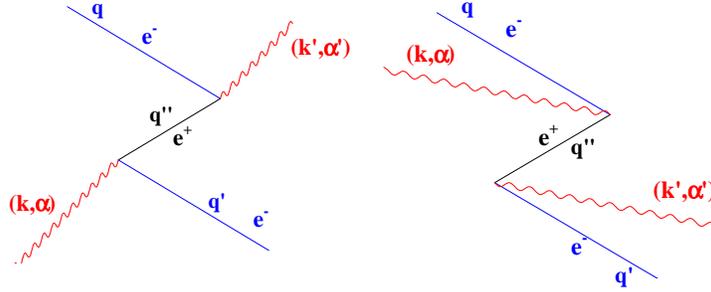
Figure 69: Processes involving positrons which contribute to Compton scattering.

into the empty negative-energy state. In this process, the negative-energy states are completely occupied in the initial and final state, and the energy of the initial and final states are conserved. By re-ordering the factors in the matrix elements and noting that since

$$E_{\underline{q}} \; + \; \hbar \, \omega_{\underline{k}} \; = \; E_{\underline{q}'} \; + \; \hbar \, \omega_{\underline{k}'} \tag{2089}$$

the contribution to the matrix element of these two descriptions are identical (apart from an over all negative sign).

The second contribution to the matrix elements can be viewed as originating from an electron which initially occupies a negative-energy state $\underline{q}''$ that absorbs the photon $(\underline{k}, \alpha)$ and makes a transition to the positive-energy state $\underline{q}'$. This is followed by the electron in the positive-energy state $\underline{q}$ emitting the photon $(\underline{k}', \alpha')$ and then falling into the empty negative-energy state $\underline{q}''$. Again, on re-ordering the matrix elements and noting that

$$E_{\underline{q}} \; - \; \hbar \, \omega_{\underline{k}'} \; = \; E_{\underline{q}'} \; - \; \hbar \, \omega_{\underline{k}} \tag{2090}$$

one finds an identical expression (and the multiplicative factor of minus one). Hence, Dirac hole-theory does lead to the correct classical result.

The above description is quite cumbersome, but can be made more concise by adopting an anti-particle description of the unoccupied negative-energy states. The first contribution to M first involves the creation of a virtual electron-positron pair with the emission of the photon $(\underline{k}', \alpha')$. The electron which has just been created in the momentum eigenstate $(\underline{q}', \sigma')$ remains unchanged in the final state. Subsequently, the positron annihilates with the initial electron $(\underline{q}, \sigma)$ while absorbing the photon $(\underline{k}, \alpha)$. Since the intermediate state is a virtual state, energy does not have to be conserved. The second contribution to M involves the creation of a virtual electron-positron pair with the absorption of the photon $(\underline{k}, \alpha)$. The created electron $(\underline{q}', \sigma')$ remains in the final state while

the positron subsequently annihilates with the initial electron $(\underline{q}, \sigma)$ and emits the photon $(\underline{k}', \alpha')$. This process is also a virtual process if the energy of the incident light $\hbar\,\omega_k$ is less than $2\,m\,c^2$.

The perturbative expression for the Compton scattering cross-section can be evaluated exactly, without recourse to non-relativistic approximations. The exact result is

$$\left(\frac{d\sigma}{d\Omega} \right) \;=\; \frac{1}{4}\,r_e^2\,\left(\frac{\omega'}{\omega} \right)^2\,\left(\frac{\omega}{\omega'} \;+\; \frac{\omega'}{\omega} \;-\; 2 \;+\; 4\cos^2\Theta \right) \tag{2091}$$

where Θ is the angle between the polarization vectors. This result was first derived by Klein and Nishina[148] in 1928.

12.16.2   Charge Conjugation

Charge conjugation is the operation of replacing matter by anti-matter, so that, for example, electrons will be replaced by positrons and vice versa. The operation of charge conjugation consists of first taking the complex conjugate of the Dirac equation

$$\left[\gamma^\mu\,\left(i\,\hbar\,\partial_\mu \;-\; \frac{q}{c}\,A_\mu \right) \;-\; m\,c \right] \psi \;=\; 0 \tag{2092}$$

which describes a particle with charge q. We shall also assume that ψ describes a positive-energy solution. Complex conjugation yields the equation

$$\left[\gamma^{\mu*}\,\left(-\,i\,\hbar\,\partial_\mu \;-\; \frac{q}{c}\,A_\mu^* \right) \;-\; m\,c \right] \psi^* \;=\; 0 \tag{2093}$$

The complex conjugate of a positive-energy solution ψ^* has a time-dependent phase that identifies it with a negative-energy solution. The vector potential A_μ is real. In the standard representation $\gamma^{(0)}$, $\gamma^{(1)}$ and $\gamma^{(3)}$ are real, whereas $\gamma^{(2)}$ is imaginary and, therefore, satisfies

$$\gamma^{(2)*} \;=\; -\,\gamma^{(2)} \tag{2094}$$

We shall multiply the complex conjugate of the Dirac equation by $\gamma^{(2)}$ and anti-commute $\gamma^{(2)}$ with the real $\gamma^{\mu*}$ and commute $\gamma^{(2)}$ with the $\gamma^{(2)*}$ matrix. This procedure changes the sign in front of the term originating from the differential momentum operator w.r.t. the sign of the mass term. This procedure yields

$$\gamma^{(2)}\left[\gamma^{\mu*}\,\left(-\,i\,\hbar\,\partial_\mu \;-\; \frac{q}{c}\,A_\mu \right) \;-\; m\,c \right] \psi^* \;=\; 0$$

$$\left[\gamma^\mu\,\left(i\,\hbar\,\partial_\mu \;+\; \frac{q}{c}\,A_\mu \right) \;-\; m\,c \right] \gamma^{(2)}\,\psi^* \;=\; 0 \tag{2095}$$

---

[148] O. Klein and Y. Nishina, Zeit. für Physik, **52**, 843 (1928).

Hence, one sees that $\gamma^{(2)}\psi^*$ describes a Dirac particle with mass m and a charge of $-q$ moving in the presence of a vector potential A^μ. The fact that the operation of charge conjugation (in any representation) involves complex conjugation is related to gauge invariance. Charge conjugation is a new type of symmetry for particles that have complex wave functions which relates particles to particles with opposite charges. The charge conjugate field ψ^c is defined as

$$\psi^c \;=\; \hat{C}\,\psi^* \tag{2096}$$

which is the result of the complex conjugation followed by the action of a linear operator $\hat{C}$. The joint operation can be represented as an anti-unitary operator. The charge conjugation operator $\hat{C}$ is defined as the unitary and Hermitean operator

$$\hat{C} \;=\; -\,i\,\gamma^{(2)} \tag{2097}$$

The charge conjugation operator is Hermitean as

$$\hat{C}^\dagger \;=\; +\,i\,\gamma^{(2)\,\dagger} \;=\; -\,i\,\gamma^{(2)} \;=\; \hat{C} \tag{2098}$$

and it is unitary since

$$\hat{C}^\dagger\,\hat{C} \;=\; -\,\gamma^{(2)}\,\gamma^{(2)} \;=\; \hat{I} \tag{2099}$$

where the anti-commutation relations of the γ matrices have been used. It was through this type of logic that Kramers[149] discovered the form of the charge conjugation transformation which turns a particle into an anti-particle.

The expectation values of an operator $\hat{A}$ in a general charge conjugated state ψ^c are related to the expectation values in a general state ψ via

$$<\psi^c\mid\hat{A}\mid\psi^c> \;=\; -\left(\;<\psi\mid\gamma^{(2)}\,\hat{A}^*\,\gamma^{(2)}\mid\psi>\;\right)^* \tag{2100}$$

This can be shown in the position representation, by writing

$$\int d^3\underline{r}\,\psi^{c\dagger}(\underline{r})\,\hat{A}\,\psi^c(\underline{r}) \;=\; \int d^3\underline{r}\,\psi^{*\dagger}(\underline{r})\,\hat{C}^\dagger\,\hat{A}\,\hat{C}\,\psi^*(\underline{r})$$

$$=\;\left(\int d^3\underline{r}\,\psi^\dagger(\underline{r})\,\hat{C}^{\dagger*}\,\hat{A}^*\,\hat{C}^*\,\psi(\underline{r})\right)^* \tag{2101}$$

where we have used the identity $z = (z^*)^*$ in the second line. However, since $\hat{C}$ is real, one finds

$$\int d^3\underline{r}\,\psi^{c\dagger}(\underline{r})\,\hat{A}\,\psi^c(\underline{r}) \;=\; \left(\int d^3\underline{r}\,\psi^\dagger(\underline{r})\,\hat{C}\,\hat{A}^*\,\hat{C}\,\psi(\underline{r})\right)^*$$

$$=\;-\left(\int d^3\underline{r}\,\psi^\dagger(\underline{r})\,\gamma^{(2)}\,\hat{A}^*\,\gamma^{(2)}\,\psi(\underline{r})\right)^* \tag{2102}$$

[149]H. A. Kramers, Proc. Amst. Akad. Sci. **40**, 814 (1937).

This shows the relation between expectation values of a general operator $\hat{A}$ in a state $\psi(\underline{r})$ and its charge conjugated state $\psi^c(\underline{r})$.

We shall examine the effect of charge conjugation on the plane wave solutions of the Dirac equation. The plane-wave solutions can be written as

$$\psi_{\sigma,\underline{k}}(x) = \sqrt{\frac{(E + m c^2)}{2 E V}} \left(\begin{array}{c} \chi_\sigma \\ \frac{c \hbar \underline{\sigma} \cdot \underline{k}}{E + m c^2} \chi_\sigma \end{array} \right) \exp\left[- i k^\mu x_\mu \right] \quad (2103)$$

The charge conjugate wave function is given by

$$\begin{aligned} \psi^c_{\sigma,\underline{k}}(x) &= \hat{C} \, \psi^*_{\sigma,\underline{k}}(x) \\ &= \sqrt{\frac{(E + m c^2)}{2 E V}} \, \hat{C} \left(\begin{array}{c} \chi^*_\sigma \\ \frac{c \hbar \underline{\sigma}^* \cdot \underline{k}}{E + m c^2} \chi^*_\sigma \end{array} \right) \exp\left[+ i k^\mu x_\mu \right] \end{aligned}$$
$$(2104)$$

where

$$\begin{aligned} \hat{C} &= - i \gamma^{(2)} \\ &= \left(\begin{array}{cc} 0 & -i\sigma^{(2)} \\ i\sigma^{(2)} & 0 \end{array} \right) \\ &= \left(\begin{array}{cccc} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array} \right) \end{aligned} \quad (2105)$$

Therefore, the charge conjugate wave function is found to be given by

$$\psi^c_{\sigma,\underline{k}}(x) = i \, \hat{\sigma}^{(2)} \sqrt{\frac{(E + m c^2)}{2 E V}} \left(\begin{array}{c} - \frac{c \hbar \underline{\sigma}^* \cdot \underline{k}}{E + m c^2} \chi^*_\sigma \\ \chi^*_\sigma \end{array} \right) \exp\left[+ i k^\mu x_\mu \right]$$
$$(2106)$$

which has the form of a plane-wave solution with negative energy $E \rightarrow - E$, and momentum $\hbar \underline{k} \rightarrow - \hbar \underline{k}$. Furthermore, the spin of the charge conjugated wave function has been reversed[150] $\underline{\sigma} \rightarrow - \underline{\sigma}$, since when $i \sigma^{(2)}$ acts on the complex conjugated positive-eigenvalue eigenstate of the spin projected on an arbitrary direction

$$\chi_{+\sigma}(\theta, \varphi)^* = \left(\begin{array}{c} \cos\frac{\theta}{2} \exp[+i\frac{\varphi}{2}] \\ \sin\frac{\theta}{2} \exp[-i\frac{\varphi}{2}] \end{array} \right) \quad (2107)$$

---

[150]Note that the helicity is invariant under the joint transformation

$$\underline{\sigma} \rightarrow - \underline{\sigma}$$
$$\underline{k} \rightarrow - \underline{k}$$

it turns it into the negative-eigenvalue eigenstate

$$\chi_{-\sigma}(\theta,\varphi) \;=\; \left(\begin{array}{c} -\sin\frac{\theta}{2}\,\exp[-i\frac{\varphi}{2}] \\ \cos\frac{\theta}{2}\,\exp[+i\frac{\varphi}{2}] \end{array} \right) \tag{2108}$$

That is, up to an arbitrary phase factor, the lower two-component spinor is given by

$$i\,\sigma^{(2)}\,\chi_{+\sigma}(\theta,\varphi)^* \;=\; \chi_{-\sigma}(\theta,\varphi) \tag{2109}$$

Likewise, it can be shown that the upper two-component spinor is proportional to

$$\begin{aligned} i\,\sigma^{(2)}\,(\,\underline{\sigma}^*\,.\,\underline{k}\,)\,\chi_{+\sigma}(\theta,\varphi)^* &= -(\,\underline{\sigma}\,.\,\underline{k}\,)\,(\,i\,\sigma^{(2)}\,)\,\chi_{+\sigma}(\theta,\varphi)^* \\ &= -(\,\underline{\sigma}\,.\,\underline{k}\,)\,\chi_{-\sigma}(\theta,\varphi) \\ &= (\,\underline{\sigma}\,.\,(\,-\,\underline{k}\,)\,)\,\chi_{-\sigma}(\theta,\varphi) \end{aligned} \tag{2110}$$

The end result is that the charge conjugated single-particle wave function has the form

$$\psi^c_{\sigma,\underline{k}}(x) \;=\; \sqrt{\frac{(\,E\,+\,m\,c^2\,)}{2\,E\,V}}\,\left(\begin{array}{c} -\,\frac{c\,\hbar\,(\,\underline{\sigma}\,.\,(-\underline{k})\,)}{E\,+\,m\,c^2}\,\chi_{-\sigma} \\ \chi_{-\sigma} \end{array} \right)\,\exp\left[\,+\,i\,k^\mu\,x_\mu\,\right] \tag{2111}$$

The properties described above are the properties of a state of a relativistic free particle with a negative energy eigenvalue $-\,E$, momentum $-\,\hbar\,\underline{k}$ and spin $-\,\underline{\sigma}$. The absence of an electron in the charge conjugated state describes a positron, with positive energy E, momentum $\hbar\,\underline{k}$ and spin σ.

