Quantum Electrodynamics

Real Calculations

Interaction of Light with Matter-Quantum Electrodynamics

Part 1

The theory of interaction of light with matter is called quantum electrodynamics.

The subject is made to appear more difficult than it actually is by the very many equivalent methods by which it may be formulated.

One of the simplest is that of Fermi. We shall take another starting point by just postulating for the emission or absorption of photons.

In this form it is most immediately applicable.

DISCUSSION OF FERMI'S METHOD!

Suppose all the atoms of the universe are in a box.

Classically the box may be treated as having natural modes describable in terms of a distribution of harmonic oscillators with coupling between the oscillators and matter.

The transition to quantum electrodynamics involves merely the assumption that the oscillators are quantum mechanical instead of classical.

They then have energies $(n + 1/2)\hbar\omega$, $n = 0, 1$ ..., with zero-point energy $\hbar\omega/2$.

The box is considered to be full of photons with a distribution of energies $n\hbar\omega$.

The interaction of photons with matter causes the number of photons of type n to increase by ±1 (emission or absorption).

Waves in a box can be represented as plane standing waves, spherical waves, or plane running waves exp (i$\mathbf{K}\cdot\mathbf{x}$).

One can say there is an instantaneous Coulomb interaction e^2/r_{ij} between all charges plus transverse waves only.

Then the Coulomb forces may be put into the Schrödinger equation directly.

Other formal means of expression are Maxwell's equations in Hamiltonian form, field operators, etc.

Fermi's technique leads to an infinite self-energy term e^2/r_{ii}.

It is possible to eliminate this term in suitable coordinate systems but then the transverse waves contribute an infinity (interpretation more obscure).

This anomaly was one of the central problems of modern quantum electrodynamics.

Part 2 LAWS O F QUANTUM ELECTRODYNAMICS

Without justification at this time the "laws of quantum electrodynamics" will be stated as follows:

1. The amplitude that an atomic system will *absorb* a photon during the process of transition from one state to another is *exactly* the same as the amplitude that the same transition will be made under the influence of a potential equal to that of a classical electromagnetic wave representing that photon, provided:

(a) the classical wave is normalized to represent an energy density equal to $\hbar\omega$ times the probability per cubic centimeter of finding the photon;

(b) the real classical wave is split into two complex waves $e^{-i\omega t}$ and $e^{+i\omega t}$, and only the $e^{-i\omega t}$ part is kept;

(c) the potential acts only *once* in perturbation; that is, only terms to first order in the electromagnetic field strength should be retained.

Replacing the word "absorbed" b y "emit" in rule 1 requires only that the wave represented by $\exp(+i\omega t)$ be kept instead of $\exp(-i\omega t)$.

2. The number of states available per cubic centimeter of a given polarization is

$$d^3\mathbf{K}/(2\pi)^3$$

Note this is exactly the same as the number of normal modes per cubic centimeter as in classical theory.

3. Photons obey Bose-Einstein statistics.

That is, the states of a collection of identical photons must be symmetric (exchange photons, add amplitudes).

Also the statistical weight of a state of n identical photons is 1 instead of the classical n!

Thus, in general, a photon may be represented by a solution of the classical Maxwell equations if properly normalized.

Although many forms of expression are possible it is most convenient to describe the electromagnetic field in terms of plane waves.

A plane wave can always b e represented by a vector potential only (scalar potential made zero by suitable gauge transformation).

The vector potential representing a real classical wave is taken as

$$\mathbf{A} = a\mathbf{e}\cos(\omega t - \mathbf{K}\cdot\mathbf{x})$$

We want the normalization of A to correspond to unit probability per cubic centimeter of finding the photon.

Therefore the average energy density should be $\hbar\omega$.

Now

$$\mathbf{E} = (1/c)(\partial\mathbf{A}/\partial t) = (\omega a/c)\,\mathbf{e}\sin(\omega t - \mathbf{K}\cdot\mathbf{x})$$

and

$$|\mathbf{B}| = |\mathbf{E}|$$

for a plane wave.

Therefore the average energy density is equal to

$$(1/8\pi)\overline{(|\mathbf{E}|^2 + |\mathbf{B}|^2)} = (1/4\pi)(\omega^2 a^2/c^2)\,\overline{\sin^2(\omega t - \mathbf{K}\cdot\mathbf{x})}$$

$$= (1/8\pi)(\omega^2 a^2/c^2)$$

Setting this equal to $\hbar\omega$ we find that

$$a = \sqrt{8\pi\hbar c^2/\omega}$$

Thus

$$\mathbf{A} = \sqrt{8\pi\hbar c^2/\omega}\;\mathbf{e}\cos(\omega t - \mathbf{K}\cdot\mathbf{x})$$

$$= \sqrt{4\pi\hbar c/2\omega}\;\mathbf{e}\{\exp[-i(\omega t - \mathbf{K}\cdot\mathbf{x})] + \exp[+i(\omega t - \mathbf{K}\cdot\mathbf{x})]\}$$

Hence we take the amplitude that an atomic system will absorb a photon to be

$$\sqrt{4\pi\hbar c^2/2\omega}\ \exp\left[-i(\omega t - \mathbf{K}\cdot\mathbf{x})\right]$$

For emission the vector potential is the same except for a positive exponential.

Example: Suppose an atom is in an excited state ψ_i with energy E_i and makes a transition to a final state ψ_f with energy E_f.

The probability of transition per second is the same as the probability of transition under the influence of a vector potential a**e** exp[+i(ωt - **K**·**X**)] representing the emitted photon.

According to the laws of quantum mechanics (Fermi's golden rule)

$$\text{Trans. prob./sec} = 2\pi/\hbar\,|_f(\text{potential})_i\,|^2 \cdot (\text{density of states})$$

$$\text{Density of states} = \frac{K^2\,dK\ d\Omega}{(2\pi)^3\,d(\omega\hbar)} = \frac{\omega^2\,d\Omega}{(2\pi c)^3\,\hbar}$$

The matrix element $U_{fi} = |_f(\text{potential})_i|^2$ is to be computed from perturbation theory.

This is explained in more detail in the next Part.

First, however, we shall note that more than one choice for the potential may give the same physical results. (This is to justify the possibility of always choosing $\phi = 0$ for our photon.)

Part 3

The representation of the plane-wave photon by the potentials

$$\mathbf{A}(\mathbf{x}, t) = a\mathbf{e} \exp\,[-i(\omega t - \mathbf{K}\cdot\mathbf{x})]$$

$$\phi = 0$$

is essentially a choice of "gauge."

The fact that a freedom of choice exists results from the invariance of the Pauli equation to the quantum-mechanic gauge transform.

The quantum-mechanical transformation is a simple extension of the classical, where, if

$$\mathbf{E} = -\nabla\phi + \partial\phi/\partial t$$

and

$$\mathbf{B} = \nabla \times \mathbf{A}$$

and if χ is any scalar, then the substitutions

$$\mathbf{A}' = \mathbf{A} + \nabla\chi$$

$$\phi' = \phi + \partial\chi/\partial t$$

leave **E** and **B** invariant.

In quantum mechanics the additional transformation o f the wave function

$$\Psi' = e^{-i\chi}\Psi$$

is introduced.

The invariance of the Pauli equation is shown as follows.

The Pauli equation is

$$-\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} = \frac{1}{2m}\left[\sigma\cdot\left(\mathbf{p}-\frac{e}{c}\mathbf{A}\right)\right]\left[\sigma\cdot\left(\mathbf{p}-\frac{e}{c}\mathbf{A}\right)\right]\Psi + e\phi\Psi$$

Then, since

$$\frac{\partial}{\partial x}\Psi' = \frac{\partial}{\partial x}e^{-i\chi}\Psi = e^{-i\chi}\frac{\partial\Psi}{\partial x} - i\frac{\partial\chi}{\partial x}\Psi e^{-i\chi}$$

$$\mathbf{p}(e^{-i\chi}\Psi) = e^{-i\chi}(\mathbf{p}-\nabla\chi)\Psi$$

and

$$\left(\mathbf{p}-\frac{e}{c}\mathbf{A}\right)e^{-i\chi}\Psi = e^{-i\chi}\left(\mathbf{p}-\nabla\chi-\frac{e}{c}\mathbf{A}\right)\Psi$$

The partial derivative with respect to time introduces a term $(\partial\chi/\mathrm{at})\psi e^{-i\chi}$, and this may be included with $\phi e^{-i\chi}\psi$.

Therefore the substitutions

$$\Psi' = e^{-i\chi}\Psi$$

$$\mathbf{A}' = \mathbf{A} + \frac{e}{c}\nabla\chi$$

$$\phi' = \phi + (\partial\chi/\partial t)$$

leave the Pauli equation unchanged.

The vector potential **A** as defined for a photon enters the Pauli Hamiltonian as a perturbation potential for a transition from state i to state f.

Any time-dependent perturbation which can be written

$$\Delta H = e^{i\omega t}\, U(x,y,z)$$

results in the matrix element U_{fi} given by

$$U_{fi} = \int \Psi_f^* \Delta H \Psi_i \, d\,vol$$

$$= \int \phi_f^*(\mathbf{x}) \exp[i(E_f/\hbar)t]\, e^{i\omega t}\, U(\mathbf{x}) \exp[-i(E_f/\hbar)t]\, \phi_i(\mathbf{x}) d\,vol$$

This expression indicates that the perturbation has the same effect as a time-independent perturbation U(x,y,z) between initial and final states whose energies are, respectively, E_i-$\omega\hbar$ and E_f.

As is well known the most important contribution will come from the states such that $E_f = E_i - \omega\hbar$.

Using the previous results, the probability of a transition per second is

$$P_{fi}\, d\Omega = \frac{2\pi}{\hbar} |U_{fi}|^2 \frac{\omega^2\, d\Omega}{(2\pi)^3}$$

To determine U_{fi} write

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 - \frac{e\hbar}{2mc}(\boldsymbol{\sigma}\cdot\nabla\times\mathbf{A}) + eV$$

$$= \frac{1}{2m}\mathbf{p}\cdot\mathbf{p} + eV - \frac{e}{2mc}(\mathbf{p}\cdot\mathbf{A} + \mathbf{A}\cdot\mathbf{p}) - \frac{e\hbar}{2mc}(\boldsymbol{\sigma}\cdot\nabla\times\mathbf{A})$$

$$+ \frac{e^2}{2mc^2}\mathbf{A}\cdot\mathbf{A}$$

Because of the rule that the potential acts only once, which is the same as requiring only first-order terms to enter, the term in $\mathbf{A}\cdot\mathbf{A}$ does not enter this problem.

Making use of $\mathbf{A} = a\mathbf{e}\exp[-i(\omega t - \mathbf{K}\cdot\mathbf{x})]$ and the two operator relations

$$(1) \qquad \nabla\times\mathbf{A} = \mathbf{K}\times\mathbf{e}\, e^{+i\mathbf{K}\cdot\mathbf{x}}\, e^{i\omega t}$$

$$(2) \qquad \mathbf{p}\, e^{+i\mathbf{K}\cdot\mathbf{x}} = e^{+i\mathbf{K}\cdot\mathbf{x}}(\mathbf{p} - \hbar\mathbf{K})$$

or

$$\mathbf{p}\cdot\mathbf{e}\, e^{+i\mathbf{K}\cdot\mathbf{x}} = e^{+i\mathbf{K}\cdot\mathbf{x}}(\mathbf{p}\cdot\mathbf{e} - \hbar\mathbf{K}\cdot\mathbf{e})$$

where $\mathbf{K}\cdot\mathbf{e} = 0$ (which follows from the choice of gauge and the Maxwell equations), we may write

$$U_{fi} = a\int \phi_f^*\left[-(e/2mc)(\mathbf{p}\cdot\mathbf{e}\, e^{+i\mathbf{K}\cdot\mathbf{x}} + e^{+i\mathbf{K}\cdot\mathbf{x}}\,\mathbf{e}\cdot\mathbf{p})\right.$$

$$\left. + (e\hbar i/2mc)\,\boldsymbol{\sigma}\cdot(\mathbf{K}\times\mathbf{e})\, e^{+i\mathbf{K}\cdot\mathbf{x}}\right]\phi_i \; d\;\mathrm{vol}$$

This result is exact.

It can be simplified by using the so-called “dipole" approximation.

To derive this approximation consider the term $(e/2mc)(\mathbf{p}\cdot\mathbf{e}\, e^{+i\,\mathbf{K}\cdot\mathbf{x}})$, which is the order of the velocity of an electron in the atom, or the current.

The exponent can be expanded.

$$e^{+i\mathbf{K}\cdot\mathbf{x}} = 1 + i\mathbf{K}\cdot\mathbf{x} + 1/2(i\mathbf{K}\cdot\mathbf{x})^2 + \cdots$$

$\mathbf{K}\cdot\mathbf{x}$ is of the order a_0/λ, where a_0 = dimension of the atom and λ = wavelength.

If $a_0/\lambda << 1$, all terms of higher order than the first in a_0/λ may be neglected.

To complete the dipole approximation, it is also necessary to neglect the last term.

This is easily done since the last term may be taken as the order of $(\hbar K/mc) = (\hbar Kc/mc^2) = (mv^2/2mc^2)$.

Although such a term is negligible even this is an overestimate.

More correctly,

$$(e\hbar i/2mc)\boldsymbol{\sigma}\cdot(\mathbf{K}\times\mathbf{e})\, e^{+i\mathbf{K}\cdot\mathbf{x}} \approx v/c \times [\text{matrix element of } \boldsymbol{\sigma}\cdot(\mathbf{K}\times\mathbf{p})]$$

The matrix element is

$$\int \phi_f{}^*\,\boldsymbol{\sigma}\cdot(\mathbf{K}\times\mathbf{p})\,\phi_i \;\mathrm{d\,vol}$$

A good approximation allows the separation

$$\phi_f^* = \phi_f^*(\mathbf{x})\,U_f^*\,(\text{spin})$$

and

$$\phi_i = \phi_i\,(\mathbf{x})\,U_i^*(\text{spin})$$

Then to the accuracy of this approximation the integral is

$$\int \phi_f^*(\mathbf{x})\,\phi_i\,(\mathbf{x})\,U_f^*(\sigma \cdot (\mathbf{K} \times \mathbf{p}))U_i \text{ d vol} = 0$$

since the states are orthogonal.

For the present, the dipole approximation is to be used.

Then

$$U_{fi} = -a\,\frac{e}{c}\,\frac{\mathbf{p}_{fi} \cdot \mathbf{e}}{m}$$

where

$$\mathbf{p}_{fi} \cdot \mathbf{e} = \int \phi_f^*(\mathbf{p} \cdot \mathbf{e})\phi_i = \mathbf{e} \cdot \int \phi_f^* \mathbf{p} \phi_i \text{ d vol}$$

So

$$P_{fi} = \frac{2\pi}{\hbar}\left[\frac{e}{mc}\,a\right]^2 (\mathbf{p}_{fi} \cdot \mathbf{e})^2\,d\Omega\,\frac{\omega^2}{(2\pi)^3}$$

Using operator algebra, $\mathbf{p}_{fi}/m = \hbar\omega_{fi}\,\mathbf{x}_{fi}$, so that

$$P_{fi}\,d\Omega = a^2\,[e^2\omega^4/(2\pi)^2](\mathbf{e} \cdot \mathbf{x}_{fi})^2\,d\Omega$$

where $\mathbf{x}_{fi} = \int \phi_f^*\,\mathbf{x}\,\phi_i$ d vol.

The total probability is obtained by integrating P_{fi}, over $d\Omega$, thus

$$\text{Total prob./sec} = \int a^2 \frac{e^2\omega^4}{(2\pi)^2} (\mathbf{e}\cdot\mathbf{x}_{fi})^2 \, d\Omega$$

$$= a^2 \frac{e^2\omega^4}{2\pi} \int_0^{\pi} |\mathbf{x}_{fi}|^2 \sin^3\theta \, d\theta$$

$$= a^2 \, 4e^2\omega^4 |\mathbf{x}_{fi}|^2/6\pi$$

The term $\mathbf{e}\cdot\mathbf{x}_{fi}$, is resolved by noting (Fig. 3-1)

$$|\mathbf{x}_{fi} \cdot \mathbf{e}| = |\mathbf{x}_{fi}| \sin\theta$$

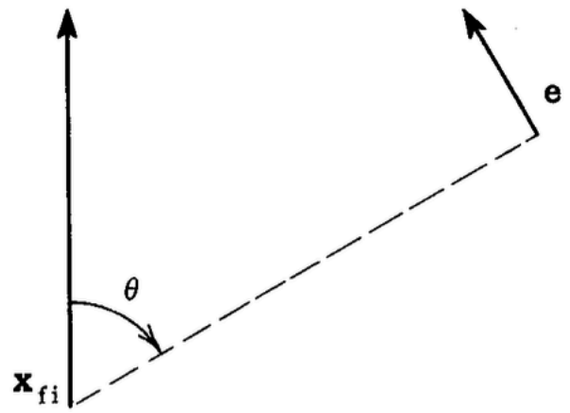


FIG. 3-1

Substituting for a^2

$$\text{Total prob./sec} = \frac{4}{3} \frac{e^2}{\hbar c} \frac{\omega^3}{c^2} |\mathbf{x}_{fi}|^2$$

Part 4

Absorption of Light. The amplitude to go from state k to state *l* in time T (Fig. 4-1) is given from perturbation theory by

$$a_{lk} = -(i/\hbar) \int_0^T \exp\left(\frac{i}{\hbar} E_l t\right) U_{lk}(t) \exp\left(-\frac{i}{\hbar} E_k t\right) dt$$

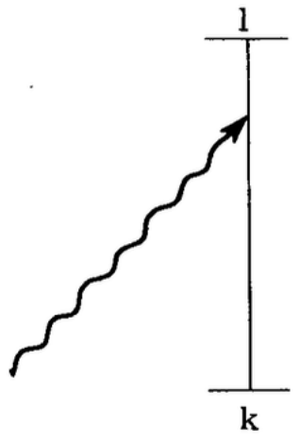


FIG. 4-1

where the time dependence of $U_{kl}(t)$ is indicated by writing

$$U_{lk}(t) = u_{lk}\, e^{-i\omega t}$$

(In accord with the rules of Part 2, the argument of the exponential is minus and only terms which are linear in the potential are included.)

Using this time dependence and performing the integration,

$$a_{lk} = -\frac{\exp\left[\frac{i}{\hbar}(E_l - \hbar\omega - E_k)\right] - 1}{E_l - \hbar\omega - E_k} u_{lk}$$

the transition probability is given b y

$$|a_{lk}|^2 = \frac{4 \sin^2(\Delta T/2\hbar)}{\Delta^2} |u_{lk}|^2 \qquad \Delta = E_l - E_k - \hbar\omega$$

$$|a_{lk}|^2 = \frac{4 \sin^2 (\Delta T/2\hbar)}{\Delta^2} |u_{lk}|^2 \qquad \Delta = E_l - E_k - \hbar\omega$$

This is the probability that a photon o f frequency ω traveling in direction (θ, ϕ) will b e absorbed.

The dependence on the photon direction is contained in the matrix element u_{lk}.

For example, see Eq. (4-1) for the directional dependence in the dipole approximation.

If the incident radiation contains a range of frequencies and directions, that is, suppose

$$\mathbf{P}(\omega,\theta,\phi)d\omega\ d\Omega = \left\{\begin{array}{l}\text{probability that a photon is present with fre-}\\ \text{quency } \omega \text{ to } \omega + d\omega \text{ and in solid angle } d\Omega\\ \text{about the direction } (\theta,\phi)\end{array}\right\}$$

and the probability of absorption of any photon traveling in the (θ,ϕ) direction is desired, it is necessary to integrate over all frequencies.

This absorption probability is

$$\int_0^\infty \frac{4 \sin^2 (\Delta T/2\hbar)}{\Delta^2} |u_{lk}|^2 P(\omega,\theta,\phi)\, d\omega\ d\Omega$$

when T is large, the factor $(\Delta)^{-2} \sin'(\Delta T/2\hbar)$ has an appreciable value only for $\hbar\omega$ near E_l-E_k, and $P(\omega, \theta, \phi)$ will be substantially constant over the small range in ω which contributes to the integral so that it may be taken out of the integral.

Similarly for u_k, so that

$$\text{Trans. prob.} = 2\pi(\hbar)^{-1}|u_{1k}|^2\, P(\omega_{1k},\theta,\phi)d\Omega \tag{4.1}$$

where

$$\hbar\omega_{1k} = (E_1 - E_k)$$

This can also be written in terms of the incident intensity (energy crossing a unit area in unit time by noting that

$$\text{Intensity} = i(\omega,\theta,\phi)d\omega\, d\Omega = \hbar\omega c\; P(\omega,\theta,\phi)\; d\omega\; d\Omega$$

Thus

$$\text{Trans. prob.} = 2\pi(\hbar)^{-1}|u_{1k}|^2\,(\hbar\omega_{1k}c)^{-1}\, i(\omega_{1k},\theta,\phi)d\Omega \tag{4.2}$$

Using the dipole approximation, in which

$$u_{1k} = \sqrt{2\pi/\omega_{1k}}\;(e/mc)(\mathbf{p}_{1k}\cdot\mathbf{e})$$

$$= \sqrt{2\pi/\omega_{1k}}\;(e/c)\,\hbar\omega_{1k}\,(\mathbf{x}_{1k}\cdot\mathbf{e})$$

the total probability o f absorption (per second) is

$$4\pi^2e^2(\hbar c)^{-1}(\mathbf{x}_{1k}\cdot\mathbf{e})^2\, i(\omega_{1k},\theta,\phi)\, d\Omega \tag{4.3}$$

It is evident that there is a relation between the probability of spontaneous emission, with accompanying atomic transition from state *l* to state k,

$$\left\{\begin{matrix}\text{Probability of spontaneous}\\ \text{emission/sec}\end{matrix}\right\} = 2\pi(\hbar)^{-1}(2\pi c)^{-3}|u_{k1}|^2\,\omega_{1k}{}^2\, d\Omega$$

and the absorption of a photon with accompanying atomic transition from state k to state l, Eq. (4-1), although the initial and final states are reversed since $|u_{lk}| = |u_{kl}|$.

This relation may be stated most simply in terms of the concept of the probability $n(\omega, \theta, \phi)$ that a particular photon state is occupied.

Since there are $(2\pi c)^{-3}\omega^2 d\omega\, d\Omega$ photon states in frequency range $d\omega$ and solid angle $d\Omega$, the probability that there is some photon within this range is

$$P(\omega,\theta,\phi)\, d\omega\, d\Omega = n(\omega,\theta,\phi)(2\omega c)^{-3}\omega^2\, d\omega\, d\Omega$$

Expressing the probability of absorption in terms of $n(\omega,\theta,\phi)$,

$$\text{Trans. prob./sec} = 2\pi(\hbar)^{-1}|u_{lk}|^2\, n(\omega,\theta,\phi)(2\pi c)^{-3}\omega_{kl}^2\, d\Omega \qquad (4.4)$$

This equation may be interpreted as follows.

Since $n(\omega, \theta, \phi)$ is the probability that a photon state is occupied, the remainder of the terms of the right-hand side must be the probability per second that a photon in that state will be absorbed.

Comparing Eq. (4-4) with the rate of spontaneous emission shows that

$$\left\{\begin{array}{l}\text{Prob./sec of absorption}\\ \text{of a photon from a state}\\ \text{(per photon in that state)}\end{array}\right\} = \left\{\begin{array}{l}\text{prob./sec of spontaneous}\\ \text{emission of a photon into}\\ \text{that state}\end{array}\right\}$$

In what follows, it will be shown that Eq. (4-4) is correct even when there is a possibility of more than one photon per state provided $n(\omega, \theta, \phi)$ is taken as the mean number of photons per state.

If the initial state consists of two photons per state, it will not be possible to distinguish them and the statistical weight of the initial state will b e 1/2!

However, the amplitude for absorption will be twice that for one photon.

Taking the statistical weight times the square of the amplitude for this process, the transition probability per second is found to be twice that for only one photon per photon state.

When there are three photons per initial photon state and one is absorbed, the following six processes (shown on Fig. 4-2) can occur.

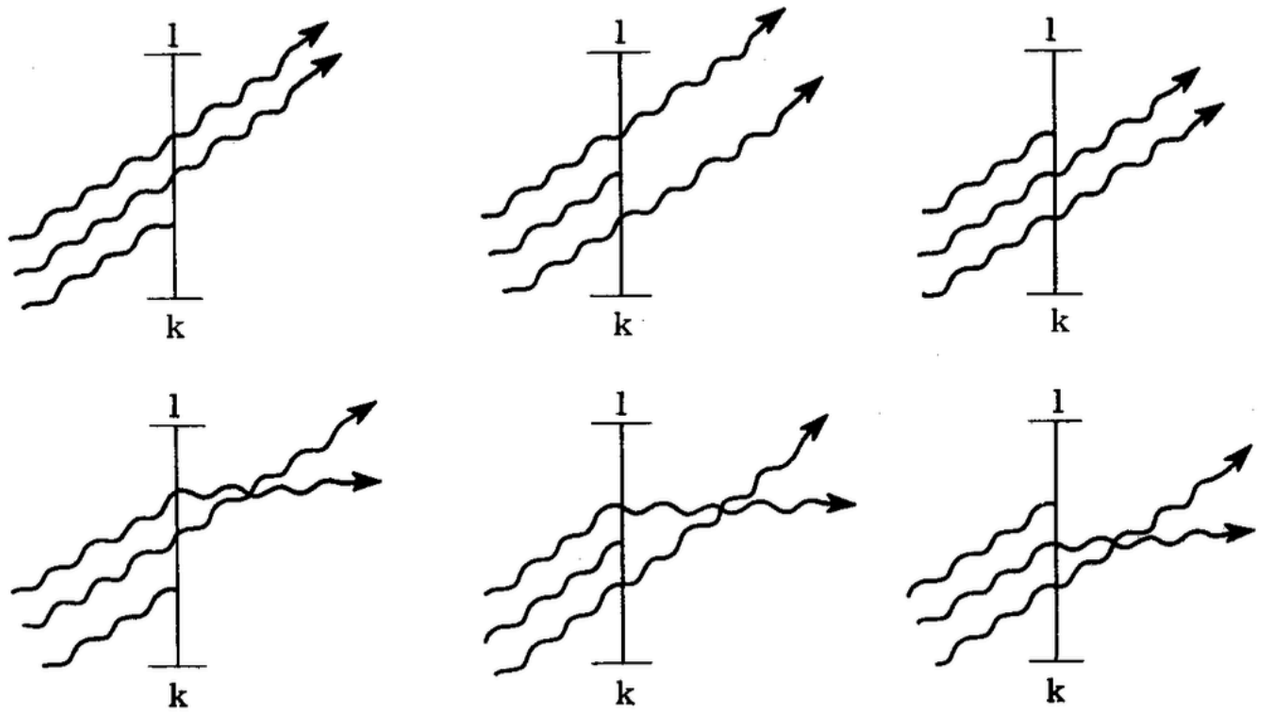


FIG. 4-2

Any of the three incident photons may be absorbed and, in addition, there is the possibility that the photons which are not absorbed may be interchanged.

The statistical weight of the initial state is 1/3!, the statistical weight of the final state is 1/2!, and the amplitude for the process is 1/3!.

Thus the transition probability is $(1/3!)(1/2!)(6)^2 = 3$ times that if there were one photon per initial state.

In general, the transition probability for n photons per initial photon state is n times that for a single photon per photon state, so Eq. (4-4) is correct if $n(\omega, \theta, \phi)$ is taken as the mean number of photons per state.

A transition that results in the emission o f a photon may be induced by incident radiation.

Such a process (involving one incident photon) could be indicated diagrammatically, as in Fig. 4-3.

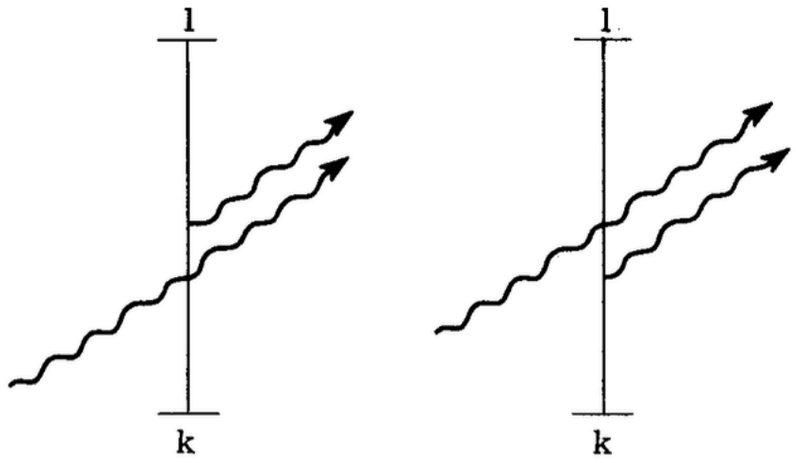


FIG. 4-3

One photon is incident on the atom and two indistinguishable photons come off.

The statistical weight o f the final state is 1/2! and the amplitude for the process is 2, so the probability of emission for this process is twice that of spontaneous emission.

For n incident photons the statistical weight of the initial state is 1/n!, the statistical weight of the final state is 1/(n + 1)!, and the amplitude for the process is (n + 1)! times the amplitude for spontaneous emission.

The probability (per second) of emission is then n+ 1 times the probability of spontaneous emission.

The n can be said to account for the induced part of the transition rate, while the 1 is the spontaneous part of the transition rate.

Since the potentials used in computing the transition probability have been normalized to one photon per cubic centimeter and the transition probability depends on the square of the amplitude of the potential, it is clear that when there are n photons per photon state the correct transition probability for absorption would be obtained by normalizing the potentials to n photons per cubic centimeter (amplitude times as large).

This is the basis for the validity of the so-called semiclassical theory of radiation.

In that theory absorption is calculated as resulting from the perturbation by a potential normalized to the actual energy in the field, that is, to energy $n\hbar\omega$ if there are n photons.

The correct transition probability for emission is not obtained this way, however, because it is proportional to n+1.

The error corresponds to omitting the spontaneous part of the transition probability.

In the semiclassical theory of radiation, the spontaneous part of the emission probability is arrived at by general arguments, including the fact that its inclusion leads to the observed Planck distribution formula.

Einstein first deduced these relationships by semiclassical reasoning.

Part 5

Selection Rules in the Dipole Approximation. In the dipole approximation the appropriate matrix element is

$$\mathbf{x}_{if} = \int \Psi_f^* \, \mathbf{x} \Psi_i \; \mathrm{d\ vol}$$

The components of of $\mathbf{x}_{if}$ are x_{if}, y_{if}, z_{if} and

$$\text{Trans. prob.} \approx |x_{if}|^2 + |y_{if}|^2 + |z_{if}|^2$$

Selection rules are determined by the conditions that cause this matrix element to vanish.

For example, if in hydrogen the initial and final states are S states (spherically symmetrical), $\mathbf{X}_{if} = 0$ and transitions between these states are "forbidden".

For transitions from P to S states, however, $\mathbf{x}_{if} \neq 0$ and they are "allowed".

In general, for single electron transitions, the selection rule is

$$\Delta L = \pm 1$$

This may be seen from the fact that the coordinates x, y , and z are essentially the Legendre polynomial P_1.

If the orbital angular momentum of the initial state is n, the wave function contains P_n.

But

$$P_1 P_n = [1/(2n+1)]\,[n P_{n-1} + (n+1) P_{n+1}]$$

Hence for the matrix element not to vanish, the angular momentum of the final state must be n+1, so that its wave function will contain either P_{n+1} or P_{n-1}.

For a complex atom (more than one electron), the Hamiltonian is

$$H = \sum_{\alpha} (1/2m)[\mathbf{P}_{\alpha} - (e/c)A(\mathbf{x}_{\alpha})]^2 + \text{Coulomb terms}$$

The transition probability is proportional to $|P_{mn}|^2 = |\sum_\alpha (\mathbf{P}_\alpha)_{mn}|^2$, where the sum is over all the electrons of the atom.

As has been shown $(\mathbf{P}_\alpha)_{mn}$ is the same, up to a constant, as $(\mathbf{x}_\alpha)_{mn}$, and the transition probability is proportional to

$$|\mathbf{x}_{mn}|^2 = \left|\sum_\alpha (\mathbf{x}_\alpha)_{mn}\right|^2$$

In particular, for two electrons the matrix element is

$$\int \Psi_f^*(\mathbf{x}_1,\mathbf{x}_2)(\mathbf{x}_1 + \mathbf{x}_2)\,\Psi_1(\mathbf{x}_1,\mathbf{x}_2)\,d\mathbf{x}_1\,d\mathbf{x}_2$$

$\mathbf{x}_1 + \mathbf{x}_2$ behaves under rotation o f coordinates similarly to the wave function of some "object" with unit angular momentum.

If the "object" and the atom in the initial state do not interact, then the product $(\mathbf{x}_1 + \mathbf{x}_2)\,\psi_1(\mathbf{x}_1,\mathbf{x}_2)$ can be formally regarded as the wave function of a system (atom + object) having possible values of $J_1 + 1$, J_1, and $J_1 - 1$ for total angular momentum.

Therefore the matrix element is nonzero only if J_f, the final angular momentum, has one of the three values $J_1 \pm 1$ or J_1.

Hence the general selection rule $\Delta J = \pm 1, 0$.

Parity. Parity is the property of a wave function referring to its behavior upon reflection of all coordinates.

That is, if

$$\Psi(-\mathbf{x}_1,\ -\mathbf{x}_2,\ \ldots) = +\Psi(\mathbf{x}_1,\mathbf{x}_2,\ \ldots)$$

parity is even; or if

$$\Psi(-\mathbf{x}_1,\ -\mathbf{x}_2,\ \ldots) = -\Psi(\mathbf{x}_1,\mathbf{x}_2,\ \ldots)$$

parity is odd.

If in the matrix elements involved in the dipole approximation one makes the change of variable of integration $\mathbf{x} = -\mathbf{x}'$, the result is

$$\mathbf{x}_{if} = \int \Psi_f^*(\mathbf{x})\,\mathbf{x}\,\Psi_i(\mathbf{x})\ d^3\mathbf{x} = \int \Psi_f^*(-\mathbf{x}')(-\mathbf{x}')\,\Psi_0(-\mathbf{x}')\ d^3\mathbf{x}'$$

If the parity of ψ_f is the same as that of ψ_i, it follows that

$$\mathbf{x}_{if} = -\mathbf{x}_{if} = 0$$

Hence the rule that parity must change in allowed transitions. For a one-electron atom, L determines the parity; therefore, $\Delta L = 0$ would be forbidden. In many-electron atoms, L does not determine the parity (determined by algebraic, not vector, sum of individual electron angular momenta), so $\Delta L = 0$ transitions can occur.

The $0 \rightarrow 0$transitions are always forbidden, however, since a photon always carries one unit of angular momentum.

All wave functions have either even or odd parity.

This can be seen from the fact that the Hamiltonian (in the absence of an external magnetic field) is invariant under the parity operation.

Then, if $H\psi(\mathbf{x}) = EX(\mathbf{x})$, it is also true that $H\psi(-\mathbf{x}) = EX(-\mathbf{x})$.

Therefore, if the state is nondegenerate, it follows that either $\psi(-\mathbf{x}) = \psi(\mathbf{x})$ or $\psi(-\mathbf{x}) = -\psi(\mathbf{x})$.

If the state is degenerate, it is possible that $\psi(-\mathbf{x}) \neq \psi(\mathbf{x})$.

But then a complete solution would be one of the linear combinations

$$\Psi(\mathbf{x}) + \Psi(-\mathbf{x}) \qquad \text{even parity}$$

$$\Psi(\mathbf{x}) - \Psi(-\mathbf{x}) \qquad \text{odd parity}$$

Forbidden Lines. Forbidden spectral lines may appear in gases it they are sufficiently rarefied.

That is, forbiddenness is not absolute in all cases.

It may simply mean that the lifetime of the state is much longer than if it were allowed, but not infinite.

Thus, if the collision rate is small enough (collisions of the second kind ordinarily cause de-excitation in forbidden cases), the forbidden transition may have sufficient time to occur.

In the nearly exact matrix element

$$\int \Psi_f^* (\mathbf{e} \cdot \mathbf{p})\, e^{-i\mathbf{K}\cdot\mathbf{x}}\, \Psi_i \, d^3\mathbf{x}$$

the dipole approximation replaces $e^{-i\mathbf{K}\cdot\mathbf{x}}$ by 1 .

If this vanishes, the transition is forbidden, as described in the foregoing.

The next higher or quadrupole approximation would then be to replace $e^{-i\mathbf{K}\cdot\mathbf{x}}$ by $1 - i/\mathbf{K}\cdot\mathbf{x}$, giving the matrix element

$$-i \int \Psi_f^* (\mathbf{e} \cdot \mathbf{p})(\mathbf{K} \cdot \mathbf{x})\, \Psi_i \, d^3\mathbf{x}$$

For light moving in the z direction and polarized in the x direction, this becomes

$$-iK \int \Psi_f^*(p_x z)\, \Psi_i\, d^3\mathbf{x} = -iK|_f(p_x z)_i|^2$$

and the transition probability is proportional to

$$(K)^2 |_f(p_x z)_i|^2$$

whereas in the dipole approximation it was proportional to

$$|_f(p_x)_i|^2$$

therefore the transition probability in the quadrupole approximation is at least of the order of $(Ka)^2 = a^2/\lambda$, smaller than in the dipole approximation, where a is of the order of the size of the atom, and λ the wavelength emitted.

Problem: Show that

$$H(xz) - (xz)H = (\hbar/mi)(p_x z + x p_z)$$

and consequently that

$$[(\hbar/mi)(p_x z + x p_z)]_{mn} = (xz)_{mn}(E_m - E_n)$$

note that $p_x z$ can be written as the sum

$$p_x z = 1/2(p_x z + x p_z) + 1/2(p_x z - x p_z)$$

From the preceding problem, the first part of $p_x z$ is seen to be equivalent, up to a constant, to xz, which behaves similarly to a wave function for angular momentum 2, even parity.

The second part is the angular momentum operator L_y, which behaves like a wave function for angular momentum 1, even parity.

Therefore the selection rules corresponding to the first part are seen to be $\Delta J = \pm 2, +1, 0$ with no parity change.

This type of radiation is called electric quadrupole.

The selection rules for the second part of $p_x z$ are $\Delta J = \pm 1, 0$, no parity change, and the corresponding radiation is called magnetic dipole.

Note that unless $\Delta J = \pm 2$, the two types of radiation cannot be distinguished by the change in angular momentum or parity.

If $\Delta J = \pm 1, 0$, they can only be distinguished by the polarization of the radiation.

Both types may occur simultaneously, producing interference.

In the case of electric quadrupole radiation, it is implicit in the rules that $1/2 \rightarrow 1/2$ and $0 \rightarrow 1$ transitions are forbidden (even though ΔJ may be ± 1), since the required change of 2 for the vector angular momentum is impossible in these cases.

Continuing to higher approximations, it is possible by similar reasoning to deduce the vector change in angular momentum, or angular momentum of the photon, and the selection rules for parity change and change of total angular momentum ΔJ associated with the various multipole orders (Table 5-1).

Actually all the implicit selection rules for ΔJ, which become numerous for the higher multipole orders, can be expressed explicitly by writing the selection rule as

$$|J_f - J_i| \leq l \leq J_f + J_i$$

where 2^l is the multipole order or l is the vector change in angular momentum.

TABLE 5-1. Classification of Transitions and Their Selection Rules

	Multipole	Electric dipole	Magnetic dipole	Electric quadrupole	Magnetic quadrupole	Electric octupole
Character of photon	Angular momentum	1	1	2	2	3
	Parity	Odd	Even	Even	Odd	Odd
Selection rule for emitting system	Parity change	Yes	No	No	Yes	Yes
	Change of total angular momentum ΔJ	$\pm 1, 0$	$\pm 1, 0$	$\pm 2, \pm 1, 0$	$\pm 2, \pm 1, 0$	$\pm 3, \pm 2, \pm 1, 0$
		No $0 \rightarrow 0$	No $0 \rightarrow 0$	No $0 \rightarrow 0$	No $0 \rightarrow 0$	No $0 \rightarrow 0$
				$\frac{1}{2} \rightarrow \frac{1}{2}$	$\frac{1}{2} \rightarrow \frac{1}{2}$	$\frac{1}{2} \rightarrow \frac{1}{2}$
				$0 \rightarrow 1$	$0 \rightarrow 1$	etc. (see following)

It turns out that in so-called parity-favored transitions, wherein the product of the initial and final parities is $(-1)^{J_f - J_i}$ and the lowest possible multipole order is $J_f - J_i$, the transition probabilities for multipole types contained within the dashed vertical lines in Table 5-1 are roughly equal.

In parity-unfavored transitions, where the parity product is $(-1)^{J_f - J_i + 1}$ and the lowest multipole order is $|J_f - J_i| + 1$, this may not be true.

Part 6

Equilibrium of Radiation. If a system is in equilibrium, the relative number of atoms per cubic centimeter in two states, say *l* and k, is given by

$$N_l / N_k = e^{-(E_l - E_k)/kT} = e^{-h\omega/kT}$$

according to statistical mechanics, when the energies differ by $\hbar\omega$.

Since the system is in equilibrium, the number of atoms going from state k to l per unit time by absorption of photons $\hbar\omega$ must equal the number going from l to k by emission.

If n_ω photons of frequency ω are present per cubic centimeter, these probabilities of absorption are proportional to n_ω and probability of emission is proportional to $n_\omega+1$.

Thus

$$N_k n_\omega = N_l (n_\omega + 1)$$

or

$$(n_\omega + 1)/n_\omega = N_k/N_l = e^{\hbar\omega/kT}$$

$$n_\omega = 1/(e^{\hbar\omega/kT} - 1)$$

This is the Planck black-body distribution law.

The Scattering of Light. We discuss here the phenomena of an incident photon being scattered by an atom into a new direction (and possibly energy)(see Fig. 6-1).

This may be considered as the absorption of the incoming photon and the emission of a new photon by the atom.

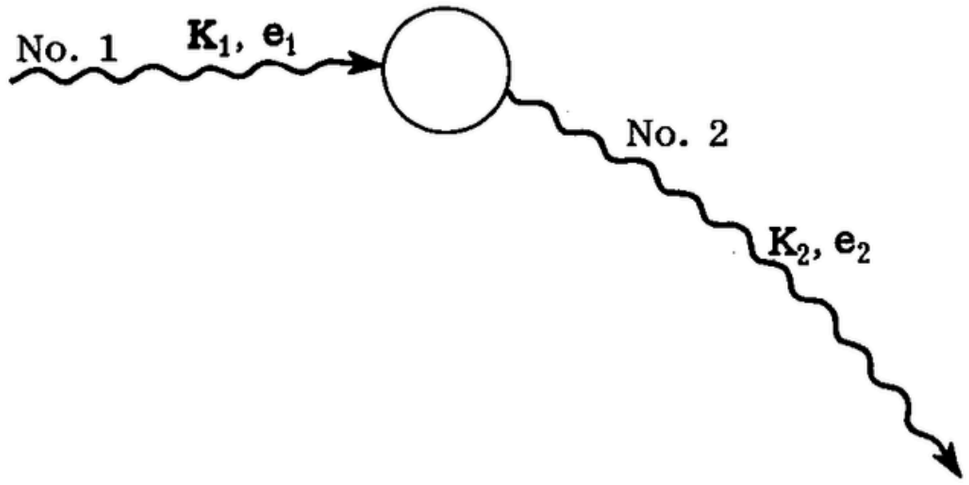


FIG. 6-1

The two photons taking part in the phenomenon are represented by the vector potentials.

$$\mathbf{A}_1 = (2\pi/\omega_1)^{1/2}\, \mathbf{e}_1\, e^{+i(\omega_1 t - \mathbf{K}\cdot\mathbf{x})}$$

$$\mathbf{A}_2 = (2\pi/\omega_2)^{1/2}\, \mathbf{e}_2 e^{-i(\omega_2 t - \mathbf{K}\cdot\mathbf{x})}$$

the number to be determined is the probability that an atom initially in state k will be left in state l by the action of the perturbation $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2$ in the time T.

This probability can be computed just as any transition probability with the use of A_{lk}, where

$$A_{1k} = \delta_{k1} \exp[-i(E_1/\hbar)T] - (i/\hbar)$$

$$\times \int_0^T \exp[-i(E_1/\hbar)(T - t_3)]\, U_{1k}(t_3) \exp[-i(E_k/\hbar)t_3]\, dt_3$$

$$+ \sum_n \int_0^T \int_0^{t_4} \exp[-i(E_1/\hbar)(T - t_4)]$$

$$\times U_{1n}(t_4) \exp[-i(E_n/\hbar)(t_4 - t_3)]\, U_{nk}(t_3) \exp[-i(E_k/\hbar)t_3]\, dt_3\, dt_4$$

The dipole approximation is to be employed and

$$U = \Delta H = (e/2mc)(\mathbf{p} \cdot \mathbf{A}) + (e^2/mc^2)(\mathbf{A} \cdot \mathbf{A})$$

where spins are neglected.

In each integral defining A_{lk}, each of the two vector potentials must appear once and only once.

Thus, in the first integral the term $\mathbf{p} \cdot \mathbf{A}$ of U will not appear in U_{lk}.

The product $\mathbf{A} \cdot \mathbf{A} = (\mathbf{A}_1 + \mathbf{A}_2) \cdot (\mathbf{A}_1 + \mathbf{A}_2)$ will contribute only its cross-product term $2\mathbf{A}_1 \cdot \mathbf{A}_2$.

The second integral will have no contribution from $\mathbf{A} \cdot \mathbf{A}$, but will be the sum of two terms.

The first term contains a U_{ln} based on $\mathbf{p} \cdot \mathbf{A}_2$ and a U_{nk} based on $\mathbf{p} \cdot \mathbf{A}_1$.

The second has U_{ln} based on $\mathbf{p} \cdot \mathbf{A}_1$, and U_{nk} based on $\mathbf{p} \cdot \mathbf{A}_2$.

The time sequences resulting in these two terms can be represented schematically in Fig. 6-2.

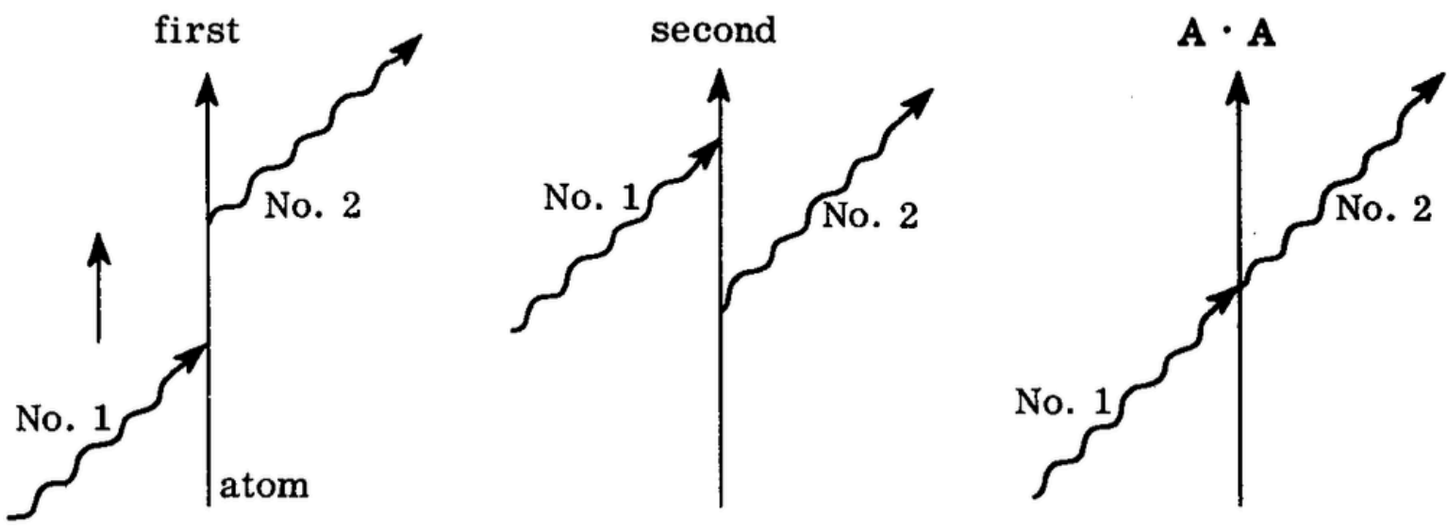


FIG. 6-2

The integral resulting from the first term will now be developed in detail.

$$(\mathbf{p}\cdot\mathbf{A}_1)_{nk} = (2\pi/\omega_1)^{1/2}\,(\mathbf{p}\cdot\mathbf{e}_1)_{nk}\,e^{-i\omega_1 t}$$

$$(\mathbf{p}\cdot\mathbf{A}_2)_{1n} = (2\pi/\omega_2)^{1/2}\,(\mathbf{p}\cdot\mathbf{e}_2)_{1n}\,e^{i\omega_2 t}$$

Then the resulting integral is

$$\sum_n 2\pi/(\omega_1\omega_2)^{1/2}(p\cdot e_2)_{1n}(p\cdot e_1)_{nk}$$

$$\times \int_0^T \int_0^{t_4} \exp[-i(E_1/\hbar)(T-t_4) + i\omega_2 t_4]$$

$$\times \exp[-i(E_n/\hbar)(t_4-t_3) - i\omega_1 t_3]\,\exp[-i(E_k/\hbar)t_3]\,dt_3\,dt_4$$

The integral is similar to the integrals considered previously with regard to transition probabilities, and the sum becomes

$$\sum_n 2\pi/(\omega_1\omega_2)^{1/2}(\mathbf{p}\cdot\mathbf{e}_1)_{1n}(\mathbf{p}\cdot\mathbf{e}_2)_{nk}\,e^{i\phi}$$

$$\times [\sin(T\cdot\Delta/\hbar)/(E_k - E_n + \hbar\omega_1)\cdot\Delta]$$

where $\Delta = (E_1 + \hbar\omega_2 - E_k - \hbar\omega_1)$, and the phase angle ϕ is independent of n.

A term with the denominator given by $(E_n - \hbar\omega_1 - E_k)(E_1 + \hbar\omega_2 - E_n)$ has been neglected, since previous results show that only energies such that $E_1 + \hbar\omega_2 \approx E_k + \hbar\omega_1$ are important.

The final result can be written

$$\text{Trans. prob./sec} = (2\pi/\hbar)|M|^2 [\omega_2{}^2 \, d\Omega_2/(2\pi^3)] = \sigma c \tag{6-1}$$

where |M| is determined from A_{lk} by integrating over ω_2 and averaging over $\mathbf{e}_2$.

Then the complete expression for the cross section σ is

$$\sigma \, d\Omega_2 = \frac{e^4}{m^2c^4}\frac{\omega_2}{\omega_1} \, d\Omega_2 \left| \frac{1}{m}\sum_n \frac{(\mathbf{p}\cdot\mathbf{e}_2)_{ln}(\mathbf{p}\cdot\mathbf{e}_1)_{nk}}{E_k + \hbar\omega_1 - E_n} + \frac{(\mathbf{p}\cdot\mathbf{e}_1)_{ln}(\mathbf{p}\cdot\mathbf{e}_2)_{nk}}{E_k - E_n - \hbar\omega_2} + \frac{1}{mc^2}(\mathbf{e}_1\cdot\mathbf{e}_2)\delta_{lk} \right|^2 \tag{6-2}$$

The first term under the summation comes from the "first term" previously referred to and the second from the "second term."

The last term in the absolute brackets comes from $\mathbf{A}\cdot\mathbf{A}$.

If $l \neq k$, the scattering is incoherent, and the result is called the ‘Raman effect."

If $l = k$, the scattering is coherent.

the the

Further, note that if all the atoms are in the ground state and l = k, then the energy of the atom can only increase and the frequency of the light ω can only decrease.

This gives rise to "Stokes lines".

The opposite effect gives "anti-Stokes lines."

Suppose $\omega_1 = \omega_2$ (coherent scattering) but further hwy is very nearly equal to $E_k - E_n$, where E_n is some possible energy level of the atom.

Then one term in the sum over n becomes extremely large and dominates the remainder.

The result is called "resonance scattering."

If σ is plotted against w, then at such values of ω the cross section has a sharp maximum (see Fig. 6-3).