More generally, even when an electromagnetic field is present, the charge conjugated wave function of a positive-energy particle corresponds to the wave function of a state with reversed energy $E\,\rightarrow\,-\,E$, reversed spin $\underline{\sigma}\,\rightarrow\,-\,\underline{\sigma}$ and reversed charge $q\,\rightarrow\,-\,q$. Therefore, the charge conjugated state corresponds to the (negative-energy) state which when unoccupied is described as an anti-particle.

Exercise:

Consider massless Dirac particles, $m\rightarrow 0$. (i) Show that the energy-helicity eigenstates coincide with the eigenstates of $\gamma^{(4)}$. (ii) Hence, show that the operators $\frac{1}{2}\,(\,\hat{I}\,\pm\,\gamma^{(4)}\,)$ project onto helicity eigenstates. These projection operators relate the four-component Dirac spinors onto the independent two-component Weyl spinors ϕ^L and ϕ^R. (iii) Show that charge conjugation transforms ϕ^L into ϕ^R.

Exercise:

Prove the completeness relation for the set of solutions for the Dirac equation for a free particle

$$\sum_\alpha \left(\phi_\alpha^\dagger(\underline{r})_\lambda \, \phi_\alpha(\underline{r}')_\rho \; + \; \phi_\alpha^{c\,\dagger}(\underline{r})_\lambda \, \phi_\alpha^c(\underline{r}')_\rho \right) \;\; = \;\; \delta^3(\underline{r} - \underline{r}') \, \delta_{\lambda,\rho} \quad (2112)$$

where λ and ρ denote the components of the Dirac spinor[151].

13   Local and Global Gauge Symmetries

Global transformations are defined to be transformations which are the same at every point in space-time, whereas local transformations vary from point to point. Global symmetries, such as invariance under rotation, spatial translation, and time translations, lead to laws of physics such as conservation of angular momentum, linear momentum and energy. Local transformations do not lead to new physical laws, but do lead to constraints on the form of the interactions.

A familiar example of a local transformation is the gauge transformation in Quantum Electrodynamics. The Lagrangian density for Quantum Electrodynamics $\mathcal{L}_{QED}$ given by the sum of the field-free electrodynamic Lagrangian density and the Dirac Lagrangian density

$$\mathcal{L}_{QED} \;\; = \;\; - \, \frac{1}{16\,\pi} \, F^{\mu,\nu} \, F_{\mu,\nu} \; + \; c \, \overline{\psi}^\dagger \left(\gamma^\mu \, (\, \hat{p}_\mu \; - \; \frac{q}{c} \, A_\mu \,) - \; m \, c \right) \psi \quad (2113)$$

The role of the source term for the electromagnetic field is played by the term in the Dirac Lagrangian which represents the coupling to the vector potential. The expression for the QED action

$$S_{QED} \;\; = \;\; \int dx^4 \, \mathcal{L}_{QED} \qquad (2114)$$

is invariant under an infinite number of infinitesimal gauge transformations with the form

$$\begin{aligned} A_\mu \;\; &\rightarrow \;\; A_\mu' \; = \; A_\mu \; + \; \partial_\mu \, \Lambda \\ \psi \;\; &\rightarrow \;\; \psi' \; = \; \psi \; - \; \left(\frac{i\,q}{\hbar\,c} \right) \Lambda \, \psi \\ \psi^\dagger \;\; &\rightarrow \;\; \psi^{\dagger\prime} \; = \; \psi^\dagger \; + \; \left(\frac{i\,q}{\hbar\,c} \right) \Lambda \, \psi^\dagger \end{aligned} \qquad (2115)$$

---

[151] Frequently, the relativistic free electron states are given a manifestly covariant normalization, in order to facilitate covariant perturbation theory. The use of different normalization conventions results in changes the form of the completeness relation.

Applying Noether's theorem to the action yields an infinite number of continuous currents, which, up to a multiplicative factor, are given by

$$j_\Lambda^\nu = \left(\frac{\partial \mathcal{L}_{QED}}{\partial \left(\partial_\nu A_\mu \right)} \right) \partial_\mu \Lambda - \left(\frac{\partial \mathcal{L}_{QED}}{\partial \left(\partial_\nu \psi \right)} \right) \left(\frac{i\,q}{\hbar\,c} \right) \Lambda\,\psi \tag{2116}$$

which satisfy the continuity conditions

$$\partial_\nu\,j_\Lambda^\nu = 0 \tag{2117}$$

The currents are identified as

$$j_\Lambda^\nu = \frac{1}{4\,\pi}\,F^{\mu,\nu}\,\partial_\mu\,\Lambda + q\,\overline{\psi}^\dagger\,\gamma^\nu\,\Lambda\,\psi \tag{2118}$$

On substituting the Euler-Lagrange equation for the electromagnetic field

$$\partial_\mu\,F^{\mu,\nu} = \frac{4\,\pi}{c}\,q\,c\,\overline{\psi}^\dagger\,\gamma^\nu\,\psi \tag{2119}$$

into the second term of the four-vector current density, one finds that the continuous current has the form

$$j_\Lambda^\nu = \frac{1}{4\,\pi}\,\partial_\mu\left(F^{\mu,\nu}\,\Lambda \right) \tag{2120}$$

The conserved charge Q can be defined as a volume integral of the temporal component of the four-vector current density

$$
\begin{aligned}
Q_\Lambda &= \frac{1}{c}\int d^3\underline{r}\;j_\Lambda^{(0)}(\underline{r}) \\
&= \frac{1}{4\,\pi\,c}\int d^3\underline{r}\;\partial_\mu\left(F^{\mu,0}\,\Lambda \right) \\
&= \frac{1}{4\,\pi\,c}\int d^3\underline{r}\;\underline{\nabla}\,.\,(\,\underline{E}\,\Lambda\,)
\end{aligned}
\tag{2121}
$$

For a global transformation for which Λ is constant over all space-time, one can write

$$Q_{\Lambda=const.} = \frac{\Lambda}{4\,\pi\,c}\int d^3\underline{r}\,(\,\underline{\nabla}\,.\,\underline{E}\,) \tag{2122}$$

On using Gauss's law

$$\underline{\nabla}\,.\,\underline{E} = 4\,\pi\,\rho \tag{2123}$$

one finds that global gauge invariance ensures conservation of electrical charge

$$Q_{\Lambda=const.} = \frac{\Lambda}{c}\int d^3\underline{r}\;\rho(\underline{r}) \tag{2124}$$

However, for any of the infinite number of spatially varying Λ which vanish on the boundaries

$$
\begin{aligned}
Q_\Lambda &= \frac{1}{4\,\pi\,c}\int d^3\underline{r}\;\underline{\nabla}\,.\,(\,\underline{E}\,\Lambda\,) \\
&= \frac{1}{4\,\pi\,c}\int d^2\underline{S}\,.\,\underline{E}\,\Lambda \\
&= 0
\end{aligned}
\tag{2125}
$$

Hence, although local transformations lead to an infinite number of symmetries and an infinite number of continuous currents, all the conserved charges are zero. Thus, local gauge invariance does not leads to new conserved quantities. It does, however, constrain the form of the interactions.

14 The Many-Particle Dirac Field

14.1 The Algebra of Fermion Operators

Second quantization of fermions amounts to adopting an occupation number representation. Therefore, we shall examine No. accounting for fermions[152].

Fermion operators satisfy anti-commutation relations. The anti-commutator of two operators $\hat{A}$ and $\hat{B}$ is defined as

$$\{\ \hat{A}\ ,\ \hat{B}\ \}_+ \ \equiv\ \hat{A}\ \hat{B}\ +\ \hat{B}\ \hat{A} \tag{2126}$$

The fermion creation and annihilation operators, $\hat{c}_\alpha^\dagger$ and $\hat{c}_\alpha$, satisfy the anti-commutation relations

$$\begin{aligned}
\{\ \hat{c}_\alpha^\dagger\ ,\ \hat{c}_\beta^\dagger\ \}_+ \ &=\ 0 \\
\{\ \hat{c}_\alpha\ ,\ \hat{c}_\beta\ \}_+ \ &=\ 0
\end{aligned} \tag{2127}$$

and

$$\{\ \hat{c}_\alpha^\dagger\ ,\ \hat{c}_\beta\ \}_+ \ =\ \delta_{\alpha,\beta} \tag{2128}$$

where the quantum numbers α and β describe a complete set of single-particle states.

The anti-commutation relation

$$\hat{c}_\alpha^\dagger\ \hat{c}_\beta^\dagger \ =\ -\ \hat{c}_\beta^\dagger\ \hat{c}_\alpha^\dagger \tag{2129}$$

is merely a re-statement of the anti-symmetric nature of a fermionic many-particle wave function under the permutation of a pair of particles, as is the Hermitean conjugate relation

$$\hat{c}_\alpha\ \hat{c}_\beta \ =\ -\ \hat{c}_\beta\ \hat{c}_\alpha \tag{2130}$$

The number operator $\hat{n}_\alpha$ is defined as

$$\hat{n}_\alpha \ =\ \hat{c}_\alpha^\dagger\ \hat{c}_\alpha \tag{2131}$$

---

[152]P. Jordan and E. Wigner, Zeit. für Physik, **47**, 631 (1928).

The choice of anti-commutation relations results in the eigenvalues of the number operator to be restricted to either $n_\alpha = 1$ or $n_\alpha = 0$. This can be seen by examining the identity

$$\hat{n}_\alpha \, \hat{n}_\alpha \;=\; \hat{n}_\alpha \tag{2132}$$

which follows from

$$
\begin{aligned}
\hat{n}_\alpha \, \hat{n}_\alpha \;&=\; \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \, \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \\
&=\; \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \;-\; \hat{c}_\alpha^\dagger \, \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \, \hat{c}_\alpha \\
&=\; \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \;+\; \hat{c}_\alpha^\dagger \, \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \, \hat{c}_\alpha
\end{aligned}
\tag{2133}
$$

where we have used the anti-commutation relation for the creation and annihilation operator to obtain the second line and used the anti-commutation relation for two annihilation operators to obtain the last line. On comparing the second and third lines, one recognizes that

$$\hat{c}_\alpha^\dagger \, \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \, \hat{c}_\alpha \;=\; 0 \tag{2134}$$

Hence, we have

$$
\begin{aligned}
\hat{n}_\alpha \, \hat{n}_\alpha \;&=\; \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \\
&=\; \hat{n}_\alpha
\end{aligned}
\tag{2135}
$$

Thus, the eigenstates of the number operator satisfy the equation

$$
\begin{aligned}
\hat{n}_\alpha \, \hat{n}_\alpha \mid n_\alpha > \;&=\; \hat{n}_\alpha \mid n_\alpha > \\
n_\alpha^2 \mid n_\alpha > \;&=\; n_\alpha \mid n_\alpha >
\end{aligned}
\tag{2136}
$$

Therefore, for there to be non-trivial eigenstates the eigenvalues must satisfy the equation

$$n_\alpha \, (\, n_\alpha \, - \, 1 \,) \;=\; 0 \tag{2137}$$

which only has the solutions $n_\alpha = 0$ and $n_\alpha = 1$. Thus the choice of anti-commutation relations for the creation and annihilation operators results in the Pauli exclusion principle. The Pauli exclusion principle states that a non-degenerate quantum state can not be occupied by more than one fermion.

The number operator satisfies the commutation relations

$$
\begin{aligned}
[\, \hat{n}_\alpha \, , \, \hat{c}_\beta^\dagger \,] \;&=\; \delta_{\alpha,\beta} \, \hat{c}_\beta^\dagger \\
[\, \hat{n}_\alpha \, , \, \hat{c}_\beta \,] \;&=\; - \, \delta_{\alpha,\beta} \, \hat{c}_\beta
\end{aligned}
\tag{2138}
$$

as can be seen by using the fermion anti-commutation relations. The hierarchy of eigenstates of the number operator can be found from the action of the creation operator. In particular, one can define the eigenstate of the annihilation operator with eigenvalue zero by

$$\hat{c}_\alpha \mid 0 > \; = \; 0 \tag{2139}$$

The state $| \ 0 \ >$ is also an eigenstate of the number operator with eigenvalue zero since

$$
\begin{aligned}
\hat{n}_\alpha \ | \ 0 \ > \ &= \ \hat{c}_\alpha^\dagger \ \hat{c}_\alpha \ | \ 0 \ > \\
&= \ 0
\end{aligned}
\tag{2140}
$$

A general eigenstate of the number operator with eigenvalue n_α can be expressed as

$$
| \ n_\alpha \ > \ = \ \frac{(\ \hat{c}_\alpha^\dagger \ )^{n_\alpha}}{\sqrt{n_\alpha!}} \ | \ 0 \ >
\tag{2141}
$$

as can be seen by using the commutation relation

$$
[\ \hat{n}_\alpha \ , \ (\ \hat{c}_\alpha^\dagger \ )^{n_\alpha} \ ] \ = \ n_\alpha \ (\ \hat{c}_\alpha^\dagger \ )^{n_\alpha}
\tag{2142}
$$

If this operator equation acts on the state where the quantum state α is unoccupied $| \ 0 \ >$, and using the condition

$$
\hat{n}_\alpha \ | \ 0 \ > \ = \ 0
\tag{2143}
$$

one finds the state of equation(2140) satisfies the eigenvalue equation

$$
\hat{n}_\alpha \ | \ n_\alpha \ > \ = \ n_\alpha \ | \ n_\alpha \ >
\tag{2144}
$$

with eigenvalue of either unity or zero.