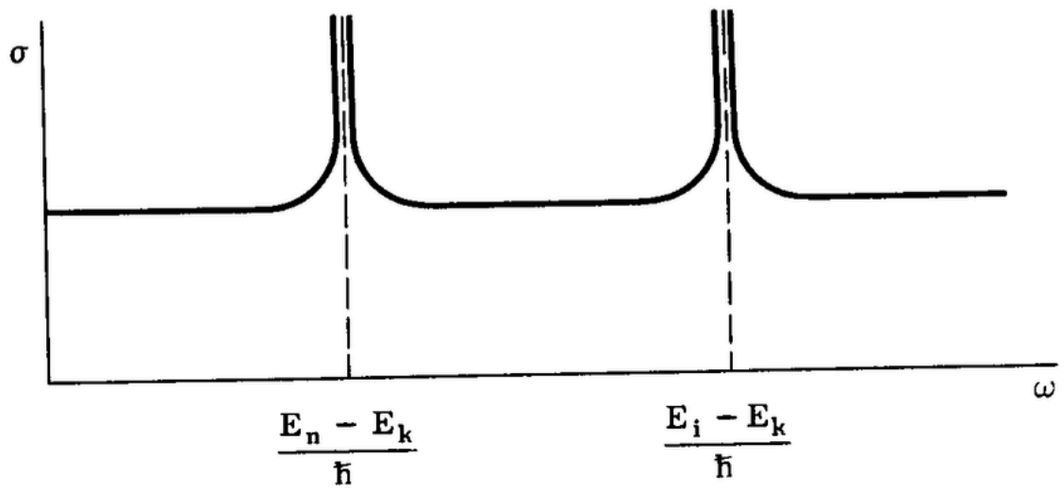


FIG. 6-3

The "index" of refraction of a gas can be obtained by our scattering formula.

It can be obtained, as for other types of scattering, by considering the light scattered in the forward direction.

Self-Energy. Another phenomenon that must be considered in quantum electrodynamics is the possibility of an atom emitting a photon and reabsorbing the same photon.

This affects the diagonal element A_{kk}.

Its effect is equivalent to a shift of energy of the level.

One finds

$$\Delta E = \sum_n \int \frac{(\mathbf{p}\cdot\mathbf{e})_{kn}(\mathbf{p}\cdot\mathbf{e})_{nk}}{E_k - E_n - \omega} \frac{d^3\mathbf{K}}{(2\pi\hbar)^3} \frac{2\pi}{\omega}$$

where **e** is the direction of polarization.

This integral diverges.

A more exact relativistic calculation also gives a divergent integral.

This means that our formulation of electromagnetic effects is not really a completely satisfactory theory.

The modifications required to avoid this difficulty of the infinite self-energy will be discussed later.

The net result is a very small shift ΔE in position of energy levels.

This shift has been observed by Lamb and Rutherford.

Résumé of the Principles and Results of Special Relativity

Part 7

The principle of relativity is the principle that all physical phenomena would appear to be exactly the same if all the objects concerned were moving uniformly together at velocity v; that is, no experiments made entirely inside of a closed spaceship moving uniformly at velocity v (relative to the center of gravity of the matter in the universe, for example) can determine this velocity.

The principle has been verified experimentally.

Newton's laws satisfy this principle; for they are unchanged when subject to a Gallilean transformation,

$$x' = x - vt \qquad y' = y \qquad z' = z \qquad t' = t$$

because they involve only second derivatives.

The Maxwell equations are changed, however, when subjected to this transformation, and early workers this field attempted to make an absolute determination of velocity of the earth using this feature (Michelson-Morley experiment).

Failure to detect any effects of this type ultimately led to Einstein's postulate that the Maxwell equations are of the same form in any coordinate system; and, in particular, that the velocity of light is the same in all coordinate systems.

The transformation between coordinate systems which leaves the Maxwell equations invariant is the Lorentz transformation:

$$x' = \frac{x - vt}{\sqrt{1 - (v^2/c^2)}} = x \cosh u - ct \sinh u$$

$$y' = y$$

$$z' = z$$

$$t' = \frac{t - (xv/c^2)}{\sqrt{1 - (v^2/c^2)}} = -\frac{x}{c}\sinh u + t\cosh u$$

where tanh u = v/c.

Henceforth we shall use time units so that the speed of light c is unity.

The latter form is written to demonstrate the analogy with rotation of axes,

$$x' = x\cos\theta + y\sin\theta$$

$$y' = -x\sin\theta + y\cos\theta$$

Successive transformations v_1 and v_2 or u_1 and u_2 add in the sense that a single transformation v_3 or u_3 will give the same final system if

$$v_3 = v_1 + v_2 \quad \text{or} \quad \tanh u_3 = \tanh (u_1 + u_2)$$

Einstein postulated (theory of special relativity) that the Newton laws must be modified in such a way that they, too, are unchanged in form under a Lorentz transformation.

An interesting consequence of the Lorentz transformation is that clocks appear to run slower in moving systems; that is called time dilation.

In transforming from one coordinate system to another it is convenient to use tensor analysis.

To this end, a four-vector will be defined as a set of four quantities that transforms in the same way as x, y , z and ct.

The subscript μ will be used to designate which of the four components is being considered; for example,

$$x_1 = x \qquad x_2 = y \qquad x_3 = z \qquad x_4 = t$$

The following quantities are four-vectors:

$$-\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}, +\frac{\partial}{\partial t} \qquad (\nabla_\mu) \text{ four-dimensional gradient}$$

$$j_x, j_y, j_z, \rho \qquad (j_\mu) \text{ current (and charge) density}$$

$$A_x, A_y, A_z, \varphi \qquad (A_\mu) \text{ vector (and scalar) potential}$$

$$p_x, p_y, p_z, E \qquad (p_\mu) \text{ momentum and total energy}$$

The energy E, here, is the total energy including the rest energy mc^2.

An invariant is a quantity that does not change under a Lorentz transformation.

If a_μ and b_μ are two four-vectors, the "product"

$$a \cdot b \equiv \sum_n a_\mu b_\mu \equiv a_4 b_4 - a_1 b_1 - a_2 b_2 - a_3 b_3$$

is an invariant.

To avoid writing the summation symbol, the following summation convention will be used.

when the same index occurs twice, sum over it, placing minus in front of first, second, and third components.

The Lorentz invariance of the continuity equation is easily demonstrated by writing it as a "product" of four-vectors ∇_μ and j_μ :

$$\nabla_\mu j_\mu = \nabla_4 j_4 - \nabla_1 j_1 - \nabla_2 j_2 - \nabla_3 j_3 = \frac{\partial \rho}{\partial t} + \frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z}$$

Conservation of charge in all systems if it is conserved in one system is a consequence of the invariance of this "product, " the four-dimensional divergence $\nabla \cdot j$.

Another invariant is

$$p_\mu p_\mu = p \cdot p = E^2 - p_x{}^2 - p_y{}^2 - p_z{}^2 = E^2 - p^2 = m^2$$

E = total energy, m = rest mass, mc^2 = rest energy, p = momentum.

Thus,

$$E^2 = p^2c^2 + m^2c^4$$

It is also interesting to note that the phase of a free particle wave function $\exp[(-i/\hbar)Et - \mathbf{p} \cdot \mathbf{X})]$ is invariant since

$$Et - \mathbf{p} \cdot \mathbf{X} = Et - p_x x - p_y y - p_z z = p_\mu p_\mu$$

The invariance of $p_\mu p_\mu$ can be used to facilitate converting laboratory energies to center-of-mass energies (Fig. 6-4) in the following way (consider identical particles, for simplicity)

FIG. 6-4

$$p_\mu p_\mu = E_{lab} m = E_0^2 + p_0^2$$

but

$$p_0^2 = E_0^2 - m \qquad \text{so} \qquad E_{lab} m = 2E_0^2 - m^2$$

and

$$E_0 = \left[\frac{1}{2} m (E_{lab} + m)\right]^{1/2}$$

The equations of electrodynamics $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{E} = -(1/c)(\partial \mathbf{A}/\partial t) - \nabla\phi$ are easily written in tensor notation

$$B_x = \partial A_z/\partial y - \partial A_y/\partial z = -\nabla_y A_z + \nabla_z A_y \qquad E_x = -\partial A_x/\partial t - \partial \phi/\partial x = -\nabla_t A_x + \nabla_x A_t$$

$$B_y = \partial A_x/\partial z - \partial A_z/\partial x = -\nabla_z A_x + \nabla_x A_z \qquad E_y = -\partial A_y/\partial t - \partial \phi/\partial y = -\nabla_t A_y + \nabla_y A_t$$

$$B_z = \partial A_y/\partial x - \partial A_x/\partial y = -\nabla_x A_y + \nabla_y A_x \qquad E_z = -\partial A_z/\partial t - \partial \phi/\partial z = -\nabla_t A_z + \nabla_z A_t$$

where use is made of the fact that ϕ is the fourth component o f the four-vector potential A_μ.

From the foregoing it can be seen that B_x, B_y, B_z, E_x, E_y and E_z are the components of a second-rank tensor:

$$\mathbf{F}_{\mu\nu} = \nabla_\mu \mathbf{A}_\nu - \nabla_\nu \mathbf{A}_\mu \tag{7-1}$$

This tensor is antisymmetric ($F_{\mu\upsilon} = - F_{\upsilon\mu}$) and the diagonal terms ($\mu = \upsilon$) are zero; thus there are only six independent components (three components of **E** and three components of **B**) instead of sixteen.

$$\mathbf{F}_{\mu\nu} = \begin{vmatrix} 0 & -\mathbf{B}_z & \mathbf{B}_y & \mathbf{E}_x \\ \mathbf{B}_z & 0 & -\mathbf{B}_x & \mathbf{E}_y \\ -\mathbf{B}_y & \mathbf{B}_x & 0 & \mathbf{E}_z \\ -\mathbf{E}_x & -\mathbf{E}_y & -\mathbf{E}_z & 0 \end{vmatrix}$$

The Maxwell equations $\nabla \times \mathbf{B} = 4\mathrm{m}\mathbf{J} + (\partial \mathbf{E}/\partial \mathrm{t})$ and $\nabla \cdot \mathbf{E} = 4\pi\rho$ are written

$$\nabla_\mu \mathbf{F}_{\mu\nu} = 4\pi \mathrm{j}_\nu \tag{7-2}$$

where $\upsilon = 1, 2, 3, 4$, that is, $j_1 = j_x$, $j_2 = J_y$, $j_3 = j_z$, ja = p, and μ is a dummy index of summation.

The $\upsilon = 1, 2,$ and 3 gives the three components of the curl equation, and $\upsilon = 4$ gives the divergence equation.

The equation satisfied by the potential A_μ is found, by substituting Eq.(7-1) into Eq. (7-2), to be

$$\nabla_\mu \nabla_\mu \mathbf{A}_\nu - \nabla_\mu \nabla_\nu \mathbf{A}_\mu = 4\pi \mathrm{j}_\nu$$

The potential A_υ is not unique, however, since the potential

$$A'_\mu = A_\mu + \nabla_\mu \chi \tag{7-3}$$

the χ = any scalar function of position) also satisfies this relation.

Such a change or transformation of potential is called a gauge transformation (for historical reasons).

We shall make the potentials more definite by assuming that all potentials have been transformed so as to satisfy the so-called Lorentz condition

$$\nabla_\mu A_\mu = 0 \tag{7-4}$$

This is convenient, because it simplifies the equation for A_υ to

$$(\nabla \cdot \nabla) A_\nu = 4\pi j_\nu \tag{7-5}$$

since $\nabla \cdot \nabla = \nabla_\mu \nabla_\mu$, which can be recognized as the wave equations

$$\nabla^2 \mathbf{A} - \partial^2 \mathbf{A}/\partial t^2 = -4\pi \mathbf{j}$$
$$\nabla^2 \phi - \partial^2 \phi/\partial t^2 = -4\pi\rho \tag{7-5′}$$

Sometimes Eq. (7-5) is written $\Box^2 A_\mu = -4\pi j_\mu$ ($\Box^2$ = D'Alembertian operator = ∇^2-$(\partial/\partial t)^2$ = $-\nabla \cdot \nabla$).

This choice of gauge ($\nabla_\mu A_\mu = 0$) is the usual one made in classical electrodynamics,

$$\nabla \cdot \mathbf{A} - \partial\phi/\partial t = 0 \tag{7-4′}$$

Part 8

SOLUTION O F THE MAXWELL EQUATION I N EMPTY SPACE

In empty space the plane wave solution of the wave equation

$$\Box^2 A_\mu = -4\pi j_\mu = 0$$

is

$$A_\mu = e_\mu e^{-ik\cdot X}$$

where e_μ and k_μ are constant vectors, and k_μ is subject to the condition

$$k_\mu k_\mu = k \cdot k = 0$$

This may be seen from the fact that ∇_υ operating on $e^{-ik\cdot x}$ has the effect of multiplying b y ik_μ (∇_υ does not operate on e_μ since the coordinates are rectangular).

Thus,

$$-\Box^2 A_\mu = \nabla_\upsilon (\nabla_\upsilon A_\mu) = \nabla_\upsilon (-i\, e_\mu k_\upsilon e^{-ik\cdot X})$$

$$= -e_\mu (k_\upsilon k_\upsilon) e^{-ik\cdot X}$$

Note that in these operations $\nabla_\upsilon A_\mu$ actually forms a second-rank tensor, $\nabla_\upsilon (\nabla_\upsilon A_\mu)$ a third-rank tensor, and then contraction on the index υ yields a first-rank tensor or vector.

The k_μ is the propagation vector with components

$$k_\mu = \omega, K_x, K_y, K_z = \omega, \mathbf{K}$$

so that in ordinary notation

$$\exp(-ik \cdot x) = \exp[-i(\omega t - \mathbf{K} \cdot \mathbf{X})]$$

and the condition $k \cdot k = 0$ means

$$\omega^2 - \mathbf{K} \cdot \mathbf{K} = 0$$

It is straightforward to show that the Lorentz condition

$$\nabla_\mu A_\mu = 0$$

implies that $k \cdot e = 0$.

When working in three dimensions it is customary to take the polarization vector **e** such that $\mathbf{K} \cdot \mathbf{e} = 0$ and to let the scalar potential $\phi = 0$.

But this is not a unique condition; that is, it is not relativistically invariant and will be true only in a one-coordinate system.

This would seem to be a paradox attaching some uniqueness to the system in which $\mathbf{K} \cdot \mathbf{e} = 0$, a situation incompatible with relativity theory.

The "paradox, " however, is resolved by the fact that one can always make a so-called gauge transformation, which leaves the field $F_{\mu\upsilon}$ unaltered but which does change **e** .

Therefore, choosing $\mathbf{K} \cdot \mathbf{e} = 0$ in a particular system amounts to selecting the certain gauge.

The gauge transformation, Eq. (7-3), is

$$\mathbf{A}' = \mathbf{A} + \nabla\chi$$

$$\phi' = \phi + (\partial\chi/\partial t)$$

where χ is a scalar.

But $\nabla \cdot A = 0$, the Lorentz condition, Eq. (7-4), will still hold if

$$\nabla \cdot A' = \nabla \cdot A + \nabla \bullet \chi = 0$$

or if

$$\Box^2 \chi = 0$$

This equation has a solution $\chi = \alpha\, e^{-ik\cdot X}$, so

$$A'_\mu = A_\mu + \nabla_\mu(\alpha\ e^{-ik\cdot X}) = (e_\mu + \alpha k_\mu)e^{-ik\cdot x}$$

where α is an arbitrary constant.

Therefore,

$$e'_\mu = e_\mu + \alpha k_\mu$$

is the new polarization vector obtained by gauge transformation.

In ordinary notation

$$\mathbf{e}' = \mathbf{e} + \alpha \mathbf{K}$$

$$e'_4 = e_4 + \alpha\omega$$

Thus, no matter what coordinate system is used,

$$\mathbf{K} \cdot \mathbf{e}' = \mathbf{K} \cdot \mathbf{e} + \alpha\, \mathbf{K} \cdot \mathbf{K} = \mathbf{K} \cdot \mathbf{e} + \alpha\omega^2$$

can be made to vanish by choice of the constant α.

Clearly the field is left unchanged by a gauge transformation for

$$F'_{\mu\nu} = \nabla_\mu A'_\nu - \nabla_\nu A'_\mu = \nabla_\mu A_\nu + \nabla_\mu \nabla_\nu \chi - \nabla_\nu A_\mu - \nabla_\nu \nabla_\mu \chi = F_{\mu\nu}$$

the $\nabla_\mu \nabla_\upsilon \chi = \nabla_\upsilon \nabla_\mu \chi$ because the order of differentiations is immaterial.

RELATIVISTIC PARTICLE MECHANICS

The components of ordinary velocity do not transform in such a manner that they can be components of a four-vector.

But another quantity

$$dz_\mu/ds = dt/ds,\ dx/ds,\ dy/ds,\ dz/ds$$

where

$$dz_\mu = dt,\ dx,\ dy,\ dz$$

is an element of path of the particle and ds is the proper time defined by

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2$$

is a four-vector and is called the four-velocity u_μ.

Dividing ds^2 by dt^2 gives the relation between proper time and local time to be

$$(ds/dt)^2 = 1 - v^2$$

The components of ordinary velocity are related as follows:

$$dx/ds = (dx/dt)(dt/ds) = v_x/(1 - y^2)^{1/2}$$

$$dy/ds = v_y/(1 - v^2)^{1/2}$$

$$dz/ds = v_z/(1 - v^2)^{1/2}$$

$$dt/ds = 1/(1 - v^2)^{1/2}$$

It is evident that $u_\mu u_\mu = 1$ for

$$u_\mu u_\mu = \frac{1}{1 - v^2} - \frac{v_x^2}{1 - v^2} - \frac{v_y^2}{1 - v^2} - \frac{v_z^2}{1 - v^2} = \frac{1 - v^2}{1 - v^2} = 1$$

The four-momentum is defined

$$p_\mu = mu_\mu = m/(1 - v^2)^{1/2},\ mv_x/(1 - v^2)^{1/2},\ mv_y/(1 - v^2)^{1/2},$$

$$mv_z/(1 - v^2)^{1/2}$$

Note that $p_4 = m/(1 - v^2)^{1/2}$ is the total energy E, so that in ordinary notation the momentum **P** is given by

$$\mathbf{P} = E\mathbf{v}$$

where **v** is the ordinary velocity.

Like the velocity, the components of ordinary force defined by d/dt (momentum) cannot form the components of a four-vector.

But the quantity

$$f_\mu = dp_\mu/ds$$

does form a four-vector with the components

$$f_\mu = d/dt\,(mv_\mu/\sqrt{1 - v^2})\; dt/ds = F_\mu/\sqrt{1 - v^2} \qquad \mu = 1, 2, 3$$

where F_μ is the ordinary force.

The fourth component is

$$f_4 = \frac{\text{power}}{\sqrt{1 - v^2}} = \frac{\text{rate of change of energy}}{\sqrt{1 - v^2}} = \frac{d/dt(m/\sqrt{1 - v^2})}{\sqrt{1 - v^2}}$$

This is seen from the fact that $m/\sqrt{1 - v^2}$ is the total energy and also from the ordinary identity

$$\text{Power} = \mathbf{F} \cdot \mathbf{V} = \left[\frac{d}{dt}\ \frac{m\mathbf{V}}{\sqrt{1 - v^2}}\right] \cdot \mathbf{V}$$

$$= \frac{m}{2}\ \frac{v^2}{(1 - v^2)^{3/2}} + \frac{1}{(1 - v^2)^{1/2}}\ \frac{dv^2}{dt}$$

$$= \frac{mv}{(1 - v^2)^{3/2}}\ \frac{dv}{dt} = \frac{d}{dt}\ \frac{m}{\sqrt{1 - v^2}}$$

Thus the relativistic analogue of the Newton equations is

$$d/ds\ (p_\mu) = f_\mu = m\ d^2z_\mu/ds^2 \tag{8-1}$$

The ordinary Lorentz force is

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{8-2}$$

and the rate of change of energy is

$$\mathbf{F} \cdot \mathbf{v} = e\mathbf{E} \cdot \mathbf{v}$$

Then from the preceding definition of four-force,

$$\mathbf{f} = e/(1 - v^2)^{1/2}(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

and

$$f_4 = e/(1 - v^2)^{1/2}\ \mathbf{E} \cdot \mathbf{v}$$

We can show that the expressions just given for f and f_4 are equivalent to

$$f_\nu = eu_\mu\ F_{\mu\nu}$$

so that the relativistic analogue of the Newton equation becomes

$$m\ d^2z_\mu/ds^2 = e(dz_\nu/ds)\ F_{\mu\nu} \tag{8-3}$$

We can also show that this implies

$$d/ds\ [(dz_\mu/ds)^2] = 0$$

In ordinary terms the equation of motion is

$$d/dt(mv/\sqrt{1 - v^2}) = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{8-4}$$

It can also be shown by direct application of the Lagrange equations

$$d/dt\,(\partial L/\partial v_\mu) - (\partial L/\partial x_\mu) = 0$$

that the Lagrangian

$$L = -m\sqrt{1 - v^2} - e\phi + e\mathbf{A} \cdot \mathbf{v} \tag{8-5}$$

leads to these equations of motion.

Also the momenta conjugate to x is given by $\partial L/\mathbf{v}$ or

$$P = m\mathbf{v}/(1 - v^2)^{1/2} + e\mathbf{A}$$

The corresponding Hamiltonian is

$$H = e\phi + [(\mathbf{P} - e\mathbf{A})^2 + m^2]^{1/2} \tag{8-6}$$

which satisfies $(H\text{-}e\phi)^2 - (\mathbf{P}\text{-}e\mathbf{A})^2 = m^2$.

It is difficult to convert the Hamiltonian idea to a covariant or four-dimensional formulation.

But the principle of least action, which states that the action

$$S = \int L\, dt$$

shall be a minimum, will lead to the relativistic form of the equations of motion directly when expressed as

$$S = \int L\,dt = m\int ds + e\int A_\mu\,(dz_\mu/ds)\,ds$$

$$= \int [m(dz/d\alpha \cdot dz/d\alpha)^{1/2} + eA_\mu\ dz_\mu/d\alpha]\,d\alpha$$

Note that by definition

$$(ds/d\alpha)^2 = (dz_\mu/d\alpha)(dz_\mu/d\alpha)$$

It is interesting that another "action, " defined

$$S' = m/2\int (dz_\mu/d\alpha)^2\,d\alpha + e\int A_\mu(z_\mu)(dz_\mu/d\alpha)\,d\alpha$$

leads to the same result as for S in the foregoing.

the

the

Part 9

Relativistic Wave Equation

UNITS

The following convention will be used hereafter.

W e define the units of mass and time and length such that

$$c = 1 \qquad (c = 2.99793 \times 10^{10} \text{ cm/sec})$$

$$\hbar = 1 \qquad (\hbar = 1.0544 \times 10^{-27} \text{ erg/sec})$$

Table 9-1 is given as a useful reference for conversion to customary units.

The following numerical values are useful:

M_p = mass of proton = 1836.1 m = 938.2 Mev
Mass unit of atomic weights = 931.2 Mev
M_H = Mass of hydrogen atom = 1.00815 mass units
M_N = Mass of neutron = 784 kev + M_H
kT = 1 ev when T = 11,606° K
N_a = Avogadro's number = 6.025×10^{23}
$N_a e$ = 96,520 coulombs

TABLE 9-1. Notations and Units

Present notation	Meaning	Customary notation	Value
m	Mass of electron	m	
	Energy	mc^2	510.99 kev
	Momentum	mc	1704 gauss cm
	Frequency	$mc^2/\hbar$	
	Wave number	$mc/\hbar$	
1/m	Length (Compton wavelength)/2π	$\hbar/mc$	3.8615×10^{-11}cm
	Time	$\hbar/mc^2$	
e^2	Fine-structure constant (dimensionless)	$e^2/\hbar c$	1/137.038
e^2/m	Classical radius of the electron	e^2/mc^2	2.8176×10^{-11} cm
$1/me^2$	Bohr radius	$a_0 = \hbar^2/me^2$	0.52945 A

KLEIN-GORDON, PAULI, AND DIRAC EQUATIONS

According to relativistic classical mechanics, the Hamiltonian is given by

$$H = \sqrt{(\mathbf{p} - e\mathbf{A})^2 + m^2} + e\phi \tag{9-1}$$

If the quantum-mechanical operator - i∇ is used for **p**, the operation determined by the square root is undefined.

Thus the relativistic quantum-mechanical Hamiltonian has not been obtained directly from the classicalequation, Eq. (9-1).

However, it is possible to define the square of the operator and to write

$$(H - e\phi)^2 - (p - eA)^2 = m^2$$

Then, if H= i∂/∂t,

$$[-(\hbar/i)\,\partial/\partial t - e\phi]^2\Psi - [(\hbar/i)(\partial/\partial x) - e/cA_x]^2\Psi - \cdots = m^2\Psi \qquad (9\text{-}2)$$

where the square of an operator is evaluated by ordinary operator algebra.

This equation was first discovered by Schrödinger as a possible relativistic equation.

It is usually referred to as the Klein-Gordon equation.

In relativistic notation it is

$$(i\nabla_\mu - eA_\mu)(i\nabla_\mu - eA_\mu)\Psi = m^2\Psi \qquad (9\text{-}2')$$

This equation does not allow for "spin" and therefore fails to describe the fine structure of the hydrogen spectrum.

It is proposed now for application to the π meson, a particle with no spin.

To demonstrate its application to the hydrogen atom, let $\mathbf{A} = 0$ and $\phi = -Ze/r$, then let $\Psi = \chi(r) \exp(-iEt)$.

Then the equation is

$$(E - Ze^2/r)^2\chi + \nabla^2\chi = m^2\chi$$

Let $E = m + W$, where $W \ll m$, and substituting $V = Ze^2/r$,

$$(W - V)\chi + \nabla^2\chi/2m = -(W - V)^2\chi/2m$$

Neglecting the term on the right in comparison with the first term on the left gives the ordinary Schrödinger equation.

By using $(W - V)^2/2m$ as a perturbation potential, you can obtain the fine-structure splitting for hydrogen and compare with the correct values.

As an example, for the Klein-Gordon equation, if we let

$$\rho = i(\Psi^*\partial\Psi/\partial t - \Psi\partial\Psi^*/\partial t) - e\phi\Psi\Psi^* = \text{charge density}$$

$$\mathbf{j} = -i(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) - e\mathbf{A}\Psi\Psi^* = \text{current density}$$

we can then show, using the four-vector $(\rho, \mathbf{j})$, that $\nabla_\mu \mathbf{j}_\mu = 0$.

The Klein-Gordon equation leads to a result that seemed so unreasonable at the time it was first brought to light that it was considered a valid basis for rejecting the equation.

This result is the possibility of negative energy states.

To see that the Klein-Gordon equation predicts such energy states, consider the equation for a free particle, which can be written

$$\Box^2 \Psi = m^2 \Psi$$

where $\Box^2$ is the D'Alembertian operator.

In four-vector notation, this equation has the solution $\Psi = A \exp(-ip_\mu x_\mu)$, where $p_\mu p_\mu = m^2$.

Then, since

$$p_\mu p_\mu = p_4 p_4 - p_x p_x - p_y p_y - p_z p_z = E^2 - p \cdot p$$

there results

$$E = \pm (m^2 + p \cdot p)^{1/2}$$

The apparent impossibility of negative values of E led Dirac to the development of a new relativistic wave equation.

The Dirac equation proves to be correct in predicting the energy levels of the hydrogen atom and is the accepted description of the electron.

However, contrary to Dirac's original intent, his equation also leads to the existence of negative energy levels, which by now have been satisfactorily interpreted.

Those of the Klein-Gordon equation can also be interpreted.

If $y = \exp(-iEt)\chi(x,y,z)$ is a solution of the Klein-Gordon equation with constants **A** and ϕ, then $y = \exp(+iEt)\chi^*$ is a solution with - **A** and $-\phi$ replacing **A** and ϕ.

This indicates one manner in which "negative" energy solutions can be interpreted.

It is the solution for a particle of opposite charge to the electron, but the same mass.

Instead of following the original method in the development of the Dirac equation, a different approach will be used here.

The Klein-Gordon equation is actually the four-vector form of the Schrödinger equation.

With an analogous point of view, the Dirac equation can be developed as the four-vector form of the Pauli equation.

In following such a procedure, the terms involving "spin" will be included in the relativistic equation.

The idea of spin was first introduced by Pauli, but it was not at first clear why the magnetic moment of the electron had to b e taken a s $\hbar e/2mc$.

This value did seem to follow naturally from the Dirac equation, and it is often stated that only the Dirac equation produces as a consequence the correct value o f the electron's magnetic moment.

However, this is not true, as further work on the Pauli equation showed that the same value follows just as naturally, i.e., as the value that produces the greatest simplification.

Because spin is present in the Dirac equation, and absent in the Klein-Gordon, and because the Klein-Gordon equation was thought to be invalid, it is often stated that spin is a relativistic requirement.

This is incorrect, since the Klein-Gordon equation is a valid relativistic equation for particles without spin.

Thus the Schrödinger equation is

$$\mathbf{H\Psi = E\Psi}$$

where

$$\mathbf{H = 1/2m(-i\nabla - eA)^2 + e\phi}$$

and the Klein-Gordon equation is

$$[(H - e\phi)^2 - (-i\nabla - eA)^2]\,\Psi = m^2\Psi \tag{9-3}$$

Now the Pauli equation is also $H\Psi = E\Psi$, where

$$H = (1/2m)[\sigma \cdot (-i\nabla - eA)]^2 + e\phi \tag{9-4}$$

Thus, $[\sigma \cdot (-i\nabla - eA)]^2$ appearing in the Schrodinger equation has been replaced by $(-i\nabla - eA)^2$.

Then a possible relativistic version of the Pauli equation, in analogy to the Klein-Gordon equation, might be

$$(H - e\phi)^2\Psi - \{\sigma \cdot [(\hbar/i)\nabla - (e/c)A]\}^2\Psi = m^2\Psi$$

Actually, this is incorrect, but a very similar form (with H replaced by $i(\partial/\partial t)$] is correct, namely,

$$\begin{aligned}&[i(\partial/\partial t) - e\phi - \sigma \cdot (-i\nabla - eA)]\\ &\times [i(\partial/\partial t) - e\phi + \sigma \cdot (-i\nabla - eA)]\,\Psi = m^2\Psi\end{aligned} \tag{9-5}$$

This is one form of the Dirac equation.

The wave function Ψ on which the operations are being carried out is actually a matrix:

$$\Psi = \begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix}$$

A form closer to that originally proposed by Dirac may be obtained as follows.

For convenience, write

$$i(\partial/\partial t) - e\phi = \pi_4$$

$$-i\nabla - (e/c)\mathbf{A} = \boldsymbol{\pi}$$

Now let χ the function be defined by $(\pi_4 + \sigma\cdot\pi)\Psi = m\chi$.

Then Eq. (9-5) implies $(\pi_4 - \sigma\cdot\pi)\Psi = m\Psi$.

This pair of equations can be rewritten (only to arrive at a particular conventional form) by writing

$$\chi + \Psi = \Psi_a$$

$$\chi - \Psi = \Psi_b$$

Then adding and subtracting the pair of equations for Ψ, χ there results

$$\begin{aligned} \pi_4\Psi_a - \sigma\cdot\boldsymbol{\pi}\Psi_b &= m\Psi_a \\ -\pi_4\Psi_b + \sigma\cdot\boldsymbol{\pi}\Psi_a &= m\Psi_b \end{aligned} \tag{9-6}$$

These two equations may be written as one by employing a particular convention.

Define a new matrix wave function as

$$\Psi = \begin{pmatrix} \Psi_{a_1} \\ \Psi_{a_2} \\ \Psi_{b_1} \\ \Psi_{b_2} \end{pmatrix} \tag{9-7}$$

where the matrix character of Ψ_a and Ψ_b has been shown explicitly, i.e., actually

$$\Psi_a = \begin{pmatrix} \Psi_{a_1} \\ \Psi_{a_2} \end{pmatrix} \qquad \Psi_b = \begin{pmatrix} \Psi_{b_1} \\ \Psi_{b_2} \end{pmatrix}$$

Then, if the auxiliary definitions are made,

$$\gamma_4 = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right) \qquad \boldsymbol{\gamma} = \left(\begin{array}{cc|cc} 0 & 0 & & \\ 0 & 0 & & \boldsymbol{\sigma} \\ \hline & & 0 & 0 \\ -\boldsymbol{\sigma} & & 0 & 0 \end{array}\right) \tag{9-8}$$

(Note: An example of the latter definition is

$$\gamma_x = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \quad \text{since} \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

γ_y and γ_z are similar.)

The two equations in Ψ_a and Ψ_b can be written as one in the form

$$\gamma_4 \pi_4 \Psi - \boldsymbol{\gamma} \cdot \boldsymbol{\pi} \Psi = m\Psi$$

which is actually four equations in four wave functions.

Then using four-vector notation, the Dirac equation is

$$\gamma_\mu \pi_\mu \Psi = m\Psi$$

or

$$\gamma_\mu (i\nabla_\mu - eA_\mu)\Psi = m\Psi \tag{9-9}$$

It can be shown that

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = \begin{cases} 0 \text{ if } \mu \neq \nu \\ 2 \text{ if } \mu = \nu = 4 \\ -2 \text{ if } \nu = \mu = 1, 2, 3 \end{cases}$$

or

$$\gamma_t^2 = 1 \qquad \gamma_x^2 = \gamma_y^2 = \gamma_z^2 = -1$$

$$\gamma_t \gamma_x = -\gamma_x \gamma_t \qquad \gamma_x \gamma_y = -\gamma_y \gamma_x \qquad \text{etc.}$$

A similar form for the Dirac equation might be obtained b y a different

argument, by comparison to the Klein-Gordon equation. Thus with

$H = i(\partial/\partial t) = i\nabla_4$ and with $e\phi = eA_4$ Eq. (9-3) becomes

$$(i\nabla_\mu - eA_\mu)^2 \Psi = m^2 \Psi \tag{9-10}$$

in four-vector notation.

Using a similar notation in the Pauli equation, Eq.(9-4), but also using $\sigma = \gamma$ and setting $\sigma_4 = \gamma_4$ arbitrarily (to complete the definition of a four-vector form o f σ), Eq. (9-4) can be written in a form similar to Eq. (9-1

$$\{\gamma_\mu[(\hbar/i)\nabla_\mu - (e/c)A_\mu]\}^2\Psi = m^2\Psi \qquad (9\text{-}11)$$

This should be compared to Eq. (9-9).

Now the Pauli equation, Eq. (9-4), differs from the Schrödinger equation in the replacement of the three-dimensional scalar product $(\mathbf{p} - e\mathbf{A})^2$ by the square of a single quantity $\sigma \cdot (\mathbf{p} - e\mathbf{A})$.

Analogously one might guess that the four-vector product $(p_\mu - eA_\mu)^2$ in Eq. (9-10) must be replaced by the square of a single quantity $\gamma_\mu(p_\mu - eA_\mu)$, where we must invent four matrices γ_μ in four dimensions in analogy to the three matrices σ in three dimensions.

The resulting equation,

$$[\gamma_\mu(i\nabla_\mu - eA_\mu)]^2\Psi = m^2\Psi \qquad (9\text{-}11)$$

is essentially equivalent to Eq. (9-9) (operate on both sides of Eq. (9-9) by $\gamma_\mu(i\nabla_\mu - eA_\mu)$ and use Eq. (9-9) again to simplify the right-hand side).

Thus, Eq. (9-11) is equivalent to

$$(i\nabla_\mu - eA_\mu)^2 - \frac{i}{2}e\gamma_\mu\gamma_\nu F_{\mu\nu}\Psi = m^2\Psi$$

Part 10

ALGEBRA OF THE γ MATRICES

In the preceding part the Dirac equation,

$$\gamma_\mu (i\nabla_\mu - eA_\mu)\Psi = m\Psi \tag{10-1}$$

was obtained, together with a special representation for the γ's,

$$\gamma_t = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \gamma_{x,y,z} = \begin{pmatrix} 0 & \sigma_{x,y,z} \\ -\sigma_{x,y,z} & 0 \end{pmatrix} \tag{10-2}$$

where each element in these four-by-four matrices is another two-by-two matrix, that is,

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ unit matrix} \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{etc.}$$

The best way to define the γ's, however, is to give their commutation relationships, since this is all that is important in their use.

The commutation relationships do not determine a unique representation for the γ's, and the foregoing is only one of many possible representations.

The commutation relationships are

$$\gamma_t^2 = 1 \qquad \gamma_x^2 = \gamma_y^2 = \gamma_z^2 = -1$$

$$\gamma_t \gamma_{x,y,z} + \gamma_{x,y,z} \gamma_t = 0 \tag{10-3}$$

$$\gamma_x\gamma_y + \gamma_y\gamma_x = 0 \qquad \gamma_x\gamma_z + \gamma_z\gamma_x = 0 \qquad \gamma_y\gamma_z + \gamma_z\gamma_y = 0$$

or, in a unified notation,

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\rho = 2\delta_{\mu\nu}$$

$$\begin{aligned} \delta_{\mu\nu} &= 0 && \mu \neq \nu \\ &= +1 && \mu = \nu = 4 \\ &= -1 && \mu = \nu = 1,2,3 \end{aligned} \tag{10.4}$$

Note that with this definition of $\delta_{\mu\upsilon}$ and the rule for forming a scalar product,

$$\delta_{\mu\nu}\, a_\nu = a_\mu$$

Other new matrices may arise by forming products of the matrices already defined.

For example, the matrices of Eq. (10-5) are products of γ's taken two at a time.

The matrices

$$\gamma_x\gamma_y \qquad \gamma_x\gamma_z \qquad \gamma_y\gamma_z \qquad \gamma_x\gamma_t \qquad \gamma_y\gamma_t \qquad \gamma_z\gamma_t$$

are all independent of $\gamma_x, \gamma_y, \gamma_z, \gamma_t$. (They cannot be formed by a linear combination o f the latter.)

Similarly, for products of three matrices,

$$\gamma_x\,\gamma_y\,\gamma_z \quad (= \gamma_5\gamma_t)$$

$$\gamma_y\,\gamma_z\,\gamma_t \quad (= -\gamma_x\gamma_5)$$

$$\gamma_z\,\gamma_t\,\gamma_x \quad (= -\gamma_y\gamma_5)$$

$$\gamma_t\,\gamma_x\,\gamma_y \quad (= -\gamma_z\gamma_5)$$

These are the only new products of three.

For, if two of the matrices were equal, the product could be reduced, thus $\gamma_t\gamma_y\gamma_t = -\gamma_t\gamma_t\gamma_y = -\gamma_y$.

The only new product of four that can be formed is given a special name, γ_5.

$$\gamma_5 \equiv \gamma_x\gamma_y\gamma_z\gamma_t$$

Products of more than four must contain two equal so they can be reduced.

There are, therefore, sixteen linearly independent quantities.

Linear combinations of them may involve sixteen arbitrary constants.

This agrees with the fact that such a combination can be expressed by a four-by-four matrix. (It is mathematically interesting then that all four-by-four matrices can be expressed in the algebra of the γ's; this is called a Clifford algebra or hypercomplex algebra. A simpler example is that of two-by-two matrices, the so-called algebra of quaternions, which is the algebra of the Pauli spin matrices.)

It is straightforward to verify that

$$i\gamma_x\gamma_y = \begin{pmatrix}\sigma_z & 0\\ 0 & \sigma_z\end{pmatrix} \qquad i\gamma_y\gamma_z = \begin{pmatrix}\sigma_x & 0\\ 0 & \sigma_x\end{pmatrix} \qquad i\gamma_z\gamma_x = \begin{pmatrix}\sigma_y & 0\\ 0 & \sigma_y\end{pmatrix} \tag{10-5}$$

and that

$$\gamma_t\gamma_{x,y,z} = \begin{pmatrix}0 & \sigma_{x,y,z}\\ \sigma_{x\,y\,z} & 0\end{pmatrix} \equiv \alpha \text{ (definition of } \alpha\text{)}$$

It is convenient to define another γ matrix, since it occurs frequently:

$$\gamma_5 = \gamma_x\gamma_y\gamma_z\gamma_t = i\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \tag{10-6}$$

It is straightforward to verify that

$$\gamma_5\gamma_t = i\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \qquad \gamma_5\gamma_{x,y,z} = -i\begin{pmatrix} \sigma_{x,y,z} & 0 \\ 0 & -\sigma_{x,y,z} \end{pmatrix} \tag{10-7}$$

$$\gamma_5^2 = -1 \qquad \gamma_5\gamma_\mu + \gamma_\mu\gamma_5 = 0$$

For later use, it will be convenient to define

$$\not{a} = a_\mu\gamma_\mu \equiv a_t\gamma_t - a_x\gamma_x - a_y\gamma_y - a_z\gamma_z \tag{10-8}$$

from which it can be shown that

$$\not{a}\not{b} = -\not{b}\not{a} + 2a\cdot b \qquad (a\cdot b = a_\mu b_\mu)$$

$$a^2 = a_\mu a_\mu \tag{10-9}$$

$$\not{a}\gamma_5 = -\gamma_5\not{a}$$

For example, the first may be verified by writing

$$\not{a}\not{b} = (a_t\gamma_t - a_x\gamma_x - a_y\gamma_y - a_z\gamma_z)(b_t\gamma_t - b_x\gamma_x - b_y\gamma_y - b_z\gamma_z)$$

and, moving the second factor to the front, by using the commutation relationships.

Doing this with the first term ($b_t\gamma_t$) of the second factor produces

$$b_t\gamma_t(a_t\gamma_t + a_x\gamma_x + a_y\gamma_y + a_z\gamma_z)$$

since γ_t commutes with itself and anticommutes with γ_x, γ_y, and γ_z.

By performing this operation on all terms, one obtains

$$\begin{aligned}
\not{a}\not{b} &= b_t\gamma_t[(-a_t\gamma_t + a_x\gamma_x + a_y\gamma_y + a_z\gamma_z) + 2a_t\gamma_t] \\
&\quad + b_x\gamma_x[(a_t\gamma_t - a_x\gamma_x - a_y\gamma_y - a_z\gamma_z) + 2a_x\gamma_x] \\
&\quad + b_y\gamma_y[(a_t\gamma_t - a_x\gamma_x - a_y\gamma_y - a_z\gamma_z) + 2a_y\gamma_y] \\
&\quad + b_z\gamma_z[(a_t\gamma_t - a_x\gamma_x - a_y\gamma_y - a_z\gamma_z) + 2a_z\gamma_z] \\
&= -\not{b}\not{a} + 2(b_t a_t\gamma_t^2 + b_x a_x\gamma_x^2 + b_y a_y\gamma_y^2 + b_z a_z\gamma_z^2) \\
&= -\not{b}\not{a} + 2b \cdot a
\end{aligned}$$

It is straightforward to show that

$$\begin{aligned}
\gamma_x\not{a}\gamma_x &= \not{a} + 2a_x\gamma_x \\
\gamma_\mu\gamma_\mu &= 4 \\
\gamma_\mu\not{a}\gamma_\mu &= -2\not{a} \\
\gamma_\mu\not{a}\not{b}\gamma_\mu &= 4a \cdot b \\
\gamma_\mu\not{a}\not{b}\not{c}\gamma_\mu &= -2\not{c}\not{b}\not{a}
\end{aligned}$$

Expanding in power series one can show that

$$\exp[(u/2)\gamma_t\gamma_x] = \cosh(u/2) + \gamma_t\gamma_x \sinh(u/2)$$

$$\exp[(\theta/2)\gamma_x\gamma_y] = \cos(\theta/2) + \gamma_x\gamma_y \sin(\theta/2) \qquad (10\text{-}10)$$

and then show that

$$\exp[-(u/2)\gamma_t\gamma_z]\,\gamma_t\,\exp[+(u/2)\gamma_t\gamma_z] = \gamma_t \cosh u + \gamma_z \sinh u$$

$$\exp[-(u/2)\gamma_t\gamma_z]\,\gamma_z\,\exp[+(u/2)\gamma_t\gamma_z] = \gamma_z \cosh u + \gamma_t \sinh u$$

$$\exp[-(u/2)\gamma_t\gamma_z]\,\gamma_y\,\exp[+(u/2)\gamma_t\gamma_z] = \gamma_y \qquad (10\text{-}11)$$

$$\exp[-(u/2)\gamma_t\gamma_z]\,\gamma_x\,\exp[+(u/2)\gamma_t\gamma_z] = \gamma_x$$

EQUIVALENCE TRANSFORMATION

Suppose another representation for the γ's is obtained which satisfies the same commutation relationships, Eq. (10-3); will the form of the Dirac equation, Eq. (10-1), remain the same?

To answer this question, make the following transformation of the wave function $\Psi = S\Psi'$, where S is a constant matrix which is assumed to have an inverse S^{-1} ($S^{S-1} = 1$).

The Dirac equation becomes

$$\gamma_\mu \pi_\mu S\Psi' = mS\Psi' \qquad (10\text{-}12)$$

The π_μ and S commute, since π is a differential operator plus a function of position, so this equation may be written

$$\gamma_\mu S\pi_\mu \Psi' = mS\Psi'$$

Multiplying by the inverse matrix,

$$S^{-1}\gamma_\mu S\pi_\mu \Psi' = mS^{-1}S\Psi'$$

or

$$\gamma'_\mu \pi_\mu \Psi' = m\Psi'$$

where $\gamma'_\mu = S^{-1}\gamma_\mu S$.

The transformation $\gamma'_\mu = S^{-1}\gamma_\mu S$ is called a n equivalence transformation, and it is easily verified that the new γ's satisfy the commutation relationships, Eq. (10-3).

Products of γ's,

$$\gamma'_\mu \gamma'_\nu = (S^{-1}\gamma_\mu S)(S^{-1}\gamma_\nu S) = S^{-1}(\gamma_\mu \gamma_\nu)S$$

transform in exactly the same manner as the γ's, so that equations involving the γ's (the commutation relations specifically) are the same in the transformed representation.

This demonstrates another representation for the γ's, and the Dirac equation is in exactly the same form as the original, Eq. (10-1), and is equivalent in all its results.

RELATIVISTIC INVARIANCE

The relativistic invariance of the Dirac equation may be demonstrated by assuming, for the moment, that γ transforms similarly to a four-vector.

That is,

$$\gamma'_x = (\gamma_x - v\gamma_t)/(1 - v^2)^{1/2} \qquad \gamma'_t = (\gamma_t - v\gamma_x)/(1 - v^2)^{1/2}$$

$$\gamma'_y = \gamma_y \qquad \gamma'_z = \gamma_z$$

Also π transforms similarly to a four-vector because it is a combination of two four-vectors ∇_μ and A_μ.

The left-hand side $\gamma_\mu \pi_\mu$ of the Dirac equation is the product of two four-vectors and hence invariant under Lorentz transformations.

The right-hand side m is also invariant.

Transforming γ_μ as a four-vector means a new representation for the γ's, but Eqs. (10-11) can be used to show that the new γ's differ from the old γ's by an equivalence transformation; thus it is really not necessary to transform the γ's at all.

That is, the same special representation can be used in all Lorentz coordinate systems.

This leads to two possibilities in making Lorentz transformations:

1. Transform the γ's similarly to a four-vector and the wave function remains the same (except for Lorentz transformation of coordinates).
2. Use the standard representation in the Lorentz-transformed coordinate system, in which case the wave function will differ from that in (1) by an equivalence transformation.

HAMILTONIAN FORM O F THE DIRAC EQUATION

To show that the Dirac equation reduces to the Schrödinger equation for low velocities, it is convenient to write it in Hamiltonian form.

The original term, Eq. (10-1), may be written

$$\gamma_t [-(\hbar/i)(\partial/\partial t) - e\phi]\Psi - \boldsymbol{\gamma} \cdot [(\hbar/i)\nabla - e\mathbf{A}]\Psi = m\Psi$$

Multiplying by $c\gamma_t$ and rearranging terms gives

$$-(\hbar/i)(\partial\Psi/\partial t) = \{\gamma_t\boldsymbol{\gamma}\cdot[(\hbar/i)\nabla - e\mathbf{A}] + e\phi + \gamma_t m\}\Psi$$

$$= H\Psi$$

By Eq. (10-5), H is written

$$H = \boldsymbol{\alpha}\cdot[(\hbar/i)\nabla - e\mathbf{A}] + e\phi + m\beta$$

where $\beta = \gamma_t$, $\alpha_{x,y,z} = \gamma_t, \gamma_{x,y,z}$.

Eq. (10-5), and the α's satisfy following commutation relations: $\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = 1$ and all pairs anticommute.

It will b e noted that α, β are Hermitian matrices in our special representation, so that in this representation H is Hermitian.

One can then show that a probability density $\rho = \Psi^*\Psi$ and a probability current $j = \Psi^*\alpha\Psi$ satisfy the continuity equation

$$(\partial\rho/\partial t) + \nabla\cdot\mathbf{j} = 0$$

Note: Ψ is a four-component wave function and

$$\rho = \Psi^*\Psi = (\Psi_1^*\Psi_2^*\Psi_3^*\Psi_4^*)\begin{pmatrix}\Psi_1\\ \Psi_2\\ \Psi_3\\ \Psi_4\end{pmatrix} = \Psi_1^*\Psi_1 + \Psi_2^*\Psi_2 + \Psi_3^*\Psi_3 + \Psi_4^*\Psi$$

$$j_x = \sum_{ij}\Psi_i^*(\alpha_x)_{ij}\Psi_j$$

$$= \Psi_4^*\Psi_1 + \Psi_3^*\Psi_2 + \Psi_2^*\Psi_3 + \Psi_1^*\Psi_4$$

Part 11

It should be noted that β and α are Hermitian only in certain representations.

In particular, they are Hermitian in the representation employed thus far; this will be called the standard representation and expressions in it will be labeled S.R. when appropriate.

The Hermitian property of α and β is necessary in order to get

$$\rho = \Psi^*\Psi$$

$$\mathbf{j} = \Psi^* \boldsymbol{\alpha} \Psi \qquad \text{S.R.} \tag{11-1}$$

as the expressions for charge and current density.

Hence they are not true in all representations.

The Dirac equation is (with $\hbar$, c restored)

$$-(\hbar/i)(\partial\Psi/\partial t) = H\Psi \qquad H = \beta mc^2 + e\phi + c\alpha \cdot [(\hbar/i)\nabla - (e/c)\mathbf{A}] \tag{11-2}$$

The expected value of x is

$$\langle x \rangle = \int \Psi^* x \Psi \, d\,\text{vol}$$

$$= \int (\Psi_1^* x \Psi_1 + \Psi_2^* x \Psi_2 + \Psi_3^* x \Psi_3 + \Psi_4^* x \Psi_4) \, d\,\text{vol} \qquad \text{S.R.}$$

remembering that Ψ now is a four-component wave function.

Similarly it may be verified that

$$\langle \alpha \rangle = \int \Psi^* \boldsymbol{\alpha} \Psi \, \mathrm{d\,vol}$$

$$\langle \alpha_x \rangle = \int (\Psi_4^* \Psi_1 + \Psi_3^* \Psi_2 + \Psi_2^* \Psi_3 + \Psi_1^* \Psi_4) \, \mathrm{d\,vol} \qquad \text{S.R.}$$

Also matrix elements are formally the same as before.

For example,

$$(\boldsymbol{\alpha})_{mn} = \int \Psi_m^* \boldsymbol{\alpha} \Psi_n \, \mathrm{d\,vol}$$

If A is any operator then its time derivative is

$$\dot{A} = i\,(HA - AH) + \partial A / \partial t$$

For x the result is clearly

$$\dot{\mathbf{x}} = i\,(H\mathbf{x} - \mathbf{x}H) = \alpha$$

since x commutes with all terms in H except $\mathbf{p} \cdot \alpha$.

But $\alpha^2 = 1$, so the eigenvalues of α are ± 1.

Hence the eigenvelocities of $\dot{\mathbf{x}}$ are ± speed of light.

This result is sometimes made plausible by the argument that a precise determination of velocity implies precise determinations of position at two times.

Then, by the uncertainty principle, the momentum is completely uncertain and all values are equally likely.

With the relativistic relation between velocity and momentum, this is seen to imply that velocities near the speed of light are more probable, so that in the limit the expected value of the velocity is the speed of light.