A general number operator eigenstate can be expressed in terms of the occupation numbers of all the single-particle states

$$
| \ \{n_\alpha\} \ > \ = \ \prod_{\alpha=1}^{\infty} \ \left\{ \frac{(\ \hat{c}_\alpha^\dagger \ )^{n_\alpha}}{\sqrt{n_\alpha!}} \right\} \ | \ 0 \ , \ 0 \ , \ 0 \ , \ \dots \ , \ 0 \ >
\tag{2145}
$$

where the allowed values of the set of occupation numbers n_α are either unity or zero. The sequencing or ordering of the creation operators in this expression is crucial, since the interchange the positions of the operators may result in a change in sign of the state. For example, the action of a creation operators on a general number eigenstate has the effect

$$
\hat{c}_\beta^\dagger \ | \ n_1 \ n_2 \ \dots \ n_\beta \ \dots \ > \ = \ (\ - \ 1 \ )^{(\sum_{i=1}^{\beta} \ n_i)} \ | \ n_1 \ n_2 \ \dots \ n_\beta + 1 \ \dots \ >
\tag{2146}
$$

where the sign occurs since this involves anti-commutating $\hat{c}_\beta^\dagger$ with $\sum_{i=1}^{\beta} \ n_i$ other creation operators to bring it into the β-th position.

14.2   Quantizing the Dirac Field

The quantization of the Dirac field proceeds exactly the same way as for non-relativistic electrons[153]. However, the negative-energy states will be described

---

[153]W. Heisenberg and W. Pauli, Zeit. für Physik, **56**, 1 (1929).
W. Heisenberg and W. Pauli, Zeit. für Physik, **59**, 168 (1930).

with a different notation from the positive-energy states. The change of notation is to reflect the intent of describing the (quasi-particle) excitations of the system and not to describe the many-particle ground state which is unobservable. The wave functions $\phi_\alpha(\underline{r})$ describing the positive-energy states of the non-interacting electrons are indexed by the set of quantum numbers $\alpha \equiv (\underline{k}, \sigma)$. The negative-energy states are described as the charge conjugates of the positive-energy states. Therefore, the negative-energy states are described by the same set of indices α and the corresponding wave functions are denoted by $\phi_\alpha^c(\underline{r})$. The annihilation operator for electrons in the positive-energy state α is denoted by $\hat{c}_\alpha$. However, the operator which removes an electron from the (negative-energy) charge conjugated state $\phi_\alpha^c(\underline{r})$ is denoted by a creation operator $\hat{b}_\alpha^\dagger$. The change from annihilation operator to creation operator merely represents that creating a positron with quantum numbers α is equivalent to creating a hole in the negative-energy state[154]. The effect of the annihilation operators on Dirac's vacuum $\mid 0 >$, in which all the negative-energy states are fully occupied are

$$
\begin{aligned}
\hat{c}_\alpha \mid 0 > &= 0 \\
\hat{b}_\alpha \mid 0 > &= 0
\end{aligned}
\tag{2147}
$$

where the first expression follows from the assumed absence of electrons in the positive-energy states, and the second expression follows from the assumption that all the negative-energy states are completely filled, so adding an extra electron to the state ϕ_α^c is forbidden by the Pauli-exclusion principle. More concisely, the above relations state that the vacuum contains neither (positive-energy) electrons nor positrons. It is seen that the form of the anti-commutation relations are unchanged by this simple change of notation. The anti-commutation relations become

$$
\begin{aligned}
\{\, \hat{c}_\alpha^\dagger \,,\, \hat{c}_\beta^\dagger \,\}_+ &= \{\, \hat{c}_\alpha \,,\, \hat{c}_\beta \,\}_+ = 0 \\
\{\, \hat{c}_\alpha^\dagger \,,\, \hat{c}_\beta \,\}_+ &= \delta_{\alpha,\beta}
\end{aligned}
\tag{2148}
$$

for the electron operators

$$
\begin{aligned}
\{\, \hat{b}_\alpha^\dagger \,,\, \hat{b}_\beta^\dagger \,\}_+ &= \{\, \hat{b}_\alpha \,,\, \hat{b}_\beta \,\}_+ = 0 \\
\{\, \hat{b}_\alpha^\dagger \,,\, \hat{b}_\beta \,\}_+ &= \delta_{\alpha,\beta}
\end{aligned}
\tag{2149}
$$

for the positron operators, and the mixed electron/positron anti-commutation relations are given by

$$
\{\, \hat{c}_\alpha^\dagger \,,\, \hat{b}_\beta^\dagger \,\}_+ = \{\, \hat{c}_\alpha \,,\, \hat{b}_\beta \,\}_+ = \{\, \hat{c}_\alpha^\dagger \,,\, \hat{b}_\beta \,\}_+ = 0
\tag{2150}
$$

The mixed electron/positron anti-commutation relations are all zero, since the operators describe electrons in different single-particle energy eigenstates. In

[154]W. H. Furry and J. R. Oppenheimer, Phys. Rev. **45**, 245 (1934).

this notation, the field operators are expressed as[155]

$$\hat{\psi}(\underline{r}) \ = \ \sum_{\alpha} \ \left(\ \phi_{\alpha}(\underline{r}) \ \hat{c}_{\alpha} \ + \ \phi_{\alpha}^{c}(\underline{r}) \ \hat{b}_{\alpha}^{\dagger} \ \right) \tag{2151}$$

and

$$\hat{\psi}^{\dagger}(\underline{r}) \ = \ \sum_{\alpha} \ \left(\ \phi_{\alpha}^{*}(\underline{r}) \ \hat{c}_{\alpha}^{\dagger} \ + \ \phi_{\alpha}^{c\,*}(\underline{r}) \ \hat{b}_{\alpha} \ \right) \tag{2152}$$

The field operators $\hat{\psi}(\underline{r})$ and $\hat{\psi}^{\dagger}(\underline{r})$ are expected to be canonically conjugate, as we shall show below.

The Lagrangian density is given by

$$\mathcal{L} \ = \ c \, \hat{\bar{\psi}}^{\dagger} \ \left(\ i \, \hbar \, \gamma^{\mu} \, \partial_{\mu} \ - \ m \, c \ \right) \, \hat{\psi} \tag{2153}$$

so the momentum field operator $\hat{\Pi}(\underline{r})$ canonically conjugate to $\hat{\psi}(\underline{r})$ is given by

$$\hat{\Pi}(\underline{r}) \ = \ \frac{1}{c} \, \frac{\delta \mathcal{L}}{\delta(\partial_0 \hat{\psi})} \ = \ i \, \hbar \, \hat{\bar{\psi}}^{\dagger}(\underline{r}) \, \gamma^{(0)} \ = \ i \, \hbar \, \hat{\psi}^{\dagger}(\underline{r}) \tag{2154}$$

Hence, one expects that the field operators $\hat{\psi}^{\dagger}(\underline{r})$ and $\hat{\psi}(\underline{r})$ are canonically conjugate and, therefore, satisfy the equal-time anti-commutation relations

$$\{ \ \hat{\psi}^{\dagger}(\underline{r})_{\lambda} \ , \ \hat{\psi}(\underline{r}')_{\rho} \ \}_{+} \ = \ \delta^{3}(\underline{r} - \underline{r}') \, \delta_{\lambda,\rho} \tag{2155}$$

where λ and ρ label the components of the Dirac spinor. The anti-commutation relations for the field operators can be verified by noting that

$$
\begin{aligned}
\{ \ \hat{\psi}^{\dagger}(\underline{r}) \ , \ \hat{\psi}(\underline{r}') \ \}_{+} \ &= \ \sum_{\alpha,\beta} \ \left(\ \{ \ \hat{c}_{\alpha}^{\dagger} \ , \ \hat{c}_{\beta} \ \}_{+} \, \phi_{\alpha}^{*}(\underline{r}) \, \phi_{\beta}(\underline{r}') \ + \ \{ \ \hat{c}_{\alpha}^{\dagger} \ , \ \hat{b}_{\beta}^{\dagger} \ \}_{+} \, \phi_{\alpha}^{*}(\underline{r}) \, \phi_{\beta}^{c}(\underline{r}') \right. \\
&\qquad \left. + \ \{ \ \hat{b}_{\alpha} \ , \ \hat{c}_{\beta} \ \}_{+} \, \phi_{\alpha}^{c\,*}(\underline{r}) \, \phi_{\beta}(\underline{r}') \ + \ \{ \ \hat{b}_{\alpha} \ , \ \hat{b}_{\beta}^{\dagger} \ \}_{+} \, \phi_{\alpha}^{c\,*}(\underline{r}) \, \phi_{\beta}^{c}(\underline{r}') \ \right) \\
&= \ \sum_{\alpha,\beta} \ \left(\ \delta_{\alpha,\beta} \, \phi_{\alpha}^{*}(\underline{r}) \, \phi_{\beta}(\underline{r}') \ + \ \delta_{\alpha,\beta} \, \phi_{\alpha}^{c\,*}(\underline{r}) \, \phi_{\beta}^{c}(\underline{r}') \ \right) \\
&= \ \sum_{\alpha} \ \left(\ \phi_{\alpha}^{*}(\underline{r}) \, \phi_{\alpha}(\underline{r}') \ + \ \phi_{\alpha}^{c\,*}(\underline{r}) \, \phi_{\alpha}^{c}(\underline{r}') \ \right) \\
&= \ \delta^{3}(\underline{r} - \underline{r}')
\end{aligned}
\tag{2156}
$$

where the fermion anti-commutation relations have been used in arriving at the second line. The positive-energy states and their charge conjugated states form a complete set of basis states for the single-particle Dirac equation, so their

---

[155]W. Heisenberg and W. Pauli, Zeit. für Physik, **56**, 1 (1929).
W. Heisenberg and W. Pauli, Zeit. für Physik, **59**, 168 (1930).

completeness condition has been used in going from the third to the fourth line. The equal-time field anti-commutation relations can be generalized to field anti-commutators at space-time points with a general type of separation. In the case where the two field points x and x' have a space-like separation

$$(x^\mu - x'^\mu) (x_\mu - x'_\mu) < 0$$

causality dictates that the anti-commutators are zero

$$\{ \hat{\psi}^\dagger(x) , \hat{\psi}(x') \}_+ = 0$$

That is, for space-like separations, there is no causal connection[156] so a measurement of a local field at x' cannot affect a measurement at x. N. Bohr and
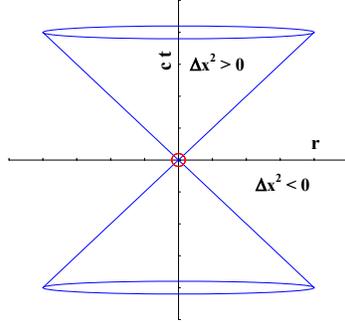

Figure 70: Due to causality, the anti-commutator of the field operator should vanish for space-like separations. The anti-commutators can be non-zero inside or on the light cone.

L. Rosenfeld[157] have put forward general arguments that the commutation relations also place limitations on the measurement of fields at time-like separations.

The Hamiltonian density for the (non-interacting) quantized Dirac field theory can be expressed as the operator

$$\hat{\mathcal{H}} = \hat{\psi}^\dagger \, \gamma^{(0)} \, c \left(- i \, \hbar \, \underline{\gamma} \, . \, \underline{\nabla} + m \, c \right) \hat{\psi} \tag{2157}$$

and the Hamiltonian operator is given by

$$\hat{H} = \int d^3\underline{r} \, \hat{\mathcal{H}} \tag{2158}$$

When the expansion of the quantized field in terms of single-particle wave functions is substituted into the Hamiltonian, one finds

$$\hat{H} = \sum_\alpha \left(E_\alpha \, \hat{c}_\alpha^\dagger \, \hat{c}_\alpha + E_\alpha^c \, \hat{b}_\alpha \, \hat{b}_\alpha^\dagger \right)$$

---

[156]Outside the light-cone there is no way to distinguish between future and past.
[157]N. Bohr and L. Rosenfeld, Kon. Dansk. Vid. Selskab., Mat.-Fys. Medd. **XII**, 8 (1933).

$$= \sum_\alpha \left(E_\alpha \, \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \; - \; E_\alpha \, \hat{b}_\alpha \, \hat{b}_\alpha^\dagger \right) \qquad (2159)$$

where the expression for the energy of the charge conjugated state

$$E_\alpha^c \; = \; - \, E_\alpha \qquad (2160)$$

has been used. On anti-commuting the positron and annihilation operators, one finds

$$\hat{H} \; = \; \sum_\alpha E_\alpha \left(\hat{c}_\alpha^\dagger \, \hat{c}_\alpha \; + \; \hat{b}_\alpha^\dagger \, \hat{b}_\alpha \; - \; 1 \right) \qquad (2161)$$

The last term, when summed over α, yields the infinitely negative energy of Dirac's vacuum in which all the negative-energy states are filled. The vacuum energy shall be used as the reference energy, so the Hamiltonian becomes

$$\hat{H} \; = \; \sum_\alpha E_\alpha \left(\hat{c}_\alpha^\dagger \, \hat{c}_\alpha \; + \; \hat{b}_\alpha^\dagger \, \hat{b}_\alpha \right) \qquad (2162)$$

which describes the energy of the excited state as the sum of the energies of the excited electrons and the excited positrons. The energies of the positrons and electrons are given by positive numbers.