Similarly,

$$(\dot{\overline{\mathbf{p} - e\mathbf{A}}})_x = i(Hp_x - p_xH) - ie(HA_x - A_xH) - e\partial A_x/\partial t$$

$$= -e(\partial\phi/\partial x) + e\boldsymbol{\alpha}\cdot(\partial\mathbf{A}/\partial x) - e(\boldsymbol{\alpha}\cdot\nabla)A_x - e(\partial A_x/\partial t)$$

The terms in $\mathbf{A}$ and A_x, except the last, expand as follows:

$$e\left(\alpha_x\frac{\partial A_x}{\partial x} + \alpha_y\frac{\partial A_y}{\partial x} + \alpha_z\frac{\partial A_z}{\partial x} - \alpha_x\frac{\partial A_x}{\partial x} - \alpha_y\frac{\partial A_y}{\partial y} - \alpha_z\frac{\partial A_x}{\partial z}\right)$$

This seen to be the x component of

$$e\boldsymbol{\alpha}\times(\nabla\times\mathbf{A}) = e\boldsymbol{\alpha}\times\mathbf{B}$$

The first and last terms form the x component of $\mathbf{E}$.

Therefore,

$$(\dot{\overline{\mathbf{p} - e\mathbf{A}}}) = e(\mathbf{E} + \boldsymbol{\alpha}\times\mathbf{B}) = \mathbf{F}$$

where $\mathbf{F}$ is the analogue of the Lorentz force.

This equation is sometimes regarded as the analogue of Newton's equations.

But, since there is no direct connection between this equation and $\dot{\mathbf{x}}$, it does not lead directly to Newton's equations in the limit of small velocities and hence is not completely acceptable as a suitable analogue.

The following relations may be verified as true but their meaning is not yet completely understood, if at all:

$$(d/dt)\,[\mathbf{x} + (i/2m)\beta\boldsymbol{\alpha}] = (\beta/m)(\mathbf{p} - e\mathbf{A})$$

$$(d/dt)\,[t + (i/2m)\,\beta] = (\beta/m)(H - e\phi)$$

$$i(d/dt)(\alpha_x\alpha_y\alpha_z) = -2m\beta\alpha_x\alpha_y\alpha_z$$

$$-(d/dt)(\beta\boldsymbol{\sigma}) = 2(\beta\,\alpha_x\alpha_y\alpha_z)(\mathbf{p} - e\mathbf{A})$$

where in the last relation σ means the matrix

$$\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}$$

so that

$$\sigma_z = -i\,\alpha_x\,\alpha_y, \text{ etc.}$$

From analogy to classical physics, one might expect that the angular momentum operator is now

$$\mathbf{L} = \mathbf{R} \times (\mathbf{p} - e\mathbf{A})$$

Note that in classical physics

$$\mathbf{p} - e\mathbf{A} = m\mathbf{v}\,(1 - v^2)^{-1/2}$$

From previous results for $\dot{\mathbf{R}}$ and $\overline{(\mathbf{p} - e\mathbf{A})}^{\cdot}$, the time derivative of $\mathbf{L}$ may be written

$$\dot{\mathbf{L}} = \dot{\mathbf{R}} \times (\mathbf{p} - e\mathbf{A}) + \mathbf{R} \times \overline{(\mathbf{p} - e\mathbf{A})}^{\cdot}$$

$$= \boldsymbol{\alpha} \times (\mathbf{p} - e\mathbf{A}) + \mathbf{R} \times \mathbf{F}$$

The last term may be interpreted as torque.

For a central force **F**, this term vanishes.

But then it is seen that $\dot{\mathbf{L}} \neq 0$ because of the first term; that is, the angular momentum **L** is not conserved, even with central forces.

But consider the time derivative of the operator σ defined as

$$\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}$$

where $\sigma_z = -\alpha_x\alpha_y$, etc.

The z component is seen to commute with the β, eϕ, and α_z terms of H but not with the α_x and α_y terms, so that $\sigma_z =+ 1\ (H\alpha_x\alpha_y - \alpha_x\alpha_y H) = +\ (\alpha_x\pi_x\alpha_x\alpha_y - \alpha_x\alpha_y\alpha_x\pi_x + \alpha_y\pi_y\alpha_x\alpha_y - \alpha_x\alpha_y\alpha_y\pi_y)$,where

$$\boldsymbol{\pi} = (-i\boldsymbol{\nabla} - e\mathbf{A})$$

But

$$\alpha_x\pi_x\alpha_x\alpha_y = \alpha_x\alpha_x\alpha_y\pi_x = \alpha_y\pi_x$$

$$-\alpha_x\alpha_y\alpha_x\pi_x = \alpha_x\alpha_x\alpha_y\pi_x = \alpha_y\pi_x$$

$$\alpha_y\pi_y\alpha_x\alpha_y = -\alpha_y\alpha_y\alpha_x\pi_y = -\alpha_x\pi_y$$

$$-\alpha_x\alpha_y\alpha_y\pi_y = -\alpha_x\pi_y$$

so that

$$\dot{\boldsymbol{\sigma}}_z = (2\alpha_y\pi_x - 2\alpha_x\pi_y)$$

This is seen to be the z component of $-2\alpha \times \pi$.

Finally then,

$$1/2(\dot{\sigma}) = -\alpha \times \pi = -\alpha \times (\mathbf{p} - e\mathbf{A})$$

and this is the first term of $\dot{\mathbf{L}}$ with negative sign.

Therefore it follows that

$$(d/dt)[\mathbf{L} + (\hbar/2)\boldsymbol{\sigma}] = \mathbf{R} \times \mathbf{F}$$

which vanishes with central forces.

The operator $\mathbf{L} + (i\hbar/2)\sigma$ may be regarded as the total angular momentum operator, where $\mathbf{L}$ represents orbital angular momentum and $(\hbar 2)\sigma$ intrinsic angular momentum for spin 1/2.

Thus total angular momentum is conserved with central forces.

In a stationary field $\phi = 0$, $\partial A/\partial t = 0$, one can show that

$$\boldsymbol{\sigma} \cdot (\mathbf{p} - e\mathbf{A})$$

is a constant of the motion.

Note that this is a consequence of the anomalous gyromagnetic ratio of the electron.

It also means that the cyclotron frequency of the electron equals its rate of precession in a magnetic field.

In a stationary magnetic field $\phi = 0$, $\partial A/\partial t = 0$, and for a stationary state, one can show that Ψ_1, Ψ_2 in

$$\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}$$

are the same as Ψ_1, Ψ_2 in the Pauli equation.

Also, if E_{pauli} is the kinetic energy in the Pauli equation and $E_{Dirac} = W + m$ is the rest plus kinetic energy in the Dirac equation, and, thus, we can show that

$$E_{Dirac} = \sqrt{2mE_{Pauli} + m^2}$$

which is an extremely simple relationship.

NONRELATIVISTIC APPROXIMATION TO THE DIRAC EQUATION

It will be assumed that all potentials are stationary and stationary states will be considered.

This makes the work simpler but is not necessary.

In this case

$$\Psi = e^{-iEt}\,\Psi(\mathbf{X})$$

$$H\Psi = E\Psi \text{ (Dirac Hamiltonian)}$$

and put

$$E = m + W$$

That is,

$$H\Psi = (m + W)\Psi = \boldsymbol{\alpha} \cdot (\mathbf{p} - e\mathbf{A})\,\Psi + \beta\, m\Psi + e\phi\Psi$$

It will be recalled with Ψ written as Eq. (9-5) and with α, β as given in Part 10, the previous equation may be written as two equations (9-4').

$$(m + W)\Psi_a = \boldsymbol{\sigma} \cdot \boldsymbol{\pi}\Psi_b + m\Psi_a + V\Psi_a \tag{11-4}$$

$$(m + W)\Psi_b = \boldsymbol{\sigma} \cdot \boldsymbol{\pi}\Psi_a - m\Psi_b + V\Psi_b \tag{11-5}$$

where, as before, $\mathbf{x} = (\mathbf{p} - e\mathbf{A})$ and $V = e\,\phi$. Simplifying and solving Eq. (11-5) for Ψ_b gives

$$\Psi_b = [1/(2m + W - V)]\,(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})\Psi_a \tag{11-6}$$

It is noted that if W and V are << 2m, then $\Psi_b \sim (v/c)\,\Psi_a$.

For this reason Ψ_a, and Ψ_b are sometimes referred to as the large and small components of Ψ, respectively.

Substitution of Ψ_b from Eq. (11-6) into Eq. (11-4) gives

$$W\Psi_a = (\sigma \cdot \pi)[1/(2m + W - V)]\,(\sigma \cdot \pi)\Psi_a + V\Psi_a \tag{11.7}$$

and, if W and V are neglected in comparison to 2m, the result is

$$W\Psi_a = (1/2m)(\sigma \cdot \pi)^2 \Psi_a + V\Psi_a$$

This is the Pauli equation, Eq. (9-4).

Now the approximation will be carried out to second order, that is, to order v^2/c^2, to determine just what error may be expected from use of the Pauli equation.

Part 12

Using the results of Part 11, given by Eqs. (11-6) and (11-7), the low-energy approximation $(w - V) << 2m$ will be made, keeping terms to order v^4.

Thus

$$(2m + W - V)^{-1} \approx 1/2m - (w - V)/(2m)^2 \tag{12.1}$$

Then Eq. (11-7) becomes

$$(W - V)\Psi_a = (1/2m)(\sigma \cdot \pi)^2 \Psi_a - (1/4m^2)(\sigma \cdot \pi)(W - V)(\sigma \cdot \pi)\Psi_a \tag{12.2}$$

while the normalizing requirement $\int(\Psi_a{}^2 + \Psi_b{}^2)\ \mathrm{d\ vol} = 1$, becomes

$$\int \Psi_a{}^*[1 + (\sigma \cdot \pi)^2/(4m^2)]\Psi_a\ \mathrm{d\ vol} = 1 \tag{12.3}$$

By use of the substitution

$$\chi = [1 + (\sigma \cdot \pi)/(8m^2)]\ \Psi_a \tag{12.4}$$

the normalizing integral can be simplified to read (to order v^2/c^2)

$$\int \chi^* \chi\ \mathrm{d\ vol} = 1$$

This substitution also allows easier interpretation of Eq. (12-2).

Rewriting Eg. (12-2),

$$[1 + (\sigma \cdot \pi)^2/(8m^2)]\ (W - V)\ [1 + (\sigma \cdot \pi)^2/(8m^2)]\Psi_a$$

$$= (1/2m)(\sigma \cdot \pi)^2 \Psi_a + (1/8m^2)[(\sigma \cdot \pi)^2(W - V) - 2(\sigma \cdot \pi)(W - V)$$

$$\times\ (\sigma \cdot \pi) + (W - V)(\sigma \cdot \pi)^2]\Psi_a$$

Then applying Eq. (12-4) and dividing by $1 + (\sigma \cdot \pi)^2/(8m^2)$, there results

$$(W - V)\chi = (1/2m)(\sigma \cdot \pi)^2 \chi - (1/8m^3)(\sigma \cdot \pi)^4 \chi$$

$$+ (1/8m^2)[(\sigma \cdot \pi)^2 (W - V) - 2(\sigma \cdot \pi)(W - V)(\sigma \cdot \pi)$$

$$+ (W - V)(\sigma \cdot \pi)^2] \chi \qquad (12.5)$$

The techniques of operator algebra may be used to convert Eq. (12-5) to a form more easily interpreted.

In particular one should recall that

$$A^2B - 2ABA + BA^2 = A(AB - BA) - (AB - BA)A$$

Then, since $\pi = (p - eA)$, and since

$$(\sigma \cdot \pi)(W - V) - (W - V)(\sigma \cdot \pi) = -i(\sigma \cdot \nabla V)$$

$$= +i(\sigma \cdot E)$$

there results [with $\sigma \cdot \pi = A$ and $(W - V) = B$ in the foregoing],

$$i(\sigma \cdot \pi)(\sigma \cdot E) - i(\sigma \cdot E)(\sigma \cdot \pi) = \nabla \cdot E + 2\sigma \cdot (\pi \times E)$$

(since $\nabla \times E \sim \partial B/\partial t = 0$ here), so Eq. (12-5) can be expanded as

$$W\chi = \underset{(1)}{V\chi} + \underset{(2)}{(1/2m)(\mathbf{p} - e\mathbf{A})\cdot(\mathbf{p} - e\mathbf{A})\chi} - \underset{(3)}{(e/2m)(\boldsymbol{\sigma}\cdot\mathbf{B})\chi}$$
$$-\underset{(4)}{(1/8m^3)(\mathbf{p}\cdot\mathbf{p})^2\chi}$$
$$+ \underset{(5)}{(e^2/8m^2)[\nabla\cdot\mathbf{E}} + \underset{(6)}{2\boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A})\times\mathbf{E}]\chi \qquad (12.6)$$

In this form the wave equation may be interpreted by considering each term of Eq. (12-6) separately.

Term (1) gives the ordinary scalar potential energy as it has appeared before.

Term (2) can be interpreted as the kinetic energy.

Term (3), the Pauli spin effect, is just as it appears in the Pauli equation.

Term (4) is a relativistic correction to the kinetic energy. The correction derives from

$$E = (m^2 + p^2)^{1/2} = m(1 + p^2/m^2)^{1/2}$$
$$= m + p^2/2m - p^4/8m^3 + \cdots$$

The last term in this expansion is equivalent to term (4).

Terms (5) and (6) express the spin-orbit coupling.

To understand this interpretation consider the part of term (6) given by $\boldsymbol{\sigma} \cdot (\mathbf{p} \times \mathbf{E})$.

In an inverse-square field this is proportional to $\boldsymbol{\sigma} \cdot (\mathbf{p} \times \mathbf{r})/r^3$.

The factor **p** x **r** can be interpreted as the angular momentum **L** to get $(\boldsymbol{\sigma} \cdot \mathbf{L})/r^3$, the spin-orbit coupling.

This term has no effect when the electron is in a s-state (L = 0).

On the other hand, (5) reduces t $\nabla \cdot \mathbf{E} = 4\pi Z\delta(\mathbf{r})$, which affects only the s-states (when the wave function is nonzero at r = 0).

So (5) and (6) together result in a continuous function for spin-orbit coupling.

The magnetic moment of the electron e/2m, appears as the coefficient of term (3), and again of terms (5) and (6), i.e., $(e/2m)(1/4m^2)$.

A classical argument can be made to interpret term (6).

A charge moving through an electric field with velocity **v** feels an effective magnetic field
$\mathbf{B} = \mathbf{v} \times \mathbf{E} = (1/m)(\mathbf{p} - e\mathbf{A}) \times \mathbf{E}$, and term (6) is just the energy (e/2m) x (σ-**B**) in this field.

We get a factor 2 too much this way, however.

Even before the development of the Dirac equation, Thomas showed that this simple classical argument is incomplete and gave the correct term (6).

The situation is different for the anomalous moments introduced by Pauli to describe neutrons and protons (see below).

In Pauli's modified equation, the anomalous moment does appear with the factor 2 when multiplying terms (5) and (6).

Part 13 Solution of the Dirac Equation for a Free Particle

It will be convenient to use the form of the Dirac equation with the γ's when solving for the free-particle wave functions

$$\gamma_\mu (i\nabla_\mu - eA_\mu)\Psi = m\Psi$$

Using the definition of Part 10, $\not{a} = \gamma_\mu a_\mu$,

$$\not{A} = \gamma_\mu A_\mu = \gamma_t A_t - \gamma_x A_x - \gamma_y A_y - \gamma_z A_z$$

$$\not{\nabla} = \gamma_\mu \nabla_\mu = \gamma_t \nabla_t - \gamma_x \nabla_x - \gamma_y \nabla_y - \gamma_z \nabla_z$$

and the Dirac equation may be written

$$(i\not{\nabla} - e\not{A})\Psi = m\Psi \tag{13.1}$$

(Recall that the quantity $\not{a} = \gamma_\mu a_\mu$ is invariant under a Lorentz transformation.)

It is necessary to put the probability density and current into a four-dimensional form.

In the special representation, the probability density and current are given by

$$\rho = \Psi^* \Psi \qquad \mathbf{j} = \Psi^* \alpha \Psi$$

If the relativistic adjoint of Ψ is defined

$$\tilde{\Psi} = \Psi^* \beta \tag{13.2}$$

in the standard representation, then the probability density and current may be written

$$\rho = \tilde{\Psi} \beta \Psi \qquad j_\mu = \tilde{\Psi} \gamma_\mu \Psi$$

Note Ψ is a four-component column vector,

$$\begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}$$

The adjoint $\tilde{\Psi}$ is the four-component row vector $\Psi_1^*, \Psi_2^* - \Psi_3^* - \Psi_4^*$ in the standard representation.

Multiplication by b changes the sign of the third and fourth components, in addition to changing Ψ from a column vector to a row vectors.

In general, the adjoint of an operator N is denoted by $\tilde{N}$, and $\tilde{N}$ is the same as N except that the order of all γ 's appearing in it is reversed, and each explicit i (not those contained in the γ 's) is replaced by - i.

The following property takes the place of the hermitian property so useful in nonrelativistic quantum mechanics:

$$(\tilde{\Psi}_2 N \Psi_1)^* = (\tilde{\Psi}_1 \tilde{N} \Psi_2) \tag{13.3}$$

One can show that the adjoint of Ψ satisfies

$$\tilde{\Psi}(-i\not{\nabla} - e\not{A}) = m\tilde{\Psi} \tag{13.4}$$

For a free particle, there are no potentials, so $\not{A} = 0$ and the Dirac equation becomes

$$i\not{\nabla}\Psi = m\Psi \tag{13.5}$$

A solution is

$$\Psi = ue^{-ip\cdot x} = ue^{-ip_\nu x_\nu} \tag{13.6}$$

Ψ is a four-component wave function and what is meant by this trial solution is that each of the four components is of this form, that is,

$$\begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} e^{-ip\cdot x}$$

Thus u_1, u_2, u_3, and u_4 are the components of a column vector, and u is called a Dirac spinor.

The problem is now to determine what restrictions must be placed on the u's and p's in order that the trial solution satisfy the Dirac equation.

The ∇_μ operation on each component of Ψ multiplies each component by $-ip_\mu$, so that the result of this operation on Ψ produces

$$\nabla_\mu \Psi = \nabla_\mu \, u e^{-i p_\nu x_\nu} = -i p_\mu \, u e^{-i p_\nu x_\nu} = -i p_\mu \Psi$$

so that Eq. (13-5) becomes

$$i\gamma_\mu(-ip_\mu)\Psi = \gamma_\mu p_\mu \Psi = \not{p}\Psi = m\Psi \qquad (13.7)$$

Thus the assumed solution will be satisfactory if $\not{p}u = mu$.

To simplify writing, it will now be assumed that the particle moves in the xy plane, so that

$$p_1 = p_x \qquad p_2 = p_y \qquad p_3 = 0 \qquad p_4 = E$$

Under these conditions, $\not{p} = \gamma_t E - \gamma_y p_y - \gamma_x p_x$.

In standard representation

$$\gamma_t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \gamma_{x,y} = \begin{pmatrix} 0 & \sigma_{x,y} \\ -\sigma_{x,y} & 0 \end{pmatrix}$$

so $\not{p} - m$ becomes

$$\begin{pmatrix} E-m & 0 & 0 & -p_x + ip_y \\ 0 & E-m & -(p_x + p_y) & 0 \\ 0 & p_x - ip_y & -(E+m) & 0 \\ p_x + ip_y & 0 & 0 & -(E+m) \end{pmatrix} \tag{13.8}$$

By components, Eq. (13-7) becomes

$$(E - m)u_1 - (p_x - ip_y)u_4 = 0 \tag{13-9a}$$

$$(E - m)u_2 - (p_x + ip_y)u_3 = 0 \tag{13-9b}$$

$$(p_x - ip_y)u_2 - (E + m)u_3 = 0 \tag{13-9c}$$

$$(p_x + ip_y)u_1 - (E + m)u_4 = 0 \tag{13-9d}$$

The ratio u_1/u_4, can be determined from Eq. (13-9a) and also from Eq (13-9d).

These two values must agree in order that Eq. (13-6) be a solution.

Thus

$$u_1/u_4 = (p_x - ip_y)/(E - m) = (E + m)/(p_x + ip_y)$$

or

$$p_x^2 + p_y^2 + m^2 = E^2 \tag{13.10}$$

This is not a surprising condition.

It states that the p, must be chosen so as to satisfy the relativistic equation for total energy.

Similarly, Eqs. (13-9b) and (13-9c) can be solved for u_2/u_3 giving

$$u_2/u_3 = (p_x + ip_y)/(E - m) = (E + m)/(p_x - ip_y)$$

which also leads to condition (13-10).

A more elegant way of obtaining exactly the same condition is to start directly with Eq. (13-7).

Then, by multiplying this equation by $\not{p}$ gives

$$\not{p}(\not{p}u) = \not{p}(mu) = m(\not{p}u) = m^2u$$

Using Eq. (10-9),

$$\not{p}\not{p} = p \cdot p = E^2 - p_x{}^2 - p_y{}^2$$

so that the condition becomes

$$E^2 - p_x{}^2 - p_y{}^2 = m^2 \qquad \text{or} \qquad u = 0$$

The former is the same condition as obtained before, and the latter is a trivial solution (no wave function). Dirac equation.

Evidently there are two linearly independent solutions of the free-particle Dirac equation.

This is so because substitution of the assumed solution, Eq. (13-6), into the Dirac equation gives only a condition on pairs of the u_1, u_4 and u_2, u_3.

It is convenient to choose the independent solutions so that each has two components which are zero.

Thus the u's for the two solutions can be taken as

$$\begin{pmatrix} F \\ 0 \\ 0 \\ p_+ \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ F \\ p_- \\ 0 \end{pmatrix} \tag{13.11}$$

where the following notation has been used:

$$\begin{aligned} F &= E + m \\ p_+ &= p_x + ip_y \\ p_- &= p_x - ip_y \end{aligned} \tag{13.12}$$

These solutions are not normalized.

DEFINITION OF THE SPIN OF A MOVING ELECTRON

What do the two linearly independent solutions mean?

There must be some physical quantity that can still be specified, which will uniquely determine the wave function.

It is known, for example, that in the coordinate system in which the particle is stationary there are two possible spin orientations.

Mathematically speaking, the existence of two solutions to the eigenvalue equation $\not{p}u = mu$ implies the existence of an operator that commutes with $\not{p}$.

This operator will have to be discovered.

Observe that γ_5 anticommutes with $\not{p}$; that is, $\gamma_5 \not{p} = -\not{p}\gamma_5$.

Also observe that any operator $\not{W}$ will anticommute with $\not{p}$ if $W \cdot p = 0$, because

$$\not{W}\not{p} = -\not{p}\not{W} + 2W \cdot p \qquad (10.9)$$

The combination $\gamma_5 \not{W}$ of these two anticommuting operators is an operator which commutes with $\not{p}$; that is,

$$(\gamma_5 \not{W})\not{p} = -\gamma_5 \not{p}\not{W} = \not{p}(\gamma_5 \not{W})$$

The eigenvalues of the operator (i $\gamma_5 \not{W}$) must now be found (the i has been added to make eigenvalues come out real in what follows).

Denoting these eigenvalues by s,

$$(i\gamma_5 \not{W})u = su \qquad (13.13)$$

To find the possible values of s, multiply Eq. (13-13) by $i\gamma \not{W}$,

$$(i\gamma_5 \not{W})(i\gamma_5 \not{W})u = -\gamma_5 \not{W}\gamma_5 \not{W}u = -W \cdot W\, u = i\gamma_5 W\, su = s^2 u$$

or

$$-W \cdot W = s^2$$

If W•W is taken to be -1, the eigenvalues of the operator i $\gamma_5 \not{W}$ are +1.

The significance of the choice W•W = -1 is as follows: In the system in which the particle is at rest, $p_x = p_y = p_z = 0$ and $P_4 = E$.

Then

$$0 = p \cdot W = p_4 W_4 \qquad \text{or} \qquad W_4 = 0$$

Thus, W•W = -**W•W** = -1 or **W•W** = 1.

This states that in the coordinate system in which the particle is at rest, W is an ordinary vector (it has zero fourth component) with unit length.

When the particle moves in the xy plane, choose $\not{W}$ to be γ_z, so the operator equation for i $\gamma_5 \not{W}$ becomes

$$i\gamma_5 \gamma_z u = su$$

Using relationships derived in Lecture 10, this becomes, for a stationary particle(for a stationary particle $\gamma_t u = u$),

$$i\gamma_5\gamma_z u = i\gamma_x\gamma_y\gamma_t u = i\gamma_x\gamma_y u = \left(\begin{array}{c|c} \sigma_z & 0 \\ \hline 0 & \sigma_z \end{array}\right) u = su$$

This choice makes $\not{W}$ the σ_z, operator, and the relationship with spin is clearly demonstrated.

If we define u to satisfy both $\not{p}u = mu$ and $i\gamma_5 \not{W} u = su$, this completely specifies u.

It represents a particle moving with momentum p_μ and having its spin (in the coordinate system moving with the particle) along the W_μ axis either positive (s = +1) or negative (s = -1).

Another way of obtaining the wave function for a freely moving electron is to perform an equivalence transformation of the wave function as in Eq. (10-12).

If the electron is initially at rest with its spin up or down in the z direction, then the spinor for an electron moving with a velocity v in the spatial direction **k** is

$$u(\mathbf{k}) = Su \qquad u = (2m)^{1/2} u_0 \qquad u_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

From Eq. (10-11), S is given by

$$S = \exp\left[(u/2)\gamma_t \gamma_k\right] \qquad \cosh u = 1/(1 - v^2)^{1/2}$$

Now

$$\exp\left[-(u/2)\gamma_t \gamma_k\right] = \cosh (u/2) + \gamma_t \gamma_k \sinh (u/2)$$

and

$$(2m)^{1/2} \cosh (u/2) = [m(1 - v^2)^{-1/2} + m]^{1/2} = (E + m)^{1/2}$$

$$(2m)^{1/2} \sinh (u/2) = (E - m)^{1/2}$$

Therefore,

$$u_{(k)} = [(E + m)^{1/2} + \gamma_t \gamma_k (E - m)^{1/2}] u_0$$

Writing $f = (E+m)$, $\alpha = \gamma_t \boldsymbol{\gamma}$, and noting $(E^2 - m^2)^{1/2} = p_k$, we get

$$u_{(k)} = (1/\sqrt{F})(E + m + \boldsymbol{\alpha} \cdot \mathbf{p})u_0$$

For the case that **p** is in the xy plane, this just gives the result, Eq. (13-11), with a normalization factor $1/\sqrt{F}$.