The momentum operator defined by Noether's theorem is found as

$$\underline{\hat{P}} \; = \; \sum_{\underline{k},\sigma} \hbar \, \underline{k} \left(\hat{c}_{\underline{k},\sigma}^\dagger \, \hat{c}_{\underline{k},\sigma} \; + \; \hat{b}_{\underline{k},\sigma}^\dagger \, \hat{b}_{\underline{k},\sigma} \right) \qquad (2163)$$

which is just the sum of the momenta of the (positive-energy) electrons and the positrons. The spin operator is defined as

$$\underline{\hat{S}} \; = \; \frac{\hbar}{2} \int d^3\underline{r} \; \hat{\psi}^\dagger \, \underline{\hat{\sigma}} \, \hat{\psi} \qquad (2164)$$

This is evaluated by substituting the expression for the field operators in terms of the single-particle wave functions and the particle creation and annihilation operators. The expectation value of the spin operator in the charge conjugated state ϕ_α^c is given by

$$\int d^3\underline{r} \; \phi_\alpha^{c\,\dagger}(\underline{r}) \, \underline{\hat{\sigma}} \, \phi_\alpha^c(\underline{r}) \;\; = \;\; - \left(\int d^3\underline{r} \; \phi_\alpha^\dagger(\underline{r}) \, \gamma^{(2)} \, \underline{\sigma}^* \, \gamma^{(2)} \, \phi_\alpha(\underline{r}) \right)^*$$

$$= \;\; \left(\int d^3\underline{r} \; \phi_\alpha^\dagger(\underline{r}) \, \sigma^{(2)} \, \underline{\hat{\sigma}}^* \, \sigma^{(2)} \, \phi_\alpha(\underline{r}) \right)^*$$

$$= \;\; - \left(\int d^3\underline{r} \; \phi_\alpha^\dagger(\underline{r}) \, \underline{\hat{\sigma}} \, \phi_\alpha(\underline{r}) \right)^*$$

$$= \;\; - \left(\int d^3\underline{r} \; \phi_\alpha^\dagger(\underline{r}) \, \underline{\hat{\sigma}} \, \phi_\alpha(\underline{r}) \right) \qquad (2165)$$

381

The third line follows from the identity

$$\sigma^{(2)} \, \hat{\underline{\sigma}}^* \, \sigma^{(2)} \; = \; - \, \hat{\underline{\sigma}} \tag{2166}$$

The last line follows since $\underline{\sigma}$ is Hermitean. Hence, the spin operator is evaluated as

$$\hat{\underline{S}} \; = \; \frac{\hbar}{2} \sum_{\underline{k};\sigma',\sigma''} \chi^\dagger_{\sigma''} \, \underline{\sigma} \, \chi_{\sigma'} \left(\hat{c}^\dagger_{\underline{k},\sigma''} \, \hat{c}_{\underline{k},\sigma'} \; + \; \hat{b}^\dagger_{\underline{k},\sigma''} \, \hat{b}_{\underline{k},\sigma'} \right) \tag{2167}$$

which is just the sums of the spins of the electrons and positrons.

Finally, the conserved Noether charge corresponding to the global gauge invariance is given by

$$\begin{aligned}
\hat{Q} \; &= \; \int d^3\underline{r} \; \hat{\psi}^\dagger(\underline{r}) \, \hat{\psi}(\underline{r}) \\
&= \; \sum_\alpha \left(\hat{c}^\dagger_\alpha \, \hat{c}_\alpha \; + \; \hat{b}_\alpha \, \hat{b}^\dagger_\alpha \right) \\
&= \; \sum_\alpha \left(\hat{c}^\dagger_\alpha \, \hat{c}_\alpha \; - \; \hat{b}^\dagger_\alpha \, \hat{b}_\alpha \; + \; 1 \right)
\end{aligned} \tag{2168}$$

The last term in the parenthesis, when summed over all states α, yields the total charge of the vacuum which is to be discarded. Hence, the observable charge is defined as

$$\hat{Q} \; = \; \sum_\alpha \left(\hat{c}^\dagger_\alpha \, \hat{c}_\alpha \; - \; \hat{b}^\dagger_\alpha \, \hat{b}_\alpha \right) \tag{2169}$$

which shows that the total electrical charge defined as the difference between the number of electrons and the number of positrons is conserved.

14.3 Parity, Charge and Time Reversal Invariance

The Lagrangian density may posses continuous symmetries and it may also posses discrete symmetries. Some of the discrete symmetries are examined below.

14.3.1 Parity

The parity eigenvalue equation for a multi-particle state with parity η_ψ can be expressed as

$$\hat{\mathcal{P}} \mid \psi \; > \; = \; \eta_\psi \mid \psi \; > \tag{2170}$$

Since the action of the parity operator on states is described by a unitary operator, operators transform under parity according to the general form of a unitary

transformation. In particular, the effect of the parity transformation on the field operator is determined as

$$\hat{\psi}(\underline{r}) \;\rightarrow\; \hat{\psi}'(\underline{r}') \;=\; \hat{\mathcal{P}} \; \hat{\psi}(\underline{r}) \; \hat{\mathcal{P}} \tag{2171}$$

The parity transformation is going to be determined in analogy with the parity transformation of a classical field, in which the creation and annihilation operators are replaced by complex numbers. The parity operation on the quantum field can be interpreted as only acting on the wave functions and not the particle creation and annihilation operators. Quantum mechanically, this corresponds to viewing the parity operator as changing the properties of the states to the properties associated with the parity reversed states. Since the field operator is expressed as

$$\hat{\psi}(\underline{r}) \;=\; \sum_{\alpha} \left(\; \hat{c}_{\alpha} \; \phi_{\alpha}(\underline{r}) \;+\; \hat{b}_{\alpha}^{\dagger} \; \phi_{\alpha}^{c}(\underline{r}) \; \right) \tag{2172}$$

one has

$$\hat{\mathcal{P}} \; \hat{\psi}(\underline{r}) \; \hat{\mathcal{P}} \;=\; \sum_{\alpha} \left(\; \hat{c}_{\alpha} \; \hat{\mathcal{P}} \; \phi_{\alpha}(\underline{r}) \; \hat{\mathcal{P}} \;+\; \hat{b}_{\alpha}^{\dagger} \; \hat{\mathcal{P}} \; \phi_{\alpha}^{c}(\underline{r}) \; \hat{\mathcal{P}} \; \right) \tag{2173}$$

However, under a parity transform a general Dirac spinor satisfies

$$\begin{aligned}
\hat{\mathcal{P}} \; \phi_{\alpha}(\underline{r}) \;&=\; \eta_{\alpha}^{P} \; \phi_{\mathcal{P}\alpha}(\underline{r}) \\
\hat{\mathcal{P}} \; \phi_{\alpha}^{c}(\underline{r}) \;&=\; \eta_{\alpha}^{Pc} \; \phi_{\mathcal{P}\alpha}^{c}(\underline{r})
\end{aligned} \tag{2174}$$

where η_{α}^{P} is a phase factor which represents the intrinsic parity of the state. Furthermore, since $\hat{\mathcal{P}}^{2} \;=\; \hat{I}$, then the intrinsic parities η_{α}^{P} and η_{α}^{Pc} have to satisfy the conditions

$$\begin{aligned}
(\, \eta_{\alpha}^{P} \,)^{2} \;&=\; 1 \\
(\, \eta_{\alpha}^{Pc} \,)^{2} \;&=\; 1
\end{aligned} \tag{2175}$$

So the intrinsic parities are ± 1. The intrinsic parity of a state $\phi_{\alpha}(\underline{r})$ and its charge conjugated state $\phi_{\alpha}^{c}(\underline{r})$ are related by

$$\eta_{\alpha}^{Pc} \;=\; -\, \eta_{\alpha}^{P} \tag{2176}$$

This follows since charge conjugation flips the upper and lower two-component spinors and these two-component spinors have opposite intrinsic parity. Therefore, the state $\phi_{\alpha}(\underline{r})$ and the charge conjugates state $\phi_{\alpha}^{c}(\underline{r})$ have opposite parities. Therefore, it follows that the field operator transforms as

$$\hat{\mathcal{P}} \; \hat{\psi}(\underline{r}) \; \hat{\mathcal{P}} \;=\; \sum_{\alpha} \left(\; \eta_{\alpha}^{P} \; \hat{c}_{\alpha} \; \phi_{\mathcal{P}\alpha}(\underline{r}) \;-\; \eta_{\alpha}^{P} \; \hat{b}_{\alpha}^{\dagger} \; \phi_{\mathcal{P}\alpha}^{c}(\underline{r}) \; \right) \tag{2177}$$

so the quantum field operators transforms in a similar fashion to the classical field.

The relations between parity reversed states and parity reversed charge conjugated states can be verified by examining the free particle solutions of the Dirac equation and noting that the parity operator consists of the product of $\gamma^{(0)}$ and spatial inversion $\underline{r} \rightarrow -\underline{r}$. This spatial inversion acting on a wave function with momentum $\underline{k}$ and spin σ becomes a wave function with momentum $-\underline{k}$ and spin σ, up to a constant of proportionality. A free particle momentum eigenstate is given by

$$\phi_{\sigma,\underline{k}}(x) = \mathcal{N} \left(\begin{array}{c} \chi_\sigma \\ \frac{c \hbar \underline{k} \cdot \underline{\sigma}}{E + m c^2} \chi_\sigma \end{array} \right) \exp\left[-i (k_0 x^{(0)} - \underline{k} \cdot \underline{r}) \right] \quad (2178)$$

The application of the parity operator to the above wave function yields

$$\begin{aligned} \hat{\mathcal{P}} \phi_{\sigma,\underline{k}}(x) &= \mathcal{N} \gamma^{(0)} \left(\begin{array}{c} \chi_\sigma \\ \frac{c \hbar \underline{k} \cdot \underline{\sigma}}{E + m c^2} \chi_\sigma \end{array} \right) \exp\left[-i (k_0 x^{(0)} + \underline{k} \cdot \underline{r}) \right] \\ &= \mathcal{N} \left(\begin{array}{c} \chi_\sigma \\ -\frac{c \hbar \underline{k} \cdot \underline{\sigma}}{E + m c^2} \chi_\sigma \end{array} \right) \exp\left[-i (k_0 x^{(0)} + \underline{k} \cdot \underline{r}) \right] \\ &= \phi_{\sigma,-\underline{k}}(x) \end{aligned} \quad (2179)$$

as anticipated. The charge conjugated state is given by

$$\begin{aligned} \phi^c_{\sigma,\underline{k}}(x) &= \hat{C} \phi^*_{\sigma,\underline{k}}(x) \\ &= \mathcal{N} \hat{C} \left(\begin{array}{c} \chi^*_\sigma \\ \frac{c \hbar \underline{k} \cdot \underline{\sigma}^*}{E + m c^2} \chi^*_\sigma \end{array} \right) \exp\left[+ i k^\mu x_\mu \right] \end{aligned} \quad (2180)$$

where

$$\begin{aligned} \hat{C} &= -i \gamma^{(2)} \\ &= \left(\begin{array}{cccc} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array} \right) \end{aligned} \quad (2181)$$

Therefore, the charge conjugate wave function is given by

$$\phi^c_{\sigma,\underline{k}}(x) = -i \hat{\sigma}^{(2)} \mathcal{N} \left(\begin{array}{c} \frac{c \hbar \underline{k} \cdot \underline{\sigma}^*}{E + m c^2} \chi^*_\sigma \\ -\chi^*_\sigma \end{array} \right) \exp\left[+ i k^\mu x_\mu \right] \quad (2182)$$

The effect of the parity operator on this state leads to

$$\begin{aligned} \hat{\mathcal{P}} \phi^c_{\sigma,\underline{k}}(x) &= -i \hat{\sigma}^{(2)} \mathcal{N} \left(\begin{array}{c} \frac{c \hbar \underline{k} \cdot \underline{\sigma}^*}{E + m c^2} \chi^*_\sigma \\ +\chi^*_\sigma \end{array} \right) \exp\left[+ i (k^{(0)} x_0 + \underline{k} \cdot \underline{r}) \right] \\ &= -i \hat{\sigma}^{(2)} \mathcal{N} \left(\begin{array}{c} -\frac{c \hbar (-\underline{k} \cdot \underline{\sigma}^*)}{E + m c^2} \chi^*_\sigma \\ +\chi^*_\sigma \end{array} \right) \exp\left[+ i (k^{(0)} x_0 + \underline{k} \cdot \underline{r}) \right] \\ &= -\phi^c_{\sigma,-\underline{k}}(x) \end{aligned} \quad (2183)$$

where in the first line the parity operator has sent $\underline{r} \rightarrow -\underline{r}$ and the factor of $\gamma^{(0)}$ has flipped the sign of the lower components. In the second line we have re-written $\underline{k}$ as $-(-\underline{k})$ in the two two-component spinor, in anticipation of the comparison with eqn(2180) which allows us to identify the factor of $\phi^c_{\sigma,-\underline{k}}(x)$. This example shows that a state and its charge conjugate have opposite intrinsic parities.

From the general form of the parity transformation on Dirac spinors, one infers that the parity transform of the field operator is given by

$$\hat{\mathcal{P}}\, \hat{\psi}(\underline{r})\, \hat{\mathcal{P}} \; = \; \sum_\alpha \left(\hat{c}_\alpha \, \eta^P_\alpha \, \phi_{\mathcal{P}\alpha}(\underline{r}) \; - \; \hat{b}^\dagger_\alpha \, \eta^P_\alpha \, \phi^c_{\mathcal{P}\alpha}(\underline{r}) \right) \qquad (2184)$$

On setting $\alpha' \; = \; \mathcal{P}\,\alpha$ and noting that $\alpha \; = \; \mathcal{P}\,\alpha'$, one finds

$$\hat{\mathcal{P}}\, \hat{\psi}(\underline{r})\, \hat{\mathcal{P}} \; = \; \sum_{\mathcal{P}\alpha'} \left(\eta^P_{\mathcal{P}\alpha'} \, \hat{c}_{\mathcal{P}\alpha'} \, \phi_{\alpha'}(\underline{r}) \; - \; \eta^P_{\mathcal{P}\alpha'} \, \hat{b}^\dagger_{\mathcal{P}\alpha'} \, \phi^c_{\alpha'}(\underline{r}) \right) \qquad (2185)$$

and on transforming the summation index from α' to α

$$\hat{\mathcal{P}}\, \hat{\psi}(\underline{r})\, \hat{\mathcal{P}} \; = \; \sum_\alpha \left(\eta^P_{\mathcal{P}\alpha} \, \hat{c}_{\mathcal{P}\alpha} \, \phi_\alpha(\underline{r}) \; - \; \eta^P_{\mathcal{P}\alpha} \, \hat{b}^\dagger_{\mathcal{P}\alpha} \, \phi^c_\alpha(\underline{r}) \right) \qquad (2186)$$

Thus, the parity operation can also be interpreted as only affecting the particle creation and annihilation operators, and not the wave functions. Quantum mechanically, this interpretation corresponds to viewing that the particles as being transferred into their parity reversed states

$$\hat{\mathcal{P}}\, \hat{\psi}(\underline{r})\, \hat{\mathcal{P}} \; = \; \sum_\alpha \left(\hat{\mathcal{P}}\, \hat{c}_\alpha\, \hat{\mathcal{P}}\, \phi_\alpha(\underline{r}) \; + \; \hat{\mathcal{P}}\, \hat{b}^\dagger_\alpha\, \hat{\mathcal{P}}\, \phi^c_\alpha(\underline{r}) \right) \qquad (2187)$$

In this new interpretation, the effects of parity on the fermion operators are found by identifying the operators multiplying the single-particle wave functions in the previous two equations. The resulting operator equations are

$$\hat{\mathcal{P}}\, \hat{c}_\alpha\, \hat{\mathcal{P}} \; = \; \eta^P_{\mathcal{P}\alpha} \, \hat{c}_{\mathcal{P}\alpha} \qquad (2188)$$

and

$$\hat{\mathcal{P}}\, \hat{b}^\dagger_\alpha\, \hat{\mathcal{P}} \; = \; -\, \eta^P_{\mathcal{P}\alpha} \, \hat{b}^\dagger_{\mathcal{P}\alpha} \qquad (2189)$$

which shows that fermion particles and anti-particles have opposite intrinsic parities. Therefore, we conclude that, irrespective of which interpretation is used, the field operator transforms as

$$\hat{\mathcal{P}}\, \hat{\psi}(\underline{r})\, \hat{\mathcal{P}} \; = \; \sum_\alpha \left(\eta^P_\alpha \, \hat{c}_\alpha \, \phi_{\mathcal{P}\alpha}(\underline{r}) \; - \; \eta^P_\alpha \, \hat{b}^\dagger_\alpha \, \phi^c_{\mathcal{P}\alpha}(\underline{r}) \right) \qquad (2190)$$

which shows that the quantum field operators transforms in a similar fashion to the classical field.

14.3.2 Charge Conjugation

Under charge conjugation, the classical Dirac field transforms as

$$\psi \;\to\; \psi^c \;=\; -\,i\,\gamma^{(2)}\,\psi^* \tag{2191}$$

(up to an arbitrary phase) since this is how the single-particle wave functions transform. Classically, the (anti-linear) charge conjugation operator $\hat{\mathcal{C}}$ is the product of complex conjugation and the unitary matrix operator $\hat{C} = -\,i\,\gamma^{(2)}$. If the classical field is expressed as a linear superposition of energy eigenfunctions, the amplitudes of the eigenfunctions are represented by complex numbers. In the charge conjugated state, these amplitudes are replaced by the complex conjugates. In the quantum field, the amplitudes must be replaced by particle creation and annihilation operators. If an amplitude is associated with an annihilation operator, then the complex conjugate of the amplitude is usually associated with a creation operator. Hence, we should expect that charge conjugation will result in the creation and annihilation operators being switched.

Since the quantum field operator is expressed as

$$\hat{\psi}(\underline{r}) \;=\; \sum_\alpha \left(\, \hat{c}_\alpha\, \phi_\alpha(\underline{r}) \;+\; \hat{b}_\alpha^\dagger\, \phi_\alpha^c(\underline{r}) \,\right) \tag{2192}$$

the charge conjugate operation $\hat{\mathcal{C}}$ transforms the field operator via

$$\hat{\psi}^c(\underline{r}) \;=\; \hat{\mathcal{C}}\,\hat{\psi}(\underline{r})\,\hat{\mathcal{C}} \;=\; \sum_\alpha \left(\, \hat{c}_\alpha^\dagger\, \hat{\mathcal{C}}\, \phi_\alpha(\underline{r})\, \hat{\mathcal{C}} \;+\; \hat{b}_\alpha\, \hat{\mathcal{C}}\, \phi_\alpha^c(\underline{r})\, \hat{\mathcal{C}} \,\right) \tag{2193}$$

where, in accord with the earlier comment about the relation between the quantum and classical fields, the single-particle operators have been replaced by their Hermitean conjugates. However, under charge conjugation general Dirac spinors satisfy

$$\begin{aligned}
\hat{\mathcal{C}}\,\phi_\alpha(\underline{r}) &= \eta^c\,\phi_\alpha^c(\underline{r}) \\
\hat{\mathcal{C}}\,\phi_\alpha^c(\underline{r}) &= \eta^c\,\phi_\alpha(\underline{r})
\end{aligned} \tag{2194}$$

therefore,

$$\hat{\psi}^c(\underline{r}) \;=\; \hat{\mathcal{C}}\,\hat{\psi}(\underline{r})\,\hat{\mathcal{C}} \;=\; \sum_\alpha \eta^c \left(\, \hat{c}_\alpha^\dagger\, \phi_\alpha^c(\underline{r}) \;+\; \hat{b}_\alpha\, \phi_\alpha(\underline{r}) \,\right) \tag{2195}$$

However, if the charge conjugation operator is to be interpreted as only acting on the single-particle operators, one has

$$\hat{\psi}^c(\underline{r}) \;=\; \sum_\alpha \left(\, \hat{\mathcal{C}}\, \hat{c}_\alpha\, \hat{\mathcal{C}}\, \phi_\alpha(\underline{r}) \;+\; \hat{\mathcal{C}}\, \hat{b}_\alpha^\dagger\, \hat{\mathcal{C}}\, \phi_\alpha^c(\underline{r}) \,\right) \tag{2196}$$

For consistency, the two expressions for $\hat{\psi}^c(\underline{r})$ must be equivalent. Hence, the operator coefficients of $\phi_\alpha(\underline{r})$ and $\phi_\alpha^c(\underline{r})$ in the two expressions should be identical. Therefore, one requires that

$$
\begin{aligned}
\hat{\mathcal{C}}\,\hat{c}_\alpha\,\hat{\mathcal{C}} &= \eta^c\,\hat{b}_\alpha \\
\hat{\mathcal{C}}\,\hat{b}_\alpha^\dagger\,\hat{\mathcal{C}} &= \eta^c\,\hat{c}_\alpha^\dagger
\end{aligned}
\tag{2197}
$$

In other words, charge conjugation replaces particles by their anti-particles and their quantum numbers α are unchanged. Furthermore, we identify the charge conjugated field operator as

$$
\hat{\psi}^c = \hat{\mathcal{C}}\,\hat{\psi}\,\hat{\mathcal{C}} = -\,i\,\eta^c\,\gamma^{(2)}\,\hat{\psi}^\dagger
\tag{2198}
$$

where $\hat{\psi}^\dagger$ is the Hermitean conjugate (column) field operator. Apart from the replacement of the complex amplitudes with the Hermitean conjugates of the creation and annihilation operators, the above expression is identical to the expression for charge conjugation on the classical field.

The charge conjugation operator has the effect of reversing the current density operator

$$
\hat{\mathcal{C}}\,\hat{\bar{\psi}}^\dagger\,\gamma^\mu\,\hat{\psi}\,\hat{\mathcal{C}} = -\,\hat{\bar{\psi}}^\dagger\,\gamma^\mu\,\hat{\psi}
\tag{2199}
$$

which is understood as the result in the change of the charge's sign.

14.3.3   Time Reversal

The time-reversal operation interchanges the past with the future. Time reversal transforms the space-time coordinates via

$$
\hat{T}\,(ct, \underline{r}) = (-ct, \underline{r})
\tag{2200}
$$

Thus, under time reversal, the time and spatial components of the position four-vector have different transformational properties. Furthermore, the energy-momentum four-vector transforms as

$$
\hat{T}\,(p^{(0)}, \underline{p}) = (p^{(0)}, -\underline{p})
\tag{2201}
$$

Hence, the position four-vector and momentum four-vector have different transformational properties. Due to the above properties, angular momentum (including spin) transforms as

$$
\hat{T}\,\underline{J} = -\,\underline{J}
\tag{2202}
$$

Therefore, we find that time reversal reverses momenta and flips spins.

According to the Wigner theorem[158], time reversal can only be implemented by an anti-linear anti-unitary transformation. Since the time reversal operator

---

[158]E. P. Wigner. Göttinger. Nachr. **31**, 546 (1932).

$\hat{T}$ interchanges the initial and final states, then

$$
\begin{aligned}
< \ \hat{T} \ \psi_f \mid \hat{T} \ \psi_i \ > \ &= \ < \ \psi_i \mid \psi_f \ > \\
&= \ < \ \psi_f \mid \psi_i \ >^*
\end{aligned}
\tag{2203}
$$

Thus, $\hat{T}$ must be an anti-unitary operator. Furthermore, if the initial state is given by a linear superposition

$$
\mid \psi_i \ > \ = \ \sum_\alpha C_\alpha \mid \phi_\alpha \ >
\tag{2204}
$$

then the overlap is given by

$$
\sum_\alpha < \ \hat{T} \ \psi_f \mid \hat{T} \ C_\alpha \mid \phi_\alpha \ > \ = \ \sum_\alpha C_\alpha^* < \ \psi_f \mid \phi_\alpha \ >^*
\tag{2205}
$$

Hence, one infers that

$$
\hat{T} \ \sum_\alpha C_\alpha \mid \phi_\alpha \ > \ = \ \sum_\alpha C_\alpha^* \ \hat{T} \mid \phi_\alpha \ >
\tag{2206}
$$

which is the definition of an anti-linear operator and so we identify $\hat{T}$ as an anti-linear operator.

It can be shown that the time-reversed Dirac wave function defined by

$$
\hat{T} \ \psi(t, \underline{r}) \ = \ - \ \gamma^{(1)} \ \gamma^{(3)} \ \psi^*(-t, \underline{r})
\tag{2207}
$$

satisfies the Dirac equation with $t \ \rightarrow \ - \ t$. For example, the plane wave solutions of the Dirac equation can be shown to transform as

$$
\begin{aligned}
\hat{T} \ \phi_{\underline{\sigma}, \underline{k}}(\underline{r}, t) \ &= \ - \ \gamma^{(1)} \ \gamma^{(3)} \ \phi^*_{\underline{\sigma}, \underline{k}}(\underline{r}, -t) \\
&= \ \phi_{-\underline{\sigma}, -\underline{k}}(\underline{r}, t)
\end{aligned}
\tag{2208}
$$

which flips the momentum and the spin angular momentum. It should be noted that the matrix operator $\gamma^{(1)} \ \gamma^{(3)}$ does not couple the upper and lower two-component spinors, but nevertheless is closely related to the operator $- \ i \ \gamma^{(2)}$ which occurs in the charge conjugation operator.

Also, if the Dirac field operator is required to satisfy

$$
\hat{T} \ \hat{\psi}(t, \underline{r}) \ \hat{T} \ = \ - \ \gamma^{(1)} \ \gamma^{(3)} \ \hat{\psi}^*(-t, \underline{r})
\tag{2209}
$$

then the single-particle operators must satisfy

$$
\begin{aligned}
\hat{T} \ c_\alpha \ \hat{T} \ &= \ c_{T\alpha} \\
\hat{T} \ b_\alpha \ \hat{T} \ &= \ b_{T\alpha}
\end{aligned}
\tag{2210}
$$

which correspond to particles following time-reversed trajectories.

Table 19: **Discrete Symmetries of Particles.**
The charge conjugated of a state is a negative energy state with momentum $-\underline{p}$ and spin $-\underline{\sigma}$, that is interpreted as the state of antiparticle with momentum $\underline{p}$ and spin $\underline{\sigma}$.