Noticing that for an electron at rest $\gamma_t u_0 = u_0$, $u_{(k)}$ may be written

$$(1/\sqrt{F})(E\gamma_t - \boldsymbol{\gamma} \cdot \mathbf{P} + m)u_0$$

or

$$u_{(k)} = (1/\sqrt{F})(\not{p} + m)u_0$$

It is clear that this is a solution to the free-particle Dirac equation

$$(\not{p} - m)u_k = 0 \tag{13.7}$$

for

$$(\not{p} + m)(\not{p} - m) = p^2 - m^2 = 0 \qquad p^2 = m^2$$

NORMALIZATION OF THE WAVE FUNCTIONS

In nonrelativistic quantum mechanics, a plane wave is normalized to give unity probability of finding the particle in a cubic centimeter, that is, $\Psi^*\Psi = 1$.

An analogous normalization for the relativistic plane wave might be some-thing like

$$\Psi^*\Psi = u^*u = \tilde{u}\gamma_t u = 1$$

However, $\Psi^*\Psi$ transforms similarly to the fourth component of a four-vector (it is the fourth component of four-vector current), so this normalization would not be an invariant.

It is possible to make a relativistically invariant normalization by setting u*u equal to the fourth component of a suitable four-vector.

paFor example, E is the fourth component of the momentum four-vector p_μ, so the wave function could be normalized by

$$\tilde{u}\gamma_t u = 2E$$

The constant of proportionality (2) is chosen for convenience in later formulas.

Working out $(\tilde{u}\gamma_t u)$ for the s= + 1 state,

$$(\tilde{u}\gamma_t u) = \overbrace{F \quad 0 \quad 0 \quad -p_-}^{} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} F \\ 0 \\ 0 \\ p_+ \end{pmatrix} \times C_1^2$$

$$= \overbrace{F \quad 0 \quad 0 \quad -p_-}^{} \begin{pmatrix} F \\ 0 \\ 0 \\ -p_+ \end{pmatrix} \times C_1^2 = (F^2 + p_+p_-)C_1^2 = 2E(E + m)C_1^2$$

The C_1 is the normalizing factor multiplying the wave functions of Eq.(13-11).

In order that $(\tilde{u}\gamma_t u)$ be equal to 2E, the normalizing factor must be chosen $(E + m)^{-1/2} = (F)^{-1/2}$.

In terms of $(\tilde{u}\,u)$, this normalizing condition becomes

$$(\tilde{u}\,u) = \overbrace{F \quad 0 \quad 0 \quad -p_-}^{} \begin{pmatrix} F \\ 0 \\ 0 \\ p_+ \end{pmatrix} \times \frac{1}{F} = (F^2 - p_-p_+)\frac{1}{F}$$

$$= \frac{2m^2 + 2mE}{E + m} = 2m$$

The same result is obtained for the s = - 1 state.

Thus the normalizing condition can be taken as

$$(\tilde{u}\, u) = 2m \tag{13.14}$$

In a similar manner, the following can be shown to be true:

$$(\tilde{u}\, \gamma_x\, u) = 2p_x$$

$$(\tilde{u}\, \gamma_y\, u) = 2p_y$$

$$(\tilde{u}\, \gamma_z\, u) = 0$$

It will be convenient to have the matrix elements of all the γ's between various initial and final states, so Table 13-1 (next slide) has been worked out.

Limiting cases: To obtain the case where 1 is a positron at rest, the table gives $\sqrt{F_2}(\tilde{u}_2 N u_1)$ if one puts $F_1 = 0, P_1, = 1 = P_{r-}$ in the table.

For both at rest as positrons, the table gives $(\tilde{u}_2 N u_1)$ with $F_1 = F_2 = 0$; $P_{s+} = P_{2+} = 1$.

TABLE 13-1. Matrix Elements for Particle Moving in the xy Plane

Matrix N	$(\tilde{u}Nu)$ $s = +1$	$\sqrt{F_1F_2}\,(\tilde{u}_2Nu_1)$ $s_1 = +1$ $s_2 = +1$	$\sqrt{F_1F_2}\,(\tilde{u}_2Nu_1)$ $s_1 = +1$ $s_2 = -1$	$\sqrt{F_1F_2}\,(\tilde{u}_2Nu_1)$ $s_1 = -1$ $s_2 = -1$	$\sqrt{F_1F_2}\,(\tilde{u}_2Nu_1)$ $s_1 = -1$ $s_2 = +1$
1	$2m$	$F_2F_1 - p_{1+}p_{2-}$	0		
γ_x	$2p_x$	$F_2p_{1+} + p_{2-}F_1$	0		
γ_y	$2p_y$	$-iF_2p_{1+} + ip_{2-}F_1$	0		
γ_z	0	0	$-p_{1+}F_2 + p_{2+}F_1$		
γ_t	$2E$	$F_2F_1 + p_{1+}p_{2-}$	0		
$\gamma_y\gamma_z$	0	0	$-iF_2F_1 + ip_{1+}p_{2+}$		
$\gamma_z\gamma_x$	0	0	$F_2F_1 + p_{1+}p_{2+}$		
$\gamma_x\gamma_y$	$-2iE$	$-iF_2F_1 - ip_{1+}p_{2-}$	0		
$\gamma_t\gamma_x$	$2ip_y$	$F_2p_{1+} - p_{2-}F_1$	0		
$\gamma_t\gamma_y$	$-2ip_x$	$-iF_2p_{1+} - ip_{2-}F_1$	0		
$\gamma_t\gamma_z$	0	0	$-p_{1+}F_2 - p_{2+}F_1$		
$\gamma_5\gamma_x = \gamma_t\gamma_y\gamma_z$	0	0	$-iF_2F_1 - ip_{1+}p_{2+}$		
$\gamma_5\gamma_y = \gamma_t\gamma_z\gamma_x$	0	0	$F_2F_1 - p_{1+}p_{2+}$		
$\gamma_5\gamma_z = \gamma_t\gamma_x\gamma_y$	$-2im$	$-iF_2F_1 + ip_{1+}p_{2-}$	0		
$\gamma_5\gamma_t = \gamma_x\gamma_y\gamma_z$	0	0	$iF_2p_{1+} + iF_1p_{2+}$		
$\gamma_5 = \gamma_x\gamma_y\gamma_z\gamma_t$	0	0	$iF_2p_{1+} - iF_1p_{2+}$		

Complex conjugate of $s_1 = +1$, $s_2 = +1$ case (column 3)

Negative of complex conjugate of $s_1 = +1$, $s_2 = -1$ case (column 4)

Note: $p_{2+} = p_{2x} + ip_{2y} = p_2 \exp(i\theta_2)$; $p_{2-} = p_{2x} - ip_{2y} = p_2 \exp(-i\theta_2)$; $F_2 = E_2 + m$; $F_1 = E_1 + m$; $p^2 = (E - m)F$.

Part 14: METHODS OF OBTAINING MATRIX ELEMENTS

The matrix element of an operator M between initial state u_1 and final state u_2 will be denoted by

$$(\tilde{u}_2 M u_1)$$

The matrix element is independent of the representations used if they are related by unitary equivalence transformations.

That is,

$$u'_1 = Su_1$$

$$u'_2 = Su_2$$

$$M' = SMS^{-1}$$

$$\tilde{u}'_2 = \tilde{u}_2 \tilde{S}$$

so

$$\tilde{u}'_2 M' u'_1 = \tilde{u}_2 \tilde{S} S \; MS^{-1} S u_1 = \tilde{u}_2 M u_1$$

where the property $\tilde{S} = S^{-1}$ has been assumed for **S**.

The straightforward method to compute the matrix elements is simply to write them out in matrix form and carry out the operations.

In this way the data in Table 13-1 were obtained.

Other methods may be used, however, sometimes simpler and sometimes leading to corollary information, as illustrated by the following example.

By the normalization convention,

$$\bar{u}u = 2m$$

Hence

$$(\bar{u}\not{p}u) = 2m^2$$

since $\not{p}u = mu$.

Similarly,

$$(\bar{u}\gamma_\mu \not{p}u) = m(\bar{u}\gamma_\mu u)$$

But also note that

$$(\bar{u}\not{p}\gamma_\mu u) = m(\bar{u}\gamma_\mu u)$$

because $\bar{u}\not{p} = \not{p}\bar{u} = m\bar{u}$.

Adding the two expressions, one obtains

$$(\bar{u}(\gamma_\mu \not{p} + \not{p}\gamma_\mu)u) = 2m(\bar{u}\gamma_\mu u)$$

It is easy to prove that

$$\not{a}\not{b} = -\not{b}\not{a} + 2a \cdot b$$

so that

$$\not{p}\gamma_\mu + \gamma_\mu \not{p} = 2p_\mu \qquad \gamma_\mu = \not{1}$$

But p_μ is just a number, so it follows that

$$2p_\mu (\bar{u}u) = 2m(\bar{u}\gamma_\mu u)$$

and since $\bar{u}u = 2m$, by normalization

$$(\bar{u}\gamma_\mu u) = 2p_\mu$$

Furthermore, the general relation

$$(\bar{u}\gamma_t u)/(\bar{u}u) = p_4/m = E/m$$

is obtained.

From this it is seen why the possible normalization

$$(\bar{u}\gamma_\mu u) = E/m$$

was equivalent to $(\bar{u}u) = 1$.

Using methods analogous to this one, we can show that

$$(\bar{u}\gamma_5 u) = 0$$

INTERPRETATIONS OF NEGATIVE ENERGY STATES

It was found that a necessary condition for solution of the Dirac equation to exist is

$$E^2 = p^2 + m^2$$

$$E = \pm (p^2 + m^2)^{1/2}$$

The meaning of the positive energy is clear but that of the negative is not.

It was at one time suggested by Schrödinger that it should be arbitrarily excluded as having no meaning.

But it was found that there are two fundamental objections to the exclusion of negative energy states.

The first is physical, theoretically physical, that is.

For the Dirac equation yields the result that starting with a system in a positive energy state there is a probability of induced transitions into negative energy states.

Hence if they were excluded this would be a contradiction.

The second objection is mathematical.

That is, excluding the negative energy states leads to an incomplete set of wave functions.

It is not possible to represent an arbitrary function as an expansion in functions of an incomplete set.

This situation led Schrödinger into insurmountable difficulties.

The positive energy levels form a continuum extending from E = m to +∞, and the negative energies if accepted as such form another continuum from E = -m to -∞. Between +m and -m there are no available energy levels (see Fig. 14-1).

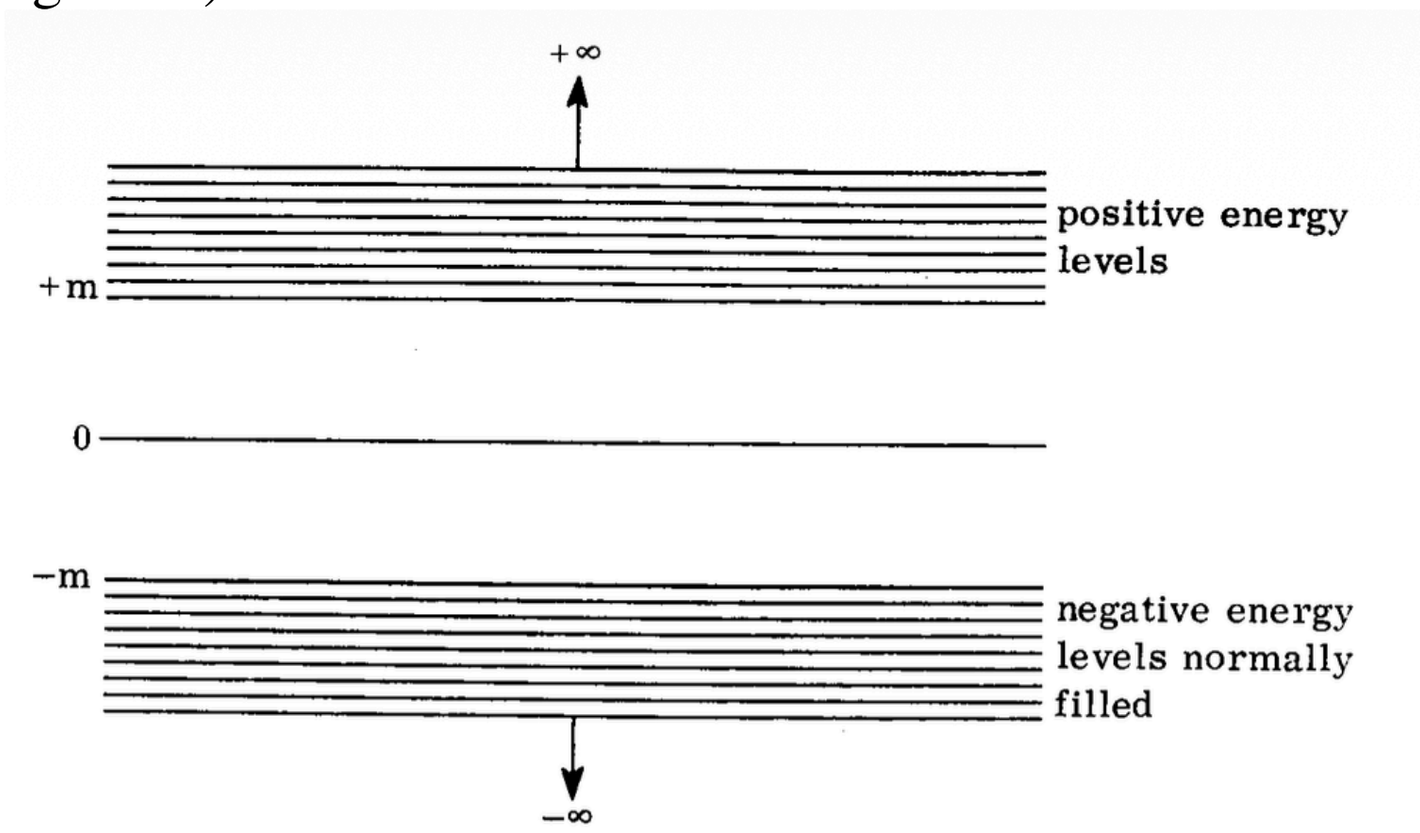


FIG. 14-1

Dirac proposed the idea that all the negative energy levels are normally filled.

Explanations for the apparent obscurity of such a sea of electrons in negative energy states, if it exists, usually contain a psychological aspect and are not very satisfactory.

But, nevertheless, if such a situation is assumed to exist, some of the important consequences are these:

1. Electrons in positive energy states will not normally be observed to make transitions into negative energy states because these states are not available; they are already full.

2. With the sea of electrons in negative energy levels unobservable, a "hole" in it produced by a transition of one of its electrons into a positive energy state should manifest itself. The manifestation of the hole is regarded as a positron and behaves like an electron with a positive charge.

3. The Pauli exclusion principle is implied in order that the negative sea may be full. That is, if any number rather than just one electron could occupy a given state, it would be impossible to fill all the negative energy states. It is in this way that the Dirac theory is sometimes considered as "proof" of the exclusion principle.

Another interpretation of negative energy states has been proposed by Feynman.

The fundamental idea is that the "negative energy" states represent the states of electrons moving backward in time.

In the classical equation of motion

$$m(d^2z_\mu/ds^2) = e(dz_\nu/ds)F_{\mu\nu}$$

reversing the direction of proper time s amounts to the same as reversing the sign of the charge so that the electron moving backward in time would look like a positron moving forward in time.

In elementary quantum mechanics, the total amplitude for an electron to go from x_1,t_1 to x_2,t_2 is computed by summing the amplitudes over a possible trajectories between x_1,t_1 and x_2,t_2, assuming that the trajectories always moved forward in time.

These trajectories might appear in one dimension as shown in Fig. 14-2.

But with the new point of view, a possible trajectory might be as shown in Fig. 14-3.

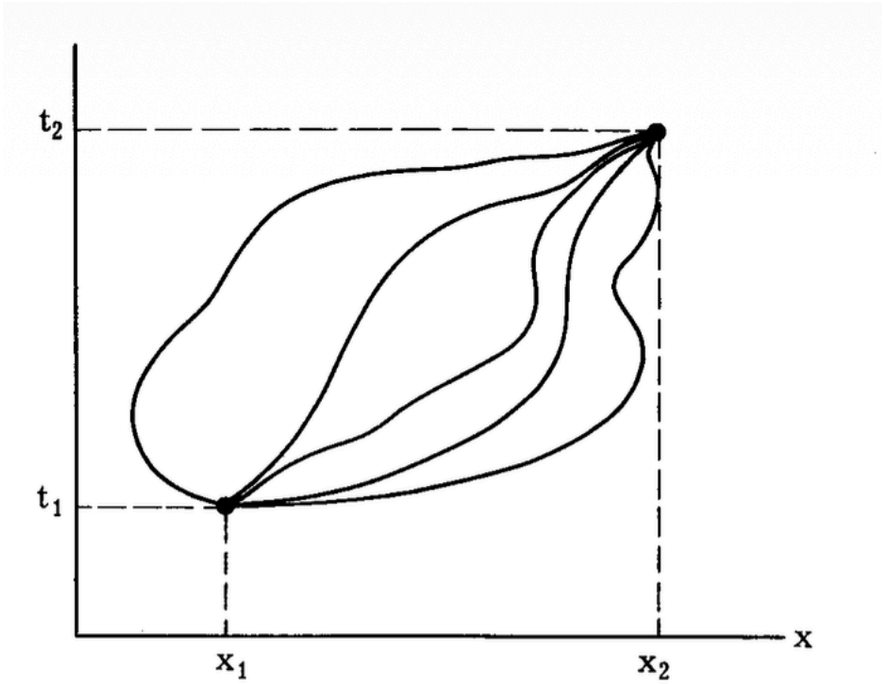


FIG. 14-2

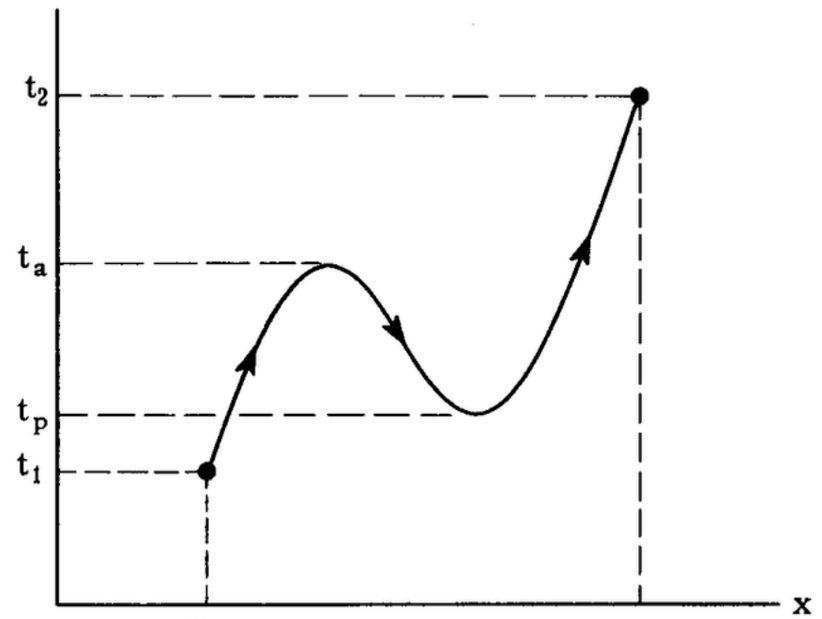


FIG. 14-3

$t_1 \rightarrow t_p$	only the initial electron present
$t_p \rightarrow$	the initial electron still present but somewhere else an electron-positron pair is formed
$t_p \rightarrow t_a$	the initial electron and newly arrived electron and positron are present
$t_a \rightarrow$	the positron meets with the initial electron, both of them annihilating, leaving only the previously created electron
$t_a \rightarrow t_2$	only one electron present

To handle this idea quantum mechanically two rules must be followed:

1. In calculating matrix elements for positrons, the positions of the initial and final wave functions must be reversed. That is, for an electron moving forward in time from a past state Ψ_{past} to a future state Ψ_{fut}, the matrix element is

$$\int \tilde{\Psi}_{fut} M \Psi_{past} \, d\,vol$$

But moving backward in time, the electron proceeds *from* Ψ_{fut} *to* Ψ_{past} so the matrix element for a positron is

$$\int \tilde{\Psi}_{past} M \Psi_{fut} \, d\,vol$$

2. If the energy E is positive, then $e^{-ip\cdot x}$ is the wave function of an electron with energy $p_4 = E$. If E is negative, $e^{-ip\cdot x}$ is the wave function of a positron with energy $-E$ or $|E|$, and of four-momentum $-p$.

Potential Problems in Quantum Electrodynamics

Part 15 PAIR CREATION AND ANNIHILATION

Two possible paths of an electron being scattered between the states Ψ_1 and Ψ_2 were discussed in the last lecture.

These are:

The four cases can be diagrammed as shown in Fig. 15-1.

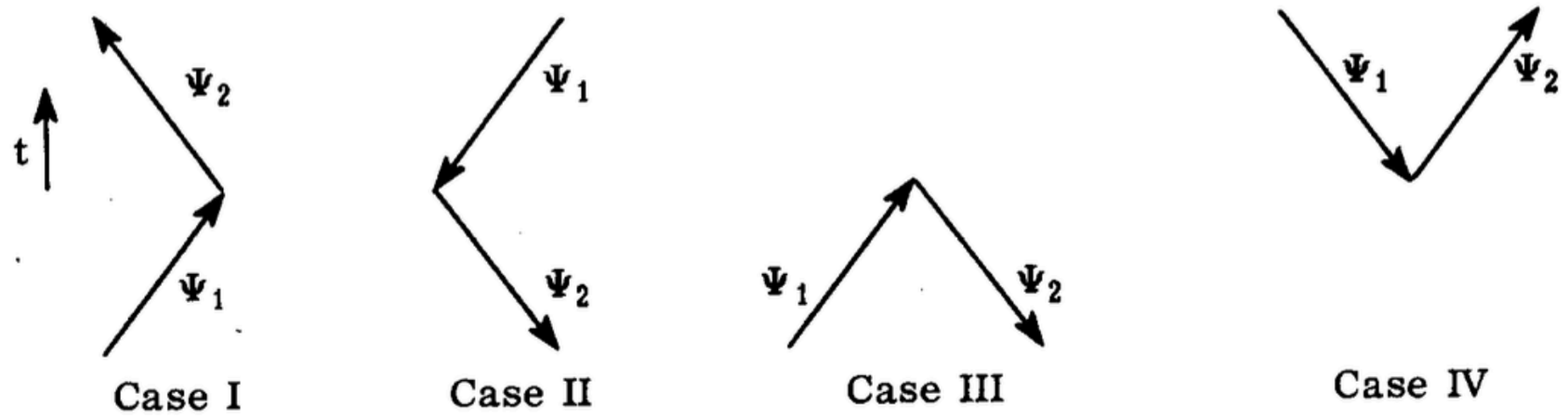


FIG. 15-1

Case I. Both Ψ_1, Ψ_2 states of positive energy, interpreted as Ψ_1 electron in "past," Ψ_2 electron in "future." This is electron scattering.

Case II. Both Ψ_1, Ψ_2 states of negative energy interpreted as Ψ_1 positron in "future," Ψ_2 positron in "past." This is positron scattering.

The existence of negative energy states makes two more types of paths possible. These are:

Case III. The Ψ_1 positive energy, Ψ_2 negative energy, interpreted as Ψ_1 in "past," Ψ_2 positron in "past." Both states are in the past, and nothing in the future. This represents pair annihilation.

Case IV. The Ψ_1 negative energy, Ψ_2 positive energy, interpreted as Ψ_1 positron in "future," Ψ_2 electron in "future." This is pair creation.

the the the the the

Note that in each diagram the arrows point from Ψ_1 to Ψ_2, although time is increasing upward in all cases.

The arrows give the direction of motion of the electron in the present interpretation of negative energy states.

In common language, the arrows point toward positive or negative time according to whether p is positive or negative, that is, whether the state represented is that of an electron or a positron.

CONSERVATION OF ENERGY

Energy relations for the scattering in case I have been established in previous lectures.

It can be seen that identical results hold for case II.

To show this, recall that in case I, if the electron goes from the energy E_1 to E_2 and if the perturbation potential is taken proportional to $e^{-i\omega t}$, then this perturbation brings in a positive energy ω.

To see this, note that the amplitude for scattering is proportional to

$$\int \exp(-iE_2 t)^* \exp(-i\omega t) \exp(-iE_1 t)\, dt$$
$$= \int \exp[(iE_2 t - i\omega t - iE_1 t)\, dt] \qquad (15\text{-}1)$$

As has been shown, there is a resonance between E_2 and $E_1 + \omega$, so that the only contributing energies are those for which $E_2 \approx E_1 + \omega$.

In case I the same integral holds but E_2 and E_1 are negative.

A positron goes from an energy (past) of $E_{past} = - E_2$ to an energy (future) of $E_{fut} = - E_1$.

With the same perturbation energy, the amplitude is large again only if $E_2 \approx E_1 + \omega$ or $-E_{past} = - E_{fut} + \omega$, so that $E_{fut} = \omega + E_{past}$, that is, the perturbation carries in a positive energy ω, just as it does for the electron case.

THE PROPAGATION KERNEL

In the nonrelativistic case (Schrödinger equation), the wave equation, including a perturbation potential, is written

$$i\partial\Psi/\partial t = H_0\Psi + V\Psi \qquad (15\text{-}2)$$

where V is the perturbation potential and H, is the unperturbed Hamiltonian.

For the free particle, the kernel giving the amplitude to go from point 1 to point 2 in space and time can be shown to be

$$\begin{aligned} K_0(2,1) &= N \exp\left[(1/2)im(x_2 - x_1)^2/(t_2 - t_1)\right] \qquad t_2 > t_1 \\ &= 0 \qquad\qquad t_2 < t_1 \end{aligned} \qquad (15\text{-}3)$$

where N is a normalizing factor depending on the time interval $t_2 - t_1$ and the mass of the particle:

$$N = [m/2\pi i(t_2 - t_1)]^{1/2}$$

Note that the kernel is defined to be 0 for $t_2 < t_1$.