	Q	$\underline{p}$	$\underline{\sigma}$	Λ
Charge Conjugation	$-$	$+$	$+$	$+$
Parity	$+$	$-$	$+$	$-$
Time Reversal	$+$	$-$	$-$	$+$
CPT	$-$	$+$	$-$	$-$

It is known that the weak interaction violates parity invariance. However, there was a slight possibility that the weak interaction conserves the combined operation of charge conjugation and spatial inversion[159]. Christenson, Cronin, Fitch and Turlay[160] performed experiments which showed that the combined operation $C\,P$ is violated in the decay of K mesons. There is reason to believe that the weak interaction is invariant under the combined symmetry operation $C\,P\,T$, since this is related to Lorentz invariance.

The combined symmetry operation $\hat{\mathcal{C}}\,\hat{\mathcal{P}}\,\hat{\mathcal{T}}$ transforms a Dirac spinor as

$$
\begin{aligned}
\psi'(x') &= \hat{\mathcal{C}}\,\hat{\mathcal{P}}\,\hat{\mathcal{T}}\,\psi(x) \\
&= -\,i\,\gamma^{(2)}\left(\hat{\mathcal{P}}\,\hat{\mathcal{T}}\,\psi(x)\right)^* \\
&= +\,i\,\gamma^{(2)}\left(\gamma^{(0)}\,\gamma^{(1)}\,\gamma^{(3)}\,\psi^*(-x)\right)^* \\
&= i\,\gamma^{(2)}\,\gamma^{(0)}\,\gamma^{(1)}\,\gamma^{(3)}\,\psi(-x) \\
&= i\,\gamma^{(0)}\,\gamma^{(1)}\,\gamma^{(2)}\,\gamma^{(3)}\,\psi(-x) \\
&= \gamma^{(4)}\,\psi(-x)
\end{aligned}
\tag{2211}
$$

[159] J. C. Wick, A. S. Wightman and E. P. Wigner, Phys. Rev. **88**, 101 (1952).
[160] J. Christenson, J. W. Cronin, V. I. Fitch and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).

14.3.4 The CPT Theorem

The CPT theorem states that any local[161] quantum field theory with a Hermitean Lorentz invariant Lagrangian which satisfies the spin-statistics theorem, is invariant under the compound operation $\hat{\mathcal{C}} \, \hat{\mathcal{P}} \, \hat{\mathcal{T}}$, where the operators can be placed in any order.

The proof of the theorem relies on the fact that any Lorentz invariant quantity must be created out of contracting the indices of bi-linear covariants (quantities such as the current density j_μ which involve products of the γ_μ) with the indices of contravariant derivatives ∂^μ. Since the joint operation $\hat{\mathcal{P}} \, \hat{\mathcal{T}}$ results in each of the contravariant derivatives ∂^μ in the product changing sign, the theorem ensures that the corresponding bi-linear covariants with which the derivatives are contracted with must undergo an equivalent number of sign changes under the compound operation $\hat{\mathcal{C}} \, \hat{\mathcal{P}} \, \hat{\mathcal{T}}$. The theorem only assumes invariance under proper orthochronous Lorentz transformations and makes no assumptions about reflection. The improper transformations are treated as analytic continuation of the Lorentz transformation into complex space-time. The theorem was first discussed by Lüders[162] and Pauli[163], and then by Lee, Oehme and Yang[164].

The theorem has several consequences, such as the equality of the masses of particles and their anti-particles. This follows since the mass mc is an eigenvalue of $\hat{p}^{(0)}$ in the particle's rest frame and since one can find simultaneous eigenstates of the commuting operators $\hat{p}^\mu$ and the product $\hat{\mathcal{C}} \, \hat{\mathcal{P}} \, \hat{\mathcal{T}}$. If one denotes the compound operator as

$$\hat{\Theta} \; = \; \hat{\mathcal{C}} \, \hat{\mathcal{P}} \, \hat{\mathcal{T}} \tag{2212}$$

then

$$
\begin{aligned}
< \, \Psi \mid \hat{H} \mid \Psi \, > \; &= \; < \, \Psi \mid \hat{\Theta}^{-1} \, \hat{\Theta} \, \hat{H} \, \hat{\Theta}^{-1} \, \hat{\Theta} \mid \Psi \, > \\
&= \; < \, \Psi \mid \hat{\Theta}^{-1} \, \hat{H} \, \hat{\Theta} \mid \Psi \, >
\end{aligned} \tag{2213}
$$

since the CPT theorem ensures that $\hat{\Theta}$ commutes with the Hamiltonian

$$\hat{\Theta} \, \hat{H} \, \hat{\Theta}^{-1} \; = \; \hat{H} \tag{2214}$$

If $\mid \Psi \, >$ represents a stable single-particle state, such as

$$\mid \Psi \, > \; = \; c_\alpha^\dagger \mid 0 \, > \tag{2215}$$

---

[161] A Local Field Theory is one expressible in terms of a local Lagrangian density in which interactions can be expressed in terms of products of fields at the same point in space-time. It would be truly remarkable if this concept were to continue to work at arbitrarily small distances!

[162] G. Lüders, Dan. Mat. Fys. Medd. **28**, 5 (1954).
G. Lüders, Ann. Phys. **2**, 1 (1957).

[163] W. Pauli in *Niels Bohr and the Development of Physics*, Pergamon Press, London (1955).

[164] T. D. Lee, R. Oehme and C. N. Yang, Phys. Rev. **106**, 340 (1957).

then the state $\hat{\Theta} \mid \Psi >$ describes an anti-particle with flipped angular momentum. This follows since the vacuum satisfies

$$\hat{\Theta} \mid 0 > = \mid 0 > \tag{2216}$$

Therefore, the single-particle state transforms as

$$\begin{aligned}
\hat{\Theta} \mid \Psi > &= \hat{\Theta} \, c_\alpha^\dagger \, \hat{\Theta}^{-1} \, \hat{\Theta} \mid 0 > \\
&= \hat{\Theta} \, c_\alpha^\dagger \, \hat{\Theta}^{-1} \mid 0 >
\end{aligned} \tag{2217}$$

By successive applications of $\hat{\mathcal{C}}$, $\hat{\mathcal{P}}$ and $\hat{\mathcal{T}}$, one finds that the operator $\hat{\Theta} \, c_\alpha^\dagger \, \hat{\Theta}^{-1}$ reduces to the creation operator for the anti-particle with reversed angular momentum. Therefore, the state $\hat{\Theta} \mid \Psi >$ describes an anti-particle with flipped angular momentum. From the equality of the expectation values

$$< \Psi \mid \hat{H} \mid \Psi > = < \Psi \mid \hat{\Theta}^{-1} \, \hat{H} \, \hat{\Theta} \mid \Psi > \tag{2218}$$

one finds that the energy of a particle is equal to the energy of an anti-particle with a reversed spin. However, as the rest mass cannot depend on the angular momentum, the mass of a particle is equal to the mass of its anti-particle. For unstable particles, the equality of the mass of the particle and anti-particle is ensured by the invariance of the S-matrix under $\hat{\mathcal{C}} \, \hat{\mathcal{P}} \, \hat{\mathcal{T}}$.

Likewise, one can use the CPT theorem to show that the total decay rate of a particle into products is equal to the total decay rate of the anti-particle into its products[165]. It should be noted that the partial decay rates into specific final states are not equivalent, only the sums over all final states are equal.

14.4 The Connection between Spin and Statics

The above result for the energy operator of the Dirac field illustrates the "Spin Statistics Theorem" proposed by Pauli[166]. The theorem states that particles with half odd-integer spins obey Fermi-Dirac Statistics and particles with integer spins obey Bose-Einstein Statistics. The Dirac spinor describes spin one-half particles, and if these particles are chosen to satisfy anti-commutation relations, then the energy of the excited states is given by

$$\hat{H}_{\text{Dirac}} = \sum_\alpha E_\alpha \left(\hat{c}_\alpha^\dagger \, \hat{c}_\alpha + \hat{b}_\alpha^\dagger \, \hat{b}_\alpha \right) \tag{2219}$$

which only has positive excitation energies. Hence, if the wave function changes sign under the interchange of a pair of spin one-half particles the energy is bounded from below. If the field operators had been chosen to obey commutation relation, then the wave function would have been symmetric under the

[165] T. D. Lee, R. Oehme and C. N. Yang, Phys. Rev. **106**, 340 (1957).
[166] W. Pauli, Phys. Rev. **58**, 716 (1940).

interchange of particles. If this were the case, there would be a negative sign in front of the positron energies so that the energy would have been unbounded from below. This would have implied that the vacuum would not be stable, and the theory is erroneous. This can be taken as implying that spin one-half particles must obey Fermi-Dirac Statistics. The other part of the theorem compels integer spin particles to be bosons. Therefore, since photons have spin one, the expression for the energy of the electromagnetic field is considered to be given by

$$\hat{H}_{\text{Photon}} \;=\; \sum_{\underline{k},\alpha} \frac{\hbar \, \omega_k}{2} \left[\, \hat{a}^{\dagger}_{\underline{k},\alpha} \, \hat{a}_{\underline{k},\alpha} \;+\; \hat{a}_{\underline{k},\alpha} \, \hat{a}^{\dagger}_{\underline{k},\alpha} \, \right] \qquad (2220)$$

This Hamiltonian represents the energy of a spin-one particle. The photon creation and annihilation operators satisfy commutation relations, therefore, the energy can be expressed as

$$\hat{H}_{\text{Photon}} \;=\; \sum_{\underline{k},\alpha} \frac{\hbar \, \omega_k}{2} \left[\, 2 \, \hat{a}^{\dagger}_{\underline{k},\alpha} \, \hat{a}_{\underline{k},\alpha} \;+\; 1 \, \right] \qquad (2221)$$

which is the sum of the vacuum energy (the zero-point energies) and the energies of each excited photon. The excitation energies are positive. If it had been assumed that the photon wave functions were anti-symmetric under the interchange of particles, then one would have found that the photon energies would have been identically equal to zero. Furthermore, the excited photons would have carried zero momentum and, therefore, be completely void of any physical consequence. Hence, one concludes that spin-one photons must obey Bose-Einstein Statistics. The generalized theorem[167] is an assertion that a non-trivial integer spin field cannot have a anti-commutator that vanishes for space-like separations and a non-trivial odd half-integer spin field cannot have a commutator that vanishes for space-like separations.

15 Massive Gauge Field Theory

Following Yang and Mills[168], we shall consider a two-component complex scalar field. The field can be expressed as a two-component field, representing states with different isospin

$$\Phi \;=\; \left(\begin{array}{c} \Phi_1 \\ \Phi_2 \end{array} \right) \qquad (2222)$$

where the Φ_i are complex scalars. That is

$$\begin{aligned} \Phi_1 &= \Re e \, \Phi_1 \;+\; i \, \Im m \, \Phi_1 \\ \Phi_2 &= \Re e \, \Phi_2 \;+\; i \, \Im m \, \Phi_2 \end{aligned} \qquad (2223)$$

[167]R. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That*, Princeton Univ. Press (2000).
[168]C. N. Yang and R. L. Mills, Phys. Rev. **96**, 191 (1964).

This is equivalent to assuming four independent real fields. The inner product is defined as

$$\Phi^\dagger \, \Phi \;=\; \Phi_1^* \, \Phi_1 \;+\; \Phi_2^* \, \Phi_2 \tag{2224}$$

15.1 The Gauge Symmetry

We shall assume that the Lagrangian is invariant under a generalized gauge transformations of the form

$$\Phi \;\to\; \Phi' \;=\; \exp\left[\,-\, i\, \alpha^{(0)} \,\right] \hat{U} \, \Phi \tag{2225}$$

where $\alpha^{(0)}$ is an arbitrary scalar. The invariance of the Lagrangian under multiplication of the wave function by the phase factor, is equivalent to the usual $U(1)$ gauge invariance which has been discussed in the context of the electromagnetic field. The operator $\hat{U}$ must be a unitary operator, if the norm of Φ is conserved by the generalized gauge transformation

$$\begin{aligned}
\Phi^{\dagger\prime} \, \Phi' \;&=\; \Phi^\dagger \, \hat{U}^\dagger \, \hat{U} \, \Phi \\
&=\; \Phi^\dagger \, \Phi
\end{aligned} \tag{2226}$$

Therefore, one requires

$$\hat{U}^\dagger \, \hat{U} \;=\; \hat{I} \tag{2227}$$

and so $\hat{U}$ must be a unitary operator. The operator $\hat{U}$ is assumed to be an arbitrary unitary matrix that acts on isospin states, that is, it acts on the two components of Φ. Furthermore, it shall be assumed that the unitary matrix has determinant $+ 1$. Hence, the Lagrangian is assumed to be invariant under a set of $SU(2)$ gauge transformations. A general transformation of $SU(2)$ is generated by the three operators

$$\begin{aligned}
\tau^{(1)} \;&=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\tau^{(2)} \;&=\; \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
\tau^{(3)} \;&=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\end{aligned} \tag{2228}$$

where these matrices generate a Lie algebra. That is, the algebra of the commutation relations is closed, since

$$[\, \tau^{(i)} \,,\, \tau^{(j)} \,] \;=\; 2\, i\, \xi^{i,j,k} \, \tau^{(k)} \tag{2229}$$

where $\xi^{i,j,k}$ is the antisymmetric Levi-Civita symbol. An arbitrary unitary transformation can be expressed as

$$\hat{U} \;=\; \exp\left[\,-\, i \sum_k \alpha^k \, \tau^{(k)} \,\right] \tag{2230}$$

393

where the α_k are three real quantities. This represents an arbitrary rotation in isospin space[169]. The $U(1)$ gauge transformation can also be represented in the same way. Namely, the $U(1)$ transformation can be expressed as

$$\hat{U}_0 \ = \ \exp\left[\ - \ i \ \alpha^{(0)} \ \tau^{(0)} \ \right] \tag{2231}$$

where $\tau^{(0)}$ is the unit matrix

$$\tau^{(0)} \ = \ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2232}$$

We should note that since $\tau^{(0)}$ commutes with all isospin operators, the $U(1)$ symmetry is decoupled from the $SU(2)$ symmetry, and hence, when a coupling to gauge fields is introduced, the $U(1)$ gauge field may have a coupling constant which is different from the coupling constant for the three $SU(2)$ gauge fields.

15.2   The Coupling to the Gauge Field

We shall start with a Lagrangian density $\mathcal{L}_{scalar}$ describing the field free two component scalar field, given by

$$\mathcal{L}_{scalar} \ = \ \partial_\mu \ \Phi^\dagger \ \partial^\mu \ \Phi \ - \ V(\Phi^\dagger \ \Phi) \tag{2233}$$

where $V(\Phi^\dagger\Phi)$ is an arbitrary scalar potential. For example, in a Klein-Gordon field theory describing particles with mass m

$$V(\Phi^\dagger \ \Phi) \ = \ \left(\frac{m \ c}{\hbar} \right)^2 \ \Phi^\dagger \ \Phi \tag{2234}$$

The Lagrangian is invariant under the combined gauge transformation if the quantities α^k are independent of x. In this case, the field is invariant under the transformation which is identical at each point in space, so the Lagrangian is said to have a global gauge invariance.

We shall alter the Lagrangian, such that it is invariant under a gauge transformation which varies from point to point in space. These are local gauge transformations, in which the $\alpha^k(x)$ depend on x. If the Lagrangian is to be invariant under local gauge transformations, then one must introduce a coupling to gauge fields A_μ. This coupling compensates for the change of the derivatives under the gauge transformation, so that

$$\left[\left(\partial_\mu \ - \ i \ g \ A_\mu \right) \Phi \right]^\dagger \left[\left(\partial^\mu \ - \ i \ g \ A^\mu \right) \Phi \right]$$

$$= \ \left[\left(\partial_\mu \ - \ i \ g \ A'_\mu \right) \Phi' \right]^\dagger \left[\left(\partial^\mu \ - \ i \ g \ A^{\mu\prime} \right) \Phi' \right] \tag{2235}$$

---

[169]We shall not stop and contemplate the question of what restricts our measurements have to be quantized along the isospin z-direction, and shall not ponder why there is a super-selection rule at work.

Since
$$\Phi = \hat{U}^\dagger \Phi' \tag{2236}$$

we require that
$$\left[\left(\partial^\mu - i\,g\,A^{\mu\prime} \right) \Phi' \right] = \hat{U} \left[\left(\partial^\mu - i\,g\,A^\mu \right) \Phi \right] \tag{2237}$$

so the fields A^μ must transform as
$$A^{\mu\prime} = \hat{U}\,A^\mu\,\hat{U}^\dagger + \frac{i}{g}\,\hat{U} \left(\partial^\mu\,\hat{U}^\dagger \right) \tag{2238}$$

where the derivative only acts on the unitary transformation. Since the $\hat{U}$ are generated by $\tau^{(k)}$, there must be four components of A^μ, i.e. the fields have four components $A^{\mu,k}$. The matrix form of A^μ is given by

$$
\begin{aligned}
A^\mu &= \sum_{k=1}^3 A^{\mu,k}\,\tau^{(k)} \\
&= \begin{pmatrix} A^{\mu,(3)} & A^{\mu,(1)} - i\,A^{\mu,(2)} \\ A^{\mu,(1)} + i\,A^{\mu,(2)} & -A^{\mu,(3)} \end{pmatrix}
\end{aligned} \tag{2239}
$$

Under a gauge transformation $\hat{U}$, the vectors $A^{\mu,k}$ are transformed in isospin space. For a global gauge transformation, the transformation is a rotation in isospin space. The gauge field A^μ is also required to transform as a four-vector under Lorentz transformations.

We shall identify the contravariant derivative for the massive scalar particles as[170] as
$$\mathcal{D}^\mu = \partial^\mu - i\,g\,A^\mu - i\,g^0\,A_0^\mu \tag{2240}$$

and one recognizes that this has the same form as the coupling of charged particles to the EM field. In that case, the coupling occurs solely via $\tau^{(0)}$, the coupling constant is given by $g^0 = \frac{q}{\hbar\,c}$ and the field $A^{\mu(0)} = A^\mu$ is the four-vector potential. Since $\tau^{(0)}$ commutes with all isospin operators, it is not necessary to consider g^0 to be identical with the g value for the $SU(2)$ gauge fields.

15.3 The Free Gauge Fields

We have four real four-vector fields $A^{\mu,k}$. These are the gauge fields. The free gauge fields exist in the absence of the particles, and has a free Lagrangian. The field strength tensors $F^{\mu,\nu}$ are given by the $SU(2)$ generalized form of the EM field tensor
$$F^{\mu,\nu} = \mathcal{D}^\mu\,A^\nu - \mathcal{D}^\nu\,A^\mu \tag{2241}$$

---

[170]This can be related to the contravariant derivative familiar in the context of general relativity, if one follows the logic adopted by Weyl and considers GR as a gauge field theory.

where $\mathcal{D}$ is the covariant derivative only involving the $SU(2)$ triplet of gauge fields. It should be noted that since the gauge fields do not commute, this involves terms which are second-order in the field amplitudes. That is

$$F^{\mu,\nu} = \left(\partial^\mu A^\nu - \partial^\nu A^\mu \right) - i\,g \left(A^\mu A^\nu - A^\nu A^\mu \right) \qquad (2242)$$

The quadratic terms can be evaluated by using the commutation relations of the isospin operators $\tau^{(k)}$. The k-th component of the $SU(2)$ triplet of gauge fields is given by

$$
\begin{aligned}
F_k^{\mu,\nu} &= \partial^\mu A_k^\nu - \partial^\nu A_k^\mu + g \sum_{\{i,j\}=1}^{3} \xi^{i,j,k} \left(A_i^\mu A_j^\nu - A_i^\nu A_j^\mu \right) \\
&= \partial^\mu A_k^\nu - \partial^\nu A_k^\mu + 2\,g \sum_{\{i>j\}=1}^{3} \xi^{i,j,k} \left(A_i^\mu A_j^\nu - A_i^\nu A_j^\mu \right)
\end{aligned}
$$
$$(2243)$$

where the indices i and j are summed over and $\xi^{i,j,k}$ is the Levi-Civita symbol. In arriving at the above expression, we have used the identity

$$\tau^{(i)}\,\tau^{(j)} = \delta^{i,j}\,\tau^{(0)} + i \sum_{k=1}^{3} \xi^{i,j,k}\,\tau^{(k)} \qquad (2244)$$

found by combining the anti-commutation and commutation relations for the Pauli spin matrices. There is no contribution to the last term in the field tensor from the $U(1)$ gauge field $A_{(0)}^\mu$ since $\tau^{(0)}$ commutes with all other matrices. The zeroth-component of the field tensor is simply given by

$$F_{(0)}^{\mu,\nu} = \partial^\mu A_{(0)}^\nu - \partial^\nu A_{(0)}^\mu \qquad (2245)$$

as expected for an electromagnetic field. Since the $SU(2)$ gauge fields don't commute, the field theory is a non-Abelian gauge field theory. Under an $SU(2)$ transformation, the field tensors transform according to

$$F^{\mu,\nu} \rightarrow F^{\mu,\nu\prime} = \hat{U}\,F^{\mu,\nu}\,\hat{U}^\dagger \qquad (2246)$$

which is just a local unitary transform in isospin space. The Lagrangian density for all the free gauge fields can be expressed as

$$\mathcal{L}_{gauge} = -\frac{1}{32\,\pi}\,\text{Trace}\,F^{\mu,\nu}\,F_{\mu,\nu} \qquad (2247)$$

where the Trace is evaluated in isospin space and takes into account that there are a total of four fields. The Lagrangian density can be expressed directly in terms of the contributions from four components of the field. The result can be expressed as

$$\mathcal{L}_{gauge} = -\frac{1}{16\,\pi} \sum_{k=0}^{3} F_{\mu,\nu}^k\,F^{\mu,\nu,k} \qquad (2248)$$

where we have decomposed the fields as

$$F_{\mu,\nu} \;=\; \sum_k F^k_{\mu,\nu}\, \tau^{(k)} \tag{2249}$$

evaluated the product of the Pauli spin matrices and used the fact that the Pauli spin matrices $\tau^{(k)}$ for $k \neq 0$ are traceless.

One can consider the k-components of the vector potential A^μ (i.e. the three real components A^μ_k for fixed μ) as forming three-vectors $\underline{A}^\mu$ in isospin space. These quantities transform as three-vectors under transformations in isospin space, and also the $\underline{A}^\mu$ transform as four-vectors under Lorentz transformations in Minkowsky space-time. The three-vector fields are spin-one bosons with isospin one. Hence, we might expect that the isospin triplet should contain two oppositely charged particles and one uncharged particle. These particles are supplemented by the particle corresponding to the single uncharged field $A^\mu_{(0)}$. In terms of this set of isospin vectors, the free gauge field Lagrangian density can be written in the form of a sum of a scalar product in isospin space and an isospin scalar

$$
\begin{aligned}
\mathcal{L}_{gauge} \;=\; & -\frac{1}{16\,\pi}\left((\partial^\mu \underline{A}^\nu - \partial^\nu \underline{A}^\mu) + 2g\,\underline{A}^\mu \wedge \underline{A}^\nu \right) \cdot \left((\partial_\mu \underline{A}_\nu - \partial_\nu \underline{A}_\mu) + 2g\,\underline{A}_\mu \wedge \underline{A}_\nu \right) \\
& -\frac{1}{16\,\pi}\left(\partial^\mu A^\nu_{(0)} - \partial^\nu A^\mu_{(0)} \right)\left(\partial_\mu A_{\nu,(0)} - \partial_\nu A_{\mu,(0)} \right)
\end{aligned}
\tag{2250}
$$

It should be noted that the Lagrangian reduces to the sum of four non-interacting electromagnetic Lagrangians in the limit $g \to 0$. However, at finite values of g, the Lagrangian density contains cubic and quartic interactions with coupling strengths that are fixed by gauge invariance in terms of the single gauge parameter g.
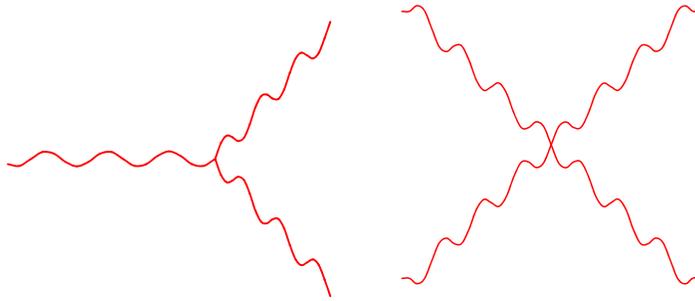

Figure 71: The interaction vertices representing the interaction of three and four isospin triplet gauge field bosons.

Exercise:

397

Determine the equations of motion for the vector gauge fields, in the presence of a source term

$$\mathcal{L}_{int} \;=\; -\; \frac{1}{c} \; \text{Trace} \;(\; A_\mu \,.\, j^\mu \;) \tag{2251}$$

where the current source j^μ has also been decomposed in terms of Pauli spin matrices.

It is convenient to introduce the two combinations

$$A_\mu^\pm \;=\; \frac{1}{\sqrt{2}} \left(\; A_\mu^{(1)} \;\mp\; i\; A_\mu^{(2)} \;\right) \tag{2252}$$

which appear in the isospin matrix form of A_μ. These combinations are mutually complex conjugate. Likewise, one can introduce the combinations of the field tensors

$$F_{\mu,\nu}^\pm \;=\; \frac{1}{\sqrt{2}} \left(\; F_{\mu,\nu}^{(1)} \;\mp\; i\; F_{\mu,\nu}^{(2)} \;\right) \tag{2253}$$

which are evaluated as

$$F_{\mu,\nu}^\pm \;=\; (\; \partial_\mu \;\mp\; 2\,i\,g\,A_\mu^{(3)} \;)\; A_\nu^\pm \;-\; (\; \partial_\nu \;\mp\; 2\,i\,g\,A_\nu^{(3)} \;)\; A_\mu^\pm \tag{2254}$$

The third component of the field tensor can be written as

$$F_{\mu,\nu}^{(3)} \;=\; (\; \partial_\mu\,A_\nu^{(3)} \;-\; \partial_\nu\,A_\mu^{(3)} \;) \;+\; 2\,i\,g\,(\; A_\mu^-\,A_\nu^+ \;-\; A_\nu^-\,A_\mu^+ \;) \tag{2255}$$

In terms of these new combinations, the free Lagrangian for the gauge fields become

$$\mathcal{L}_{gauge} \;=\; -\; \frac{1}{16\,\pi}\,F_{\mu,\nu}^{(0)}\,F^{\mu,\nu,(0)} \;-\; \frac{1}{16\,\pi}\,F_{\mu,\nu}^{(3)}\,F^{\mu,\nu,(3)} \;-\; \frac{1}{8\,\pi}\,F_{\mu,\nu}^-\,F^{\mu,\nu,+} \tag{2256}$$

where the first two terms are recognized as being similar to the Lagrangian density for the electromagnetic field. It was first hypothesized by Sheldon Glashow that the electro-weak interaction is produced by the massless vector bosons described by the above Lagrangian[171]. Masses for the gauge bosons should not be added by hand, since the resulting theory would not be renormalizable. To retain renormalizability of the theory, and to have massive vector bosons, we need to break the symmetry.