It can be shown that K_0 satisfies the equation

$$[i\partial/\partial t_2 - H_0(2)]\, \mathbf{K_0}(2,1) = i\delta(2,1) \tag{15-4}$$

The propagation kernel $\mathbf{K}_v(2,1)$ giving a similar amplitude, but in the presence of the perturbation potential V, must satisfy the equation

$$[i\partial/\partial t_2 - H_0(2) - V(2)]\, \mathbf{K_V}(2,1) = i\delta(2,1) \tag{15-5}$$

It can be shown that $\mathbf{K}_v$ can be computed from the series

$$\mathbf{K_V}(2,1) = \mathbf{K_0}(2,1) - i\int \mathbf{K_0}(2,3)V(3)\mathbf{K_0}(3,1)d^3x_3\, dt_3 \tag{15-6}$$
$$- \int \mathbf{K_0}(2,4)V(4)\mathbf{K_0}(4,3)V(3)\mathbf{K_0}(3,1)d^3x_4\, dt_4\, d^3x_3\, dt_3 + \cdots$$

In case the complete Hamiltonian $H = H_0 + V$ is independent of time, and all the stationary states n of the system are known, then $\mathbf{K}_v(2,1)$ may be obtained from the sum

$$\mathbf{K_V}(2,1) = \sum_n \exp[-iE_n(t_2 - t_1)]\, \phi_n(x_2)\phi_n^*(x_1) \tag{15-7}$$

The extension of these ideas to the relativistic case (Dirac equation) is straightforward.

By choosing a particular form for the Hamiltonian, the Dirac equation can be written

$$i\partial\Psi/\partial t = H\Psi = \boldsymbol{\alpha}\cdot(\mathbf{p} - e\mathbf{A})\Psi + e\phi\,\Psi + m\beta\Psi$$

Defining the propagation kernel as K^A, then the kernel is the solution to the equation

$$[i\partial/\partial t_2 - e\phi_2 - \boldsymbol{\alpha}\cdot(-i\nabla - e\mathbf{A}_2) - m\beta]\,\mathbf{K}^A(2,1) = i\beta\delta(2,1) \tag{15-8}$$

The matrix β is inserted in the last term in order that the kernel derived from the Hamiltonian be relativistically invariant. [Note the similarity to the nonrelativistic case, Eq. (15-6).]

Multiplying this equation by β, a simpler form results:

$$(i\not{\nabla}_2 - e\not{A}_2 - m)\mathbf{K}^A(2,1) = i\delta(2,1) \tag{15-9}$$

The equation for a free particle is obtained simply by letting $\not{A}_2 = 0$, then calling the free-particle kernel K_+,

$$(i\not{\nabla}_2 - m)\,K_+(2,1) = i\delta(2,1) \tag{15-10}$$

The notation K_+ replaces the K_0 of the nonrelativistic case, and Eq. (15-10) replaces Eq. (15-4) as the defining equation.

Just as $\mathbf{K}_v$ can be expanded in the series of Eq. (15-6), so K_+^A can be expanded as

$$\mathbf{K}^A(2,1) = \mathbf{K}_+(2,1) - i\int \mathbf{K}_+(2,3)e\not{A}(3)\mathbf{K}_+(3,1)\,d\tau_3$$
$$- \int \mathbf{K}_+(2,3)e\not{A}(3)\mathbf{K}_+(3,4)e\not{A}(4)\mathbf{K}_+(4,1)d\tau_3\,d\tau_4 + \cdots$$

Note that the kernel is now a four-by-four matrix, so that all components of Ψ we can be determined.

Since this is true, the order of the terms in Eq. (15-11) is important.

The element of integration is actually an element of volume in four-space,

$$d\tau = dx_1\, dx_2\, dx_3\, dx_4$$

The potential, $-ie\not{A}(1)$ can be interpreted as the amplitude per cubic centimeter per second for the particle to be scattered once at the point (1).

Thus the interpretation of Eq. (15-11) is completely analogous to that of Eq. (15-6).

Part 16 USE OF THE KERNEL $K_+(2,1)$

In the nonrelativistic theory it was possible to calculate the wave function at a point x_2 at time t_2 from a knowledge of the wave function at an earlier time t, (see Fig. 16-1) by means of the nonrelativistic kernel $K_0(x_2, t_2; x_1, t_1)$,

$$\Psi(\mathbf{x}_2,t_2) = \int K_0(\mathbf{x}_2,t_2;\ \mathbf{x}_1,t_1)\Psi(\mathbf{x}_1,t_1)\, d^3x_1$$

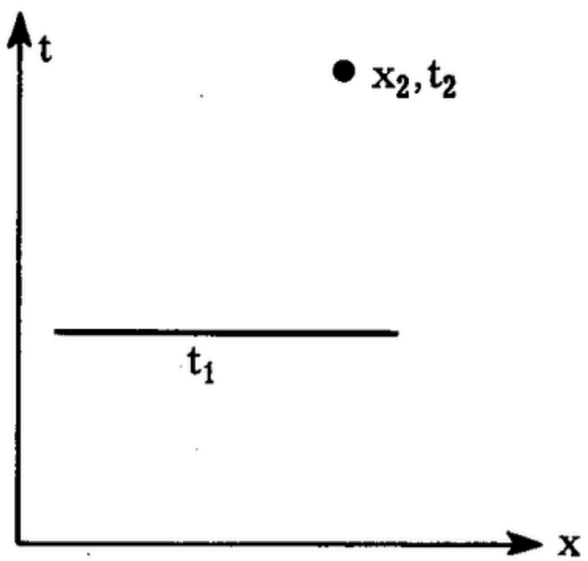


FIG. 16-1

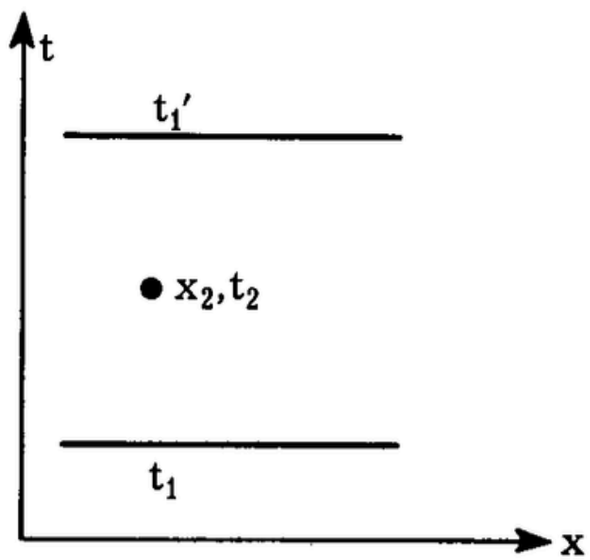


FIG. 16-2

It might be expected that a relativistic generalization of this would be

$$\Psi(\mathbf{x}_2,t_2) = \int K_+(\mathbf{x}_2,t_2;\ \mathbf{x}_1,t_1)\,\gamma_t\,\Psi(\mathbf{x}_1,t_1)\ d^3x_1$$

This turns out to be incorrect, however.

It is not sufficient, in the relativistic case, to know just the wave function at an earlier time only because $K_+(2,1)$ is not zero for $t_2 < t_1$.

When the kernel is defined in this manner (Part 15), the wave function at x_2,t_2, (see Fig. 16-2) is given by

$$\Psi(\mathbf{x}_2,t_2) = \int K_+(\mathbf{x}_2,t_2;\ \mathbf{x}_1,t_1)\,\gamma_t\,\Psi(\mathbf{x}_1,t_1)\,d^3x_1$$
$$- \int K_+(\mathbf{x}_2,t_2;\ \mathbf{x}_1,t_1')\,\gamma_t\,\Psi(\mathbf{x}_1,t_1')\,d^3x_1 \qquad t_1 < t_2 < t_1' \qquad (16\text{-}1)$$

The first term is the contribution from positive energy states at earlier times and the second term is the contribution from negative energy states at later times.

This expression can be generalized to state that it is necessary to know $\Psi(x_1, t_1)$ on a four-dimensional surface, surrounding the point x_2, t_2 (see Fig. 16-3):

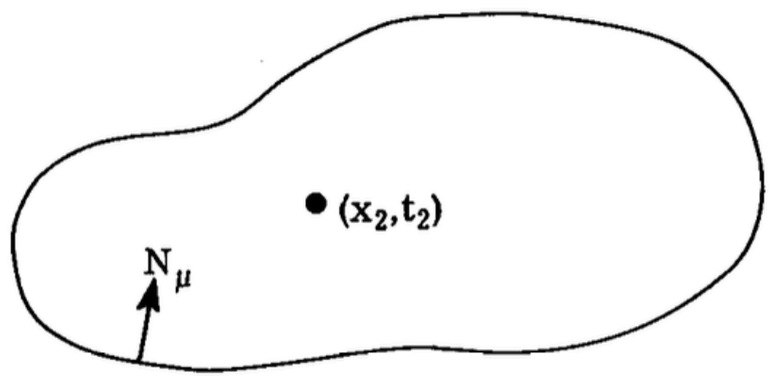


FIG. 16-3

$$\Psi(x_2,t_2) = \int K_+(2,1)\not{N}(1)\Psi(1)\,d^4x_1 \tag{16-2}$$

where $\not{N}_\mu$, is the four-vector normal to the surface that encloses x_2,t_2.

TRANSITION PROBABILITY

The amplitude to go from a state f to a state g under the action of a potential $\not{A}$ is given by an expression similar to that in nonrelativistic theory,

$$a_{21} = \int\int \tilde{g}(2)\beta\, K_+^{\,A}(2,1)\beta\, f(1) d^3x_1\; d^3x_2 \tag{16-3}$$

Using the expansion of $K_+^A(2,1)$ in terms of K(2,1), Eq. (15-12), and assuming that the amplitude for transition from state f to state g as a free particle is zero (f and g orthogonal states), the first-order amplitude for transition (Born approximation) is

$$a_{21} = -i\int\tilde{g}(2)\;\beta\int K_+(2,3)e\not{A}(3)\,K_+(3,1)\beta f(1)\,d\tau_3\;d^3x_1\;d^3x_2$$

It is convenient to let

$$f(3) = \int K_+(3,1)\beta f(1)\;d^3x_1$$

$$\tilde{g}(3) = \int \tilde{g}(2)\beta K_+(2,3)\,d^3x_2$$

These state that the particle has the free-particle wave function f just prior to scattering and the free-particle wave function g just after scattering, and that it eliminates any computation of the motion as a free particle.

The amplitude for transition, to first order, may be written

$$-i \int \tilde{g}(3) e \not{A}(3) f(3)\, d\tau_3 \tag{16-4}$$

($d\tau_3$ signifies integration over time as well as space).

The second-order term would be written

$$-(1/2) \int\int g(4) e \not{A}(4) K_+(4,3) e \not{A}(3) f(3)\, d\tau_3\, d\tau_4$$

If {(3) is a negative energy state, then it represents a positron of the future instead of an electron of the past and the process described by this amplitude is that of pair production.

SCATTERING OF AN ELECTRON FROM A COULOMB POTENTIAL

We shall make use of the theory just presented to calculate the scattering of an electron from an infinitely heavy nucleus of charge Ze.

Suppose the incident electron has momentum in the x direction and the scattered electron has momentum in the xy plane (see Fig. 16-4):

$$\not{p}_1 = \gamma_t E_1 - \gamma_x p_{1x}$$

$$\not{p}_2 = \gamma_t E_z - \gamma_x p_{2x} - \gamma_y p_{2y}$$

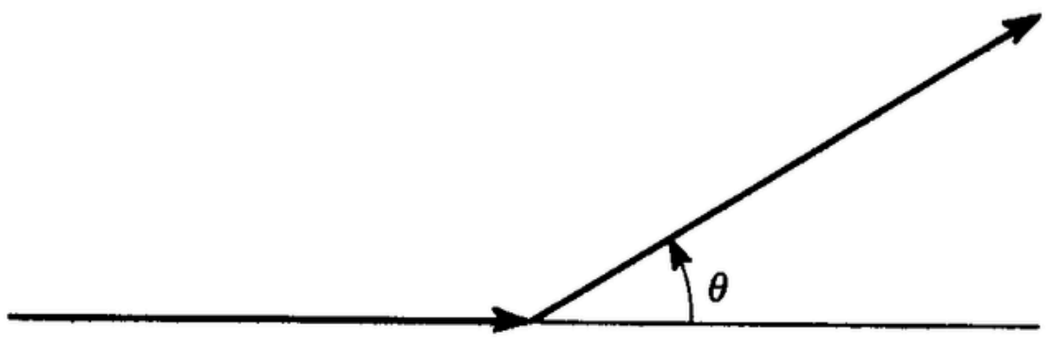


FIG. 16-4

The potential is that of a stationary charge Ze,

$$\phi = Ze/r, \quad \mathbf{A} = 0 \qquad \not{A} = \gamma_t (Ze/r)$$

The initial and final wave functions are plane waves:

$$f(1) = u_1 e^{-ip_1 \cdot x} \qquad g(2) = u_2 e^{-ip_2 \cdot x}$$ (four-component wave function)

Thus, by Eq. (16-4), the first-order amplitude for transition from state f to state g (momentum p_1 to momentum p_2) is

$$M = -i \int \tilde{u}_2 e^{ip_2 \cdot x} (Ze^2/r) \gamma_t u_1 e^{-ip_1 \cdot x} \, d^3x \, dt$$

Separating space and time dependence in the wave functions, this becomes

$$M = -i(\tilde{u}_2 \gamma_t u_1) \left[\int e^{-ip_2 \cdot x} (Ze^2/r) e^{ip_1 \cdot x} \, d^3x \right] \left[\int_0^T e^{iE_2 t} e^{-iE_1 t} \, dt \right]$$

The first integral is just $V(\mathbf{Q})$, a three-dimensional Fourier transform of the potential, which was evaluated in nonrelativistic scattering theory:

$$M = -i(\tilde{u}_2 \gamma_t u_1) [V(\mathbf{Q})] \left\{ \frac{\exp [i(E_2 - E_1)t] - 1}{i(E_2 - E_1)} \right\} \tag{16-5}$$

$$V(\mathbf{Q}) = 4\pi Ze^2/Q^2 \qquad \mathbf{Q} = \mathbf{p}_1 - \mathbf{p}_2$$

The probability of transition per second is given by

$$\text{Trans. prob./sec} = 2\pi(\Pi N)^{-1}|M|^2 \times (\text{density of final states}) \qquad (16\text{-}6)$$

This is a result from time-dependent perturbation theory, the only new factor is a normalizing factor $(\Pi N)^{-1}$ which takes account for the fact that the wave functions are not normalized to unity per unit volume.

The ΠN is a product of factors N one for each wave function, or particle in the initial state, and one for each final wave function,

$$N = (\tilde{u}\gamma_t u) \qquad (16\text{-}7)$$

for each particle in question.

In our normalization, then N = 2E.

The reason for this factor is that wave functions are normalized to

$$(\tilde{u}\,u) = 2m \qquad \text{or} \qquad (\tilde{u}\gamma_t u) = 2E$$

where, as in the computation of transition probability, they should be normalized in the conventional nonrelativistic manner $\Psi^*\Psi = 1$ or $(\tilde{u}\gamma_t u) = 1$ (so N = 1 for that case).

The matrix element M, as calculated in this manner, is relativistically invariant and in the future the chief interest will be in M.

The transition probability, knowing M, can be computed from Eq. (16-6).

Density of States, Cross Section.

For the electron scattering problem under consideration,

$$M = -i(\tilde{u}_2 \gamma_t u_1)(4\pi Ze^2/Q^2)$$

so the transition probability is

$$\text{Trans. prob./sec} = \frac{2\pi}{(2E_1)(2E_2)} |(\tilde{u}_2 \gamma_t u_1)|^2 \left| \frac{4\pi Ze^2}{Q^2} \right|^2 \frac{E_2 p_2 d\Omega}{(2\pi)^3} \qquad (16\text{-}8)$$

where the density of final states has been obtained in the following manner:

$$\text{Density of states} = \frac{d^3 \mathbf{p}_2}{(2\pi)^3 \, dE_2} = \frac{p_2^2 \, dp_2 \, d\Omega}{(2\pi)^3 \, dE_2} \qquad \hbar = 1$$

but $E_2^2 = p_2^2 + m^2$, so $dp_2/dE_2 = E_2/p_2$ and

$$\text{Density of states} = E_2 p_2 d\Omega/(2\pi)^3$$

When the incoming plane wave is normalized to one particle per cubic centimeter, the cross section is given in terms of the transition probability per second as

$$\text{Trans. prob./sec} = \sigma v_1 = \sigma (p_1/E_1)$$

† $p_1 = \frac{mv_1}{(1-v_1^2)^{1/2}} \rightarrow p_1^2 = \frac{m^2 v_1^2}{1-v_1^2} \rightarrow p_1^2 = (m^2 + p_1^2)v_1^2 = E_1^2 v_1^2$

Therefore, $v_1 = p_1/E_1$.

or

$$\sigma = (E_1/p_1) \times (\text{trans. prob./sec})$$

The essential difference between the relativistic treatment of scattering and the nonrelativistic treatment is contained in the matrix element .

From Table 13-1, for a particle moving in the xy plane and $s_1 = +1$, $s_2 = +1$,

$$|(\tilde{u}_2 \gamma_t u_1)|^2 = 1/F_1F_2|F_2F_1 + p_{1+}p_{2-}|^2$$

where

$$F_1 = F_2 = E + m$$

[$E_1 = E_2$, conservation of energy, follows from the nature of the time integral in Eq. (16-5), and

$$p_{1+} = p$$

$$p_{2-} = pe^{-i\theta}$$

(magnitude of final momentum equal to magnitude of initial momentum follows from $E_1 = E_2$).

Thus

$$|(\tilde{u}_2 \gamma_t u_1)|^2 = (E + m)^{-2} |(E+m)^2 + p^2 e^{-i\theta}|^2$$

$$= (E + m)^{-2} \{4E^2(E+m)^2[1 - (p^2/E^2) \sin^2 (\theta/2)]\}$$

$$= (2E)^2 [1 - v^2 \sin^2 (\theta/2)]$$

When $s_1 = +1, s_2 = -1$ or $s_1 = -1, s_2 = +1$, the matrix element of γ_5 is zero.

When $s_1 = -1, s_2 = -1$, the absolute value of the matrix element is the same as for $s_1 = +1, s_2 = +1$.

Thus spin does not change in scattering (in Born approximation) and the cross section is independent of spin,

$$\sigma = (4Z^2e^4E^2/Q^4)d\Omega\,[1 - v^2 \sin^2(\theta/2)] \qquad Q = 2p \sin(\theta/2)$$

The criterion for validity of the Born approximation, used in obtaining this result, is $Ze^2/\hbar v << 1$.

In the extreme relativistic limit $v \approx c$, this becomes $Z < 137$.

Just as for the nonrelativistic case, the scattering can actually be calculated exactly (correct to all orders in the potential) for the Coulomb potential.

This exact solution of the Dirac equation involves hypergeometric functions.

It was first worked out by Mott and is called Mott scattering.

For moderate energies (200 kev) there is some probability for change in spin.

Polarized electrons could be produced in this manner.

Part 17 CALCULATION OF THE PROPAGATION KERNEL FOR A FREE PARTICLE

As shown in a previous lecture, the propagation kernel, when there is no perturbing potential and the Hamiltonian of the system is constant in time, is

$$K_+(2,1) = \sum_{+n} \phi_n(\mathbf{x}_2)\,\tilde{\phi}_n(\mathbf{x}_1)\,\exp[-iE_n(t_2 - t_1)] \qquad t_2 > t_1$$

$$= -\sum_{-n} \phi_n(\mathbf{x}_2)\,\tilde{\phi}_n(\mathbf{x}_1)\,\exp[-iE_n(t_2 - t_1)] \qquad t_2 < t_1$$

For a free particle, the eigenfunctions ϕ_n are

$$u_p \exp(i\mathbf{p}\cdot\mathbf{x})$$

and the sum over n becomes an integral over p.

The u_p is the spinor corresponding to momentum p, positive or negative energy and spin up or down, as appropriate.

Then the propagation kernel for a free particle is, for $t_2 > t_1$,

$$K_+(2,1) = \sum_{\text{spins}} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2E_p}\, u_p \tilde{u}_p \exp[i\mathbf{p}\cdot(\mathbf{x}_2 - \mathbf{x}_1)]$$

$$\times \exp[-iE_p(t_2 - t_1)]$$

for $E_p = +(p^2 + m^2)^{1/2}$.

The factor $1/(2\pi)^3$ is the density of states per unit volume of momentum space per cubic centimeter.

The factor $1/2E_p$ arises from the normalization $u\tilde{u} = 2m$ or $u\gamma_t\tilde{u} = 2E_p$ used here.

The u_p are those for positive energy.

For negative energy $E_p = -(p^2 + m^2)^{1/2}$, the u_p are changed accordingly and $K_+(2, 1)$ becomes, for $t_2 < t_1$,

$$K_+(2,1) = -\sum_{\text{spins}} \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} u_p \tilde{u}_p \exp[ip \cdot (x_2 - x_1)]$$

$$\times \exp[iE_p(t_2 - t_1)]$$

The calculation will be made first for the case of $t_2 > t_1$.

We first calculate $u_p\tilde{u}_p$ for positive energy, and p in the xy plane and spin up.

Under these conditions

$$u_p = \begin{pmatrix} E + m \\ 0 \\ 0 \\ p_x + ip_y \end{pmatrix} \frac{1}{(E+m)^{1/2}}$$

$$\tilde{u}_p = \overbrace{E+m \quad 0 \quad 0 \quad -p_x + ip_y} \frac{1}{(E + m)^{1/2}}$$

Note that $u_p\tilde{u}_p$ is the opposite order to that usually encountered so that the product is a matrix, not a scalar.

That is,

$$u_p\tilde{u}_p = \begin{pmatrix} (E+m)^2 & 0 & 0 & (E+m)(-p_x+ip_y) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ (E+m)(p_x+ip_y) & 0 & 0 & (p_x+ip_y)(-p_x+ip_y) \end{pmatrix}$$

$$\times \; 1/(E+m)$$

by the usual rules for matrix multiplication.

But

$$(p_x + ip_y)(-p_x + ip_y) = -p^2 = -E^2 + m^2$$

and the matrix becomes

$$u_p\tilde{u}_p = \begin{pmatrix} E+m & 0 & 0 & -p_x+ip_y \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ p_x+ip_y & 0 & 0 & -E+m \end{pmatrix} \quad \textbf{(spin up)}$$

By the same process, the result in the spin down case is

$$u_p = \begin{pmatrix} 0 \\ E+m \\ p_x - ip_y \\ 0 \end{pmatrix} \frac{1}{(E+m)^{1/2}}$$

$$\tilde{u}_p = \overbrace{0 \qquad E-m \qquad -p_x - ip_y \qquad 0} \quad \frac{1}{(E+m)^{1/2}}$$

$$u_p\tilde{u}_p = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & E+m & -p_x-ip_y & 0 \\ 0 & p_x-ip_y & -E+m & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \textbf{(spin down)}$$

It may be verified easily that the sum of these matrices for spin up and spin down is represented by

$$E\gamma_t - p_x\gamma_x - p_y\gamma_y + m$$

In the general case when **p** is in any direction, it is clear that the only change is an additional term $-p_z\gamma_z$.

So, in general,

$$(u_p\tilde{u})_{\text{spin up}} + (u_p\tilde{u}_p)_{\text{spin down}} = E\gamma_t - \mathbf{p}\cdot\boldsymbol{\gamma} + m = \not{p} + m$$

The sign of the energy was not used in obtaining this result so it is the same for either sign.

Now put $t_2 - t_1 = t$ and $\mathbf{x}_2 - \mathbf{x}_1 = \mathbf{x}$.

For $t > 0$, the propagation kernel becomes

$$K_+(2,1) = \int (E_p\gamma_t - \mathbf{p}\cdot\boldsymbol{\gamma} + m)[d^3\mathbf{p}/(2\pi)^3](1/2E_p)$$

$$\times \exp[-i(E_p t - \mathbf{p}\cdot\mathbf{x})]$$

The appearance of p in the form $E_p = (p^2 + m^2)^{1/2}$ in the time part of the exponential makes this a difficult integral.

Note that it may also be written in the form

the the

$$K_+(2,1) = (i\gamma_t \frac{\partial}{\partial t} + i\gamma_x \frac{\partial}{\partial x} + i\gamma_y \frac{\partial}{\partial y} + i\gamma_z \frac{\partial}{\partial z} + m)$$

$$\times \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_p} \exp[-i(E_p t - \mathbf{p}\cdot\mathbf{x})]$$

$$= i(i\not{\nabla} + m)\, I_+(t,\mathbf{x})$$

where

$$I_+(t,\mathbf{x}) = -i \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_p} \exp[-i(E_p t - \mathbf{p}\cdot\mathbf{x})]$$

In this form only one integral instead of four need be done.

It may be verified as an exercise that for $t < 0$ the result is the same except that the sign of t is changed, so that putting $|\,t\,|$ in place of t in the formula for $I_+(t\ ,x)$ makes it good for all t.

This integral has been carried out with the following result:

$$I_+(t,\mathbf{x}) = -(4\pi)^{-1}\delta(s^2) + (m/8\pi s)H_1^{(2)}(ms)$$

where $s = +(t^2 - x^2)^{1/2}$ for $t > x$, and $- i(x^2 - t^2)^{1/2}$ for $t < x$.

$\delta(s^2)$ is a delta function and $H_1^{(2)}(ms)$ is a Hankel function.

Another expression for the foregoing is

$$I_+(t,\mathbf{x}) = -(1/8\pi^2) \int_0^\infty d\alpha \ \exp\{-(i/2)[(m^2/\alpha) + \alpha(t^2 - x^2)]\}$$

Both of these forms are too complicated to be of much practical use.

It will be shown shortly that a tremendous simplification results from transformation to momentum representation.

Note that $I_+(t,x)$ actually depends only on |x|, not on its direction.

In the time-space diagram (Fig. 17-1) the space axis represents |x| and the diagonal lines represent the surface of a light cone including the t axis, that is, the accessible region 0f the t - |**x**| space in the ordinary sense.

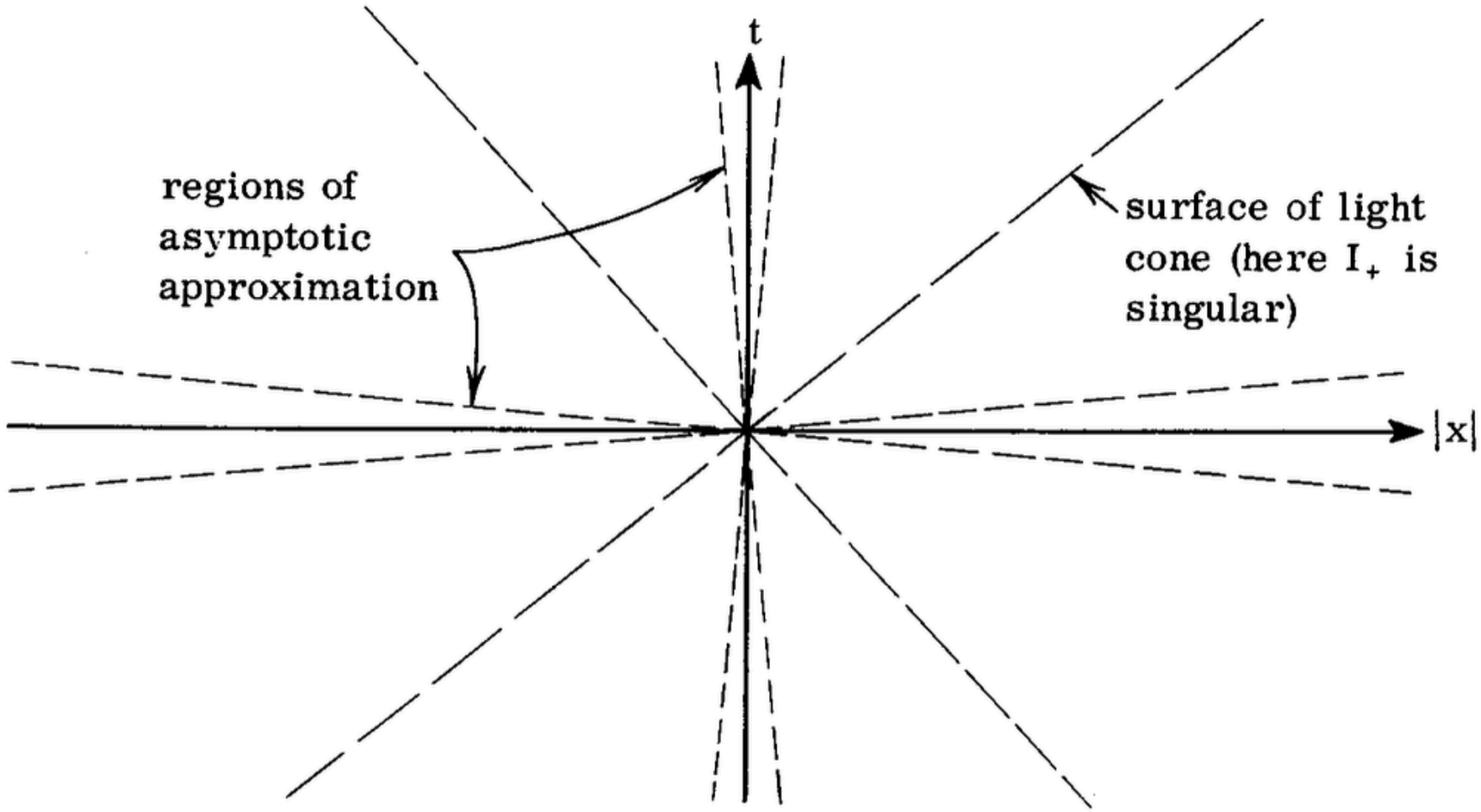


FIG. 17-1

It can be shown that the asymptotic form of $I_+(t,x)$ for large s is proportional to e^{-ims}.

When one's region of accessibility is limited to the inside of the light cone, large s implies $t^2 >> |\mathbf{x}|^2$, so that the region of the asymptotic approximation lies roughly within the dotted cone around the t axis and is

$$I_+(t,\mathbf{x}) \rightarrow e^{-ims} \approx \exp\{-im[t - (x^2/2t)]\} \approx e^{-imt}$$

The first form is seen to be essentially the same as the propagation kernel for a free particle used in nonrelativistic theory.

If, as in the new theory, possible "trajectories" are not limited to regions within the light cone, another region included in this asymptotic approximation is that within the dotted cone along the $|\mathbf{x}|$ axis where large s implies $|\mathbf{x}|^2 >> t^2$.

Hence

$$I_+(t,\mathbf{x}) \rightarrow e^{-ims} = \exp[-im(x^2 - t^2)^{1/2}] \approx e^{-m|\mathbf{x}|}$$

It is seen that the distance along $|\mathbf{x}|$ in which this becomes small is roughly the Compton wavelength (recall that $m \rightarrow mc/\hbar$ when it represents a length^{-1} as here), so that in reality not much of the t - $|\mathbf{x}|$ space outside the light cone is accessible.

The transformation to momentum representation will now be made.

This is facilitated by use of the integral formula

$$\lim_{\epsilon \to 0} 0 \int_{-\infty}^{\infty} dp_4 \frac{\exp(-ip_4 t)}{p_4^2 - E_p^2 + i\epsilon} = -\frac{\pi i}{E_p} \exp(-iE_p|t|)$$

The $i\epsilon$ term in the denominator is introduced solely to ensure passage around the proper side of the singularities at $p_4^2 = E_p^2$ along the path of integration.

Passage on the wrong side will reverse the sign in the exponential on the right.

Evaluation of the integral requires complex contour integration for those interested.

Part 18 MOMENTUM REPRESENTATION

Since the propagation kernel for a free particle is so simply expressed in momentum representation,

$$k(p) = i/(\not{p} - m)$$

it will be convenient to convert all our equations to this representation.

It is especially useful for problems involving free, fast, moving particles.

This requires four-dimensional Fourier transforms.

To convert the potential, define

$$\not{A}(q) = \int \not{A}(x) \exp(iq \cdot x)\, d^4x \qquad (18\text{-}1)$$

Then the inverse transform is

$$\not{A}(x) = (1/2\pi)^4 \int \not{A}(q) \exp(-iq \cdot x)\, d^4q \qquad (18\text{-}2)$$

The function a(q) is interpreted as the amplitude that the potential contains the momentum (q).

As an example, consider the Coulomb potential, given by $\mathbf{A} = 0$, $\varphi = Ze/r$.

Substituting into Eq. (18-1) gives

$$\not{A}(q) = 4\pi Ze/(\mathbf{Q}\cdot\mathbf{Q})\delta(q_4)\gamma_t$$

Here the vector **Q** is the space part of the momentum.

The delta function $\delta(q_4)$ arises from the time dependence of $\not{A}(x)$.

Matrix Elements. An advantage of momentum representation is the simplicity of computing matrix elements.

Recall that in space representation the first-order perturbation matrix element is given by the integral

$$M = -i\int \tilde{g}(2)e\not{A}(2)f(1)\,d\tau_2$$

For the free particle, this becomes

$$M = -i\int \tilde{u}_2 \exp(ip_2\cdot x_2)e\not{A}(2)u_1 \exp(-ip_1\cdot x_1)\,d\tau_2 \qquad (18\text{-}3)$$

In momentum representation, this is simply

$$M = i(\tilde{u}_2 e\not{a}(q)u_1) \qquad (18\text{-}3')$$

where $\not{q}$ is defined analogously to the three-vector q,

$$\not{q} = \not{p}_2 - \not{p}_1$$

The second-order matrix element in space representation is given by

$$-\int\int \tilde{g}(2)e\not{A}(2)\overline{K}_+(2,1)e\not{A}(1)f(1)\, d\tau_1\, d\tau_2$$

Substituting for a free particle and also expressing the potential functions as their Fourier transforms by means of Eq. (18-2), this becomes

$$-\iiiint \tilde{u}_2 \exp(ip_2\cdot x_2)e\not{a}(q_2)\exp(-iq_2\cdot x_2)K_+(2,1)e\not{a}(q_1)$$

$$\times \exp(-iq_1\cdot x_1)\, u_1 \exp(-ip_1\cdot x_1)d\tau_1\, d\tau_2\cdot d^4q_1/(2\pi)^4 \qquad (18\text{-}4)$$

$$\times\ d^4q_2/(2\pi)^4$$

If Eq. (18-2) is used for $K_+(2,1)$, this kernel can be written

$$K_+(2,1) = \int i/(\not{p} - m)\exp[-ip\cdot(x_2 - x_1)]\, d^4p/(2\pi)^4$$

Writing the factors that depend on τ_1, this part of the integral is

$$\int \exp(ip\cdot x_1)\exp(-iq_1\cdot x_1)\exp(-ip_1\cdot x_1)\, d\tau_1$$

$$= (2\pi)^4\delta^4(p - q_1 - p_1) \qquad (18\text{-}5)$$

where the function $\delta^4(x)$ is to be interpreted as $\delta(t_1)\delta(x_2)\delta(y_3)\delta(x_4)$.

Then the integral over τ_1, is zero for all $\not{p}$ except $\not{p} = \not{p}_1 + \not{q}_1$.

So the integral over p reduces Eq. (18-4) to

$$-\iiiint \tilde{u}_2 \exp(ip_2 \cdot x_2) e\not{A}(q_2) \exp(-ip_2 \cdot x_2) \exp[-i(p_1 + q_1) \cdot x_2]$$

$$\times i(\not{p}_1 + \not{q}_1 - m)^{-1} e\not{A}(q_1) u_1 \, d\tau_2 \, d^4q_1/(2\pi)^4 \, d^4q_2/(2\pi)^4$$

Integrating over τ_2 results in another δ function [similar to Eq. (18-5)], which differs from zero only when

$$\not{p}_2 - \not{q}_2 = \not{p}_1 + \not{q}_1$$

Then integrating over d^4q_2 gives finally

$$(-i^2)i \int \tilde{u}_2 \, e\not{A}(q_2)(\not{p}_1 + \not{q}_1 - m)^{-1} e\not{A}(q_1) u_1 \, d^4q_1/(2\pi)^4$$

These results can be written down immediately by inspection of a diagram of the interaction (see Fig. 18-1).

The electron enters the region at 1 with wave function u_1 and moves from 1 to 3 as a free particle of momentum $\not{p}_1$.

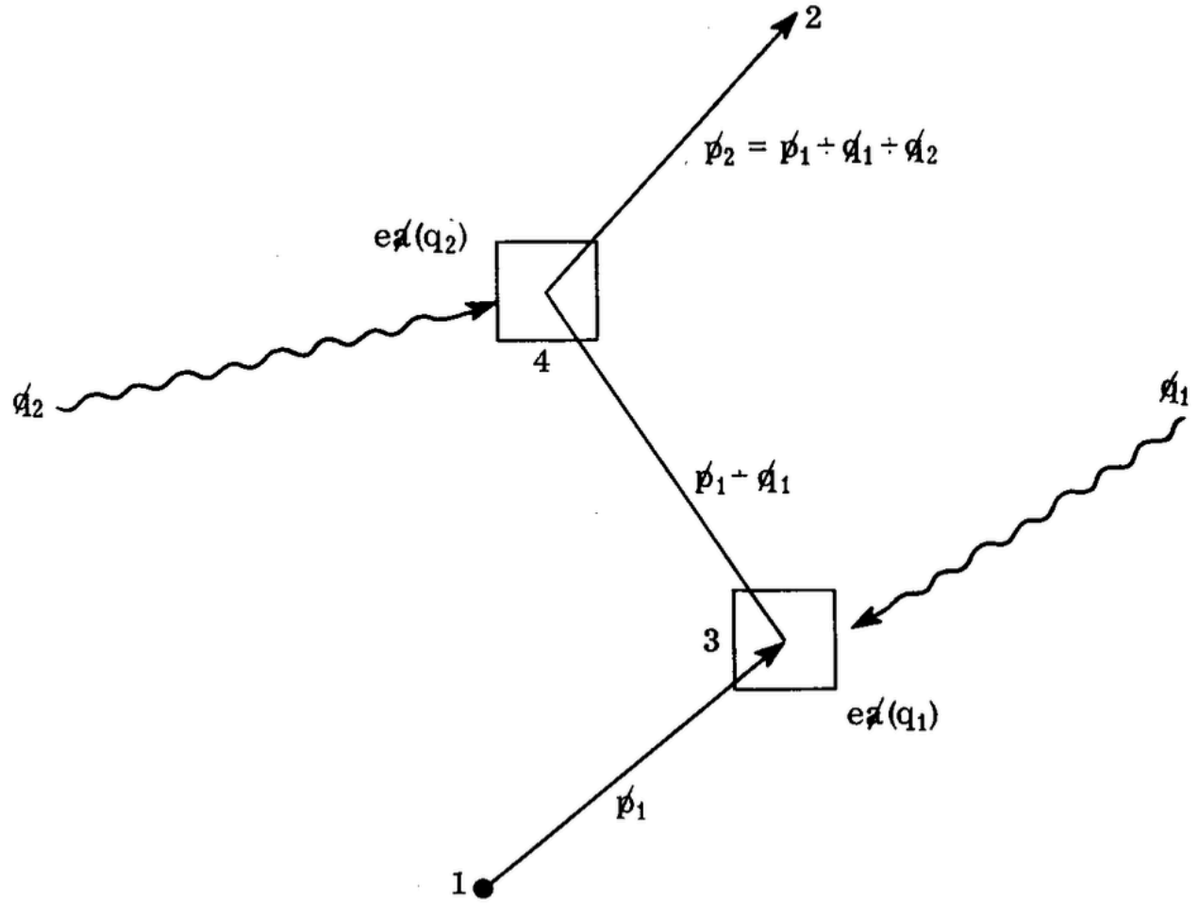


FIG. 18-1

At point 3, it is scattered by a photon of momentum $\not{q}_1$ [under the action of the potential $-ie\not{A}(q_1)$].

Having absorbed the momentum of the photon it then moves from 3 to 4 as a free particle of momentum $\not{p}_1 + \not{q}_1$, by conservation of momentum.

At point 4, it is scattered by a second photon of momentum $\not{q}_2$ [under the action of the potential $-ie\not{A}(q_2)$ absorbing the additional momentum $\not{q}_2$].

Finally, it moves from 4 to 2 as a free particle with wave function u_2 and momentum $\not{p}_2 = \not{p}_1 + \not{q}_1 + \not{q}_2$.

It is also clear from the diagram that the integral needs to be taken over q_1 only, because when $\not{p}_1$ and $\not{p}_2$ are given, $\not{q}_2$ is determined by $\not{q}_2 = \not{p}_2 - \not{p}_1 - \not{q}_1$.

The law of conservation of energy requires $p_1^2 = m^2$, $p_2^2 = m^2$; but, since the intermediate state is a virtual state, it is not necessary that $(\not{p}_1 + \not{q}_1)^2 = m^2$.

Since the operator $1/(\not{p}_1 + \not{q}_1 - m)$ may be resolved as $(\not{p}_1 + \not{q}_1 + m)/[(\not{p}_1 + \not{q}_1)^2 - m^2]$, , the importance of a virtual state is inversely proportional to the degree to which the conservation law is violated.

The results given in Eqs. (18-3') and (18-6) may be summarized by the following list of handy rules for computing the matrix element $M = (\tilde{u}_2 N u_1)$:

the the

1. An electron in a virtual state of momentum $\not{p}$ contributes the amplitude $i/(\not{p} - m)$ to N.

2. A potential containing the momentum q contributes the amplitude $-ie\not{A}(q)$ to N.

3. All indeterminate momenta q_i are summed over $d^4q_i/(2\pi)^4$.

Remember, in computing the integral, the value of the integral is desired, with the path of integration passing the singularities in a definite manner.

Thus, replace m by m - $i\epsilon$ in the integrand, then in the solution take the limit as $\epsilon \rightarrow 0$.

For relativistic work, only a few terms in the perturbation series are necessary for computation.

To assume that fast electrons (and positrons) interact with a potential only once (Born approximation) is often sufficiently accurate.

After the matrix element is determined, the probability of transition per second is given by

$$P = 2\pi/(\Pi N)|M|^2 \times \text{(density of final states)}$$

where ΠN is the normalization factor defined in Part 16.

Part 19 Relativistic Treatment of the Interaction of Particles with Light

In Part 2 the rules governing nonrelativistic interaction of particles with light were given.

The rules stated what potentials were to be used ir the calculation of transition probabilities by perturbation theory.

Those potentials are also applicable to the relativistic theory if the matrix elements are computed as described in Part 18.

For absorption of a photon, the potential used in nonrelativistic theory was

$$A_\mu = (4\pi e^2)^{1/2} (2\omega)^{-1/2} e_\mu \exp(ik \cdot x) \qquad \begin{cases} K_4 = \omega \\ K \cdot K = 0 \\ \hbar = c = 1 \end{cases} \tag{19-1}$$

For emission of a photon, the complex conjugate of this expression is used.

These potentials are normalized to one photon per cubic centimeter and hence the normalization is not invariant under Lorentz transformations.

In a manner similar to that for the normalization of electron wave functions, photon potentials will, in the future, be normalized to 2ω photons per cubic centimeter by dropping the $(2\omega)^{-1/2}$ factor in Eq. (19-1), giving

$$A_\mu = (4\pi e^2)^{1/2} e_\mu \exp(ik \cdot x) \tag{19-1'}$$

This makes any matrix element computed with these potentials invariant, but to obtain the correct transition probability in a given coordinate system, it is necessary to reinsert a factor $(2\omega)^{-1}$ for each photon in the initial and final states.

This becomes part of the normalization factor ΠN, which contains a similar factor for each electron in the initial and final states.

In the momentum representation, the amplitude to absorb (emit) a photon of polarization e_μ is $-i(4\pi e^2)^{1/2} \not{e}$.

The polarization vector e_μ is a unit vector perpendicular to the propagation vector.

Hence : $e \cdot e = -1$ and $e \cdot q = 0$.

RADIATION FROM ATOMS

The transition probability per second is

$$\text{Trans. prob./sec} = 2\pi |H|^2 \times (\text{density of final states})$$

where H is the matrix element of the relativistic Hamiltonian,

$$H = \alpha \cdot (-i\nabla - eA) \qquad \text{S.R.}$$

between initial and final states.

That is,

$$\langle f|H|i\rangle = (4\pi e^2)^{1/2} \int \Psi_f^* [\boldsymbol{\alpha} \cdot \mathbf{e} \exp(i\mathbf{k} \cdot \mathbf{x})] \Psi_i \, d\,\text{vol} \tag{19-2}$$

In the nonrelativistic limit, Eq. (19-2) reduces to

$$1/2m \int \Psi_f^* [\mathbf{e} \cdot \mathbf{p} \exp(i\mathbf{k} \cdot \mathbf{x}) + \exp(i\mathbf{k} \cdot \mathbf{x}) \mathbf{p} \cdot \mathbf{e} + \mathbf{e} \cdot (\boldsymbol{\sigma} \times \mathbf{k})$$

$$\times \exp(i\mathbf{k} \cdot \mathbf{x})] \Psi_i \, d\,\text{vol}$$

This is the same result as was obtained from the Pauli equation.

SCATTERING OF GAMMA RAYS BY ATOMIC ELECTRONS

A relativistic treatment of scattering of photons from electrons will now be given.

As an approximation, consider the electrons to be free (energies at which a relativistic treatment is necessary are, generally, much greater than atomic binding energies).

This will lead to the Klein-Nishina formula for the Compton-effect cross section.

For the incoming photon take as a potential $A_{1\mu} = e_{1\mu} \exp(-iq_1 \cdot x)$ and for the outgoing photon take $A_{2\mu} = e_{2\mu} \exp(-iq_2 \cdot x)$.

The light is polarized perpendicular to the direction of propagation (see Fig. 19-1).

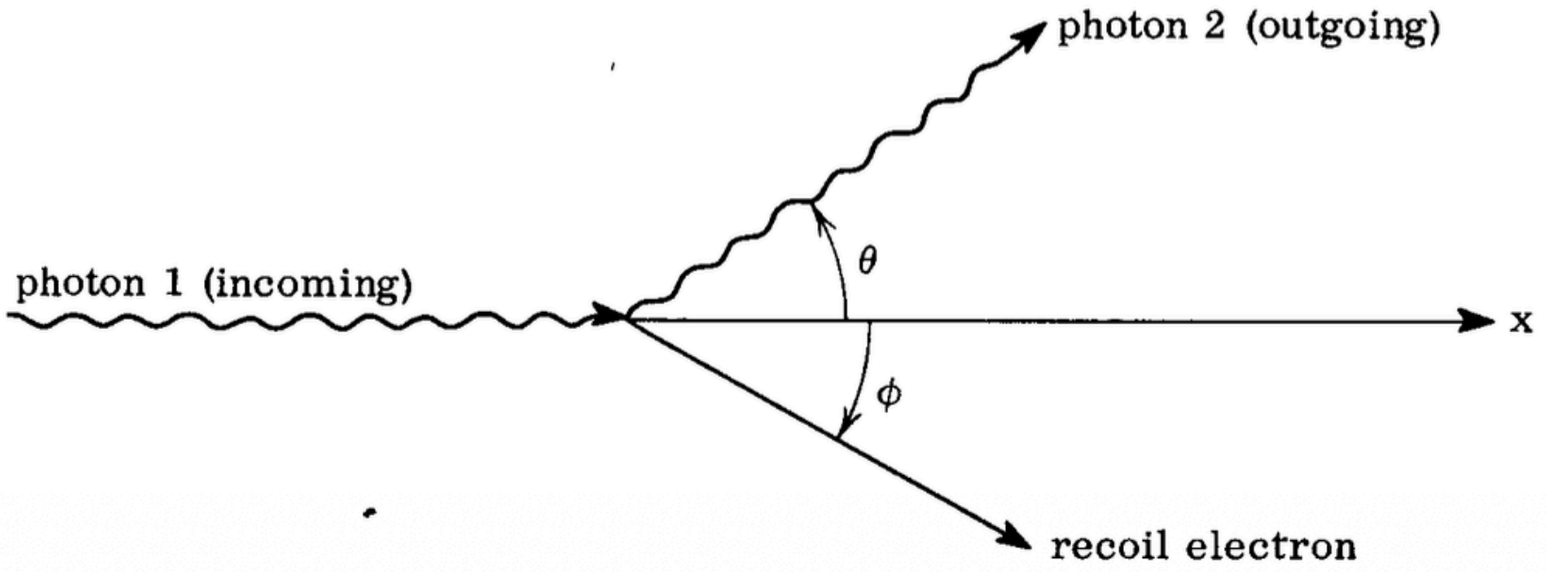


FIG. 19-1

Thus,

$$e_1 \cdot q_1 = 0 \qquad e_2 \cdot q_2 = 0$$

also

$$q_1 \cdot q_1 = {q_1}^2 = 0 \qquad \text{and} \qquad q_2 \cdot q_2 = {q_2}^2 = 0 \tag{19-3}$$

As initial and final state electron wave functions, choose

$$\Psi_1 = u_1 \exp(-ip_1 \cdot x)$$

$$\Psi_2 = u_2 \exp(-ip_2 \cdot x)$$

where u_1, u_2, p_1, and p_2 satisfy

$$\not{p}_1 u_1 = m u_1 \qquad \not{p}_2 u_2 = m u_2$$

$$p_1 \cdot p_1 = m^2 \qquad p_2 \cdot p_2 = m^2 \tag{19-4}$$

Conservation of energy and momentum (four equations) is written

$$\not{p}_1 + \not{q}_1 = \not{p}_2 + \not{q}_2 \tag{19-5}$$

If the coordinate system is chosen so that electron number 1 is at rest,

$$\not{p}_1 = m\gamma_t \tag{19-6a}$$

$$\not{p}_2 = E_2\gamma_t - p_2 \cos\phi\gamma_x + p_2 \sin\phi\gamma_y \tag{19-6b}$$

$$\not{q}_1 = \omega_1(\gamma_t - \gamma_x) \tag{19-6c}$$

$$\not{q}_2 = \omega_2(\gamma_t - \gamma_x \cos\theta - \gamma_y \sin\theta) \tag{19-6d}$$

The latter two equations follow from the fact that, for a photon, the energy and momentum are both equal to the frequency (in units in which c = 1).

The momentum has been resolved into components.

The incoming photon beam can be resolved into two types of polarization, which will be designated type A and type B:

$$\text{(A)}\ \not{e}_1 = \gamma_z \qquad \text{(B)}\ \not{e}_1 = \gamma_y$$

Type A has the electric vector in the z direction and type B has the electric vector in the y direction.

Similarly the outgoing photon beam can be resolved into two types of polarization:

$$(A')\ \not{e}_2 = \gamma_z \qquad (B')\ \not{e}_2 = \gamma_y \cos\theta - \gamma_x \sin\theta$$

Conservation of energy of momentum dictates that either the angle of the recoil electron ϕ or the angle at which the scattered photon comes off θ completely determines the remaining quantities.

If the electron direction is unimportant, its momentum can be eliminated by solving Eq. (19-5) for $\not{p}_2$ and squaring the resulting equation:

$$\not{p}_2 = \not{p}_1 + \not{q}_1 - \not{q}_2$$

$$p_2^2 = m^2 = (\not{p}_1 + \not{q}_1 - \not{q}_2)(\not{p}_1 + \not{q}_1 - \not{q}_2)$$

$$= p_1^2 + q_1^2 + q_2^2 + 2p_1 \cdot q_1 - 2p_1 \cdot q_2 - 2q_1 \cdot q_2$$

$$= m^2 + 0 + 0 + 2m\omega_1 - 2m\omega_2 - 2\omega_1\omega_2 (1 - \cos\theta)$$

where the last line was obtained from the preceding line by using Eqs. (19-3), (19-4), and (19-6a, c, d).

This can be written

$$m(\omega_1 - \omega_2) = \omega_1\omega_2(1 - \cos\theta)$$

or

$$(m/\omega_2) - (m/\omega_1) = 1 - \cos\theta \qquad (19\text{-}7)$$

This is the well-known formula for the Compton shift in wavelength (or frequency).

DIGRESSION ON THE DENSITY OF FINAL STATES

By the method discussed in the earlier part of the notes, the following final state densities (per unit energy interval) can be obtained.

When a system of total energy E and total linear momentum **p** disintegrates into a *two-particle final state*,

$$\text{Density of states} = (2\pi)^{-3} E_1 E_2 \frac{p_1^{\,3}\, d\Omega_1}{E p_1^{\,2} - E_1(\mathbf{p}\cdot\mathbf{p}_1)} \qquad \text{(D-1)}$$

where E_1 = energy of particle 1; E_2 = energy of particle 2; $\mathbf{p}_1$ = momentum of particle 1; $d\Omega_1$ = solid angle, into which particle 1 comes out; m_1 = mass of particle 1; m_2 = mass of particle 2; and $E_1 + E_2 = E$, $\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}$.

Another useful formula is in terms of the final energy of particle 1 and its azimuth ϕ_1 (instead of θ_1, ϕ_1).

It is

$$\text{Density of states} = (2\pi)^{-3} (E_1 E_2 / |