15.4   Breaking the Symmetry

We shall assume that our massive charged scalar boson field has broken symmetry[172]. The small amplitude excitations of the field with broken symmetry will

[171]S. L. Glashow, Nuclear Physics **22**, 579-588, (1961).

[172]S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967), Abdus Salam, Proc. of the 8th Nobel Symposium, Stockholm (1968).

be modified, as will be the excitations of the gauge fields. Due to the symmetry-breaking of the scalar field, the $U(1)$ vector gauge field will become coupled to the triplet of $SU(2)$ gauge fields. When the symmetry is broken, the elementary excitations of the coupled system of fields change and these new excitations will represent the observable particles.

The potential for the two-component scalar field is chosen to be given by

$$V(\Phi) = \left(\frac{m\,c}{2\,\hbar\,\phi_0} \right)^2 \left[\Phi^\dagger\,\Phi \, - \, \phi_0^2 \right]^2 \tag{2257}$$

where ϕ_0 is a fixed real constant. The lowest-energy state described by this potential is given by

$$\Phi^\dagger\,\Phi \; = \; \phi_0^2 \tag{2258}$$

This state is degenerate with respect to global rotations in the four-dimensional space of $\Re e\,\Phi_1$, $\Im m\,\Phi_1$, $\Re e\,\Phi_2$, $\Im m\,\Phi_2$ which keeps the magnitude of Φ constant and uniform over space.

The symmetry is broken by assuming that the physical ground state corresponds to one specific choice of the uniform field Φ. Given the specific ground state which the system chooses spontaneously, one can make use of the global gauge invariance to describe the ground state Φ_0 as a field which has one non-zero component which is real. That is, α^k can be chosen so that

$$\Phi_0 \;\; = \;\; \left(\begin{array}{c} \Re e\,\Phi_1 \\ 0 \end{array} \right)$$

$$= \;\; \left(\begin{array}{c} \phi_0 \\ 0 \end{array} \right) \tag{2259}$$

The excited states can be expressed as

$$\Phi \;\; = \;\; \left(\begin{array}{c} \phi_0 \, + \, \chi_1 \\ 0 \end{array} \right) \tag{2260}$$

where the local gauge degrees of freedom have been used to make χ_1 real. This excited field is invariant under the transformation

$$\Phi \;\; \rightarrow \;\; \Phi' \;\; = \;\; \hat{U}_{EM}\,\Phi \tag{2261}$$

where $\hat{U}_{EM}$ is restricted to have the form

$$\hat{U}_{EM} \;\; = \;\; \left(\begin{array}{cc} 1 & 0 \\ 0 & \exp\left[\, -\,i\,\Lambda \, \right] \end{array} \right) \tag{2262}$$

This is a transformation in which the $U(1)$ transformation is combined with a specific $SU(2)$ transformation

$$\hat{U}_{EM} \;\; = \;\; \exp\left[\, -\,i\,\frac{\Lambda}{2} \, \right] \left(\begin{array}{cc} \exp\left[\, +\,i\,\frac{\Lambda}{2} \, \right] & 0 \\ 0 & \exp\left[\, -\,i\,\frac{\Lambda}{2} \, \right] \end{array} \right) \tag{2263}$$

and will turn out to represent the residual $U(1)$ gauge invariance of the electromagnetic field.

The Lagrangian density for the isospin doublet of scalar fields and their couplings can be evaluated for the excited state as

$$\mathcal{L}_{scalar} = (\mathcal{D}_\mu \Phi)^\dagger \, \mathcal{D}^\mu \Phi \, - \, \left(\frac{m \, c}{\hbar} \right)^2 \, \chi_1^2 \tag{2264}$$

where the covariant derivative of the doublet of scalar fields is given by

$$\mathcal{D}^\mu \Phi = \left(\begin{array}{c} \partial^\mu \chi_1 \\ 0 \end{array} \right) - i \, g_0 \left(\begin{array}{c} A^{\mu,(0)} \, (\, \phi_0 \, + \, \chi_1 \,) \\ 0 \end{array} \right) - i \, g \left(\begin{array}{c} A^{\mu,(3)} \, (\, \phi_0 \, + \, \chi_1 \,) \\ \sqrt{2} \, A^{\mu,-} \, (\, \phi_0 \, + \, \chi_1 \,) \end{array} \right) \tag{2265}$$

A new interaction strength λ can be defined as

$$\lambda = \sqrt{g_0^2 + g^2} \tag{2266}$$

and on defining an angle θ via

$$\tan \theta = \frac{g}{g_0} \tag{2267}$$

the coupling constants can be represented as

$$\begin{aligned} g_0 &= \lambda \, \cos \theta \\ g &= \lambda \, \sin \theta \end{aligned} \tag{2268}$$

Thus, the covariant derivative has the connection with the field

$$A_Z^\mu = \cos \theta \, A^{\mu,(0)} \, + \, \sin \theta \, A^{\mu,(3)} \tag{2269}$$

The field A_Z^μ will turn out to be the field that describes the neutral Z particle. The field orthogonal to the Z field is defined as

$$A_{EM}^\mu = - \, \sin \theta \, A^{\mu,(0)} \, + \, \cos \theta \, A^{\mu,(3)} \tag{2270}$$

When expressed in terms of the transformed fields and constants, the covariant derivative terms become

$$\mathcal{D}^\mu \Phi = \left(\begin{array}{c} \partial^\mu \chi_1 \\ 0 \end{array} \right) - i \, (\, \phi_0 \, + \, \chi_1 \,) \left(\begin{array}{c} \lambda \, A_Z^\mu \\ g \, \sqrt{2} \, A^{\mu,-} \end{array} \right) \tag{2271}$$

The lowest-order terms in the Lagrangian density of the non-uniform scalar field and all its couplings to the gauge fields are expressed as

$$\mathcal{L}_{scalar} = \partial_\mu \chi_1 \, \partial^\mu \chi_1 \, - \, \left(\frac{m \, c}{\hbar} \right)^2 \, \chi_1^2 \, + \, \lambda^2 \, \phi_0^2 \, A_Z^\mu \, A_{\mu,Z} \, + \, 2 \, g^2 \, \phi_0^2 \, A^{\mu+} \, A_\mu^- \tag{2272}$$

The higher-order terms, which have been neglected, describe the self-interactions between the scalar field and the residual interactions between the scalar field and the gauge fields.

The Lagrangian density for the free gauge field is

$$\mathcal{L}_{gauge} = -\frac{1}{16\,\pi} F^{(0)}_{\mu,\nu} F^{\mu,\nu,(0)} - \frac{1}{16\,\pi} F^{(3)}_{\mu,\nu} F^{\mu,\nu,(3)} - \frac{1}{8\,\pi} F^{-}_{\mu,\nu} F^{\mu,\nu,+} \quad (2273)$$

has to be expressed in terms of the new fields A^{μ}_Z and A^{μ}_{EM}. The inverse transform is given by

$$\begin{aligned}
A^{\mu,(0)} &= \cos\theta\, A^{\mu}_Z \;-\; \sin\theta\, A^{\mu}_{EM} \\
A^{\mu,(3)} &= \sin\theta\, A^{\mu}_Z \;+\; \cos\theta\, A^{\mu}_{EM}
\end{aligned} \quad (2274)$$

If one defines $F^{\mu,\nu}_Z$ and $F^{\mu,\nu}_{EM}$ as the transformed field tensors evaluated to lowest-order in the fields

$$\begin{aligned}
F^{\mu,\nu}_Z &= \partial^{\mu}\, A^{\nu}_Z \;-\; \partial^{\nu} A^{\mu}_Z \\
F^{\mu,\nu}_{EM} &= \partial^{\mu}\, A^{\nu}_{EM} \;-\; \partial^{\nu} A^{\mu}_{EM}
\end{aligned} \quad (2275)$$

then the original field tensors can be expressed as

$$\begin{aligned}
F^{\mu,\nu}_{(0)} &= \cos\theta\, F^{\mu,\nu}_Z \;-\; \sin\theta\, F^{\mu,\nu}_{EM} \\
F^{\mu,\nu}_{(3)} &= \sin\theta\, F^{\mu,\nu}_Z \;+\; \cos\theta\, F^{\mu,\nu}_{EM} \;+\; 2\,i\,g\,(\, A^{\mu}_- \, A^{\nu}_+ \;-\; A^{\nu}_- \, A^{\mu}_+\,)
\end{aligned} \quad (2276)$$

The Lagrangian density describing the small amplitude excitations of the scalar field and the gauge fields can be written as

$$\begin{aligned}
\mathcal{L}_{Free} &= \partial_{\mu}\chi_1\, \partial^{\mu}\chi_1 \;-\; \left(\frac{m\,c}{\hbar}\right)^2 \chi_1^2 \\
&\quad -\frac{1}{16\,\pi}\, F_{\mu,\nu,Z}\, F^{\mu,\nu}_Z \;+\; \lambda^2\,\phi_0^2\; A^{\mu}_Z\, A_{\mu,Z} \\
&\quad -\frac{1}{16\,\pi}\, F_{\mu,\nu,EM}\, F^{\mu,\nu}_{EM} \\
&\quad -\frac{1}{8\,\pi}\, F^{-}_{\mu,\nu}\, F^{\mu,\nu,+} \;+\; 2\,g^2\,\phi_0^2\, A^{\mu+}\, A^{-}_{\mu}
\end{aligned} \quad (2277)$$

In electro-weak theory, the first term represents the free uncharged scalar boson. The second term describes an uncharged vector particle with mass M_Z proportional $\lambda\,\phi_0$. The third term describes the uncharged massless vector particle known as the photon. From the equations of motion for $A^{\mu,\pm}$, the remaining term can be shown to describe a pair of charged particles with masses M_W proportional to $g\,\phi_0 = \sin\theta\,\lambda\,\phi_0$. These particles are known as the W^+

and W^- particles. The W^+ and W^- particles are charged and the observed charges are $\pm\, e$. The interaction mediated by the massive vector bosons is found to have a finite range ($\approx 10^{-18} m$), and is responsible for the weak interaction. The experimentally determined masses[173] are $M_W c^2 \approx 80.33$ GeV and $M_Z c^2 \approx 91.187$ GeV. Nearly all the parameters of this theory have been determined through experiment, the only exception is the mass m of the scalar particle which remains to be discovered. The ratio of the masses determines the angle θ via[174]

$$\frac{M_W}{M_Z} \;=\; \sin\theta \tag{2278}$$

which yields $\sin\theta \approx 0.8810$. The $W^\pm$ particles carry electrical charges $\pm e$ since they couple to the electromagnetic field. This can be seen by examining the $W^\pm$ field tensor

$$F_\pm^{\mu,\nu} \;=\; (\,\partial^\mu \,\mp\, 2\,i\,g\,A^{\mu,(3)}\,)\,A_\pm^\nu \;-\; (\,\partial^\nu \,\mp\, 2\,i\,g\,A^{\nu,(3)}\,)\,A_\pm^\mu \tag{2279}$$

where

$$A^{\mu,(3)} \;=\; \sin\theta\, A_Z^\mu \;+\; \cos\theta\, A_{EM}^\mu \tag{2280}$$

Therefore, the covariant derivatives of the fields $\mathcal{D}^\mu A_\pm^\nu$ couple them to the electromagnetic field A_{EM}^μ with either a positive or negative coupling constant of magnitude $2\ g\ \cos\theta$. Since only electrically charged particles couple to the electromagnetic field, one can make the identification

$$\left(\frac{e}{\hbar\ c}\right) \;=\; 2\ g\ \cos\theta \;=\; \lambda\ \sin 2\theta \tag{2281}$$

which determines the coupling strengths. Furthermore, because the coupling strengths have been completely determined, the observed masses can be used to determine ϕ_0. This leads to the identification

$$\phi_0^2 \;=\; \frac{\sin^2 2\theta}{8\ \pi\ \hbar\ c}\,\frac{(\,M_Z c^2\,)^2}{(\frac{e^2}{\hbar c})} \tag{2282}$$

which leads to $\phi_0 \approx 178$ GeV $/\sqrt{\hbar c}$, where $\hbar c \approx 197$ MeV fm. Hence, the only undetermined parameter is the mass of the Higgs particle m. Recent experiments[175] have found a narrow resonance with an energy of approximately 126 GeV. The resonance has properties consistent with those expected of the Higgs particle. The resonance decays either into two photons or two vector bosons. Yang Mills theories, even if symmetry is spontaneously broken as in the Weinberg-Salam theory, were shown to be renormalizable by G. t' Hooft[176].

---

[173] G. Arnison, A. Astbury, B. Aubert, *et al.*, Phys. Lett. B, **122** 103-116 (1983).
G. Arnison, O. C. Allkofer, A. Astbury, *et al.*, Phys. Lett. B **147**, 493-508 (1984).
[174] It is customary to define the Weinberg angle θ_W via $\cos\theta_W \;=\; \frac{M_W}{M_Z}$.
[175] Physics Letters B, **716**, 1 (2012), Physics Letters B, **716**, 30 (2012).
[176] G. t' Hooft, Nuclear Physics, **B 35**, 167 (1971), G. 't Hooft and M. Veltman Nuclear Physics B 44, 189 (1972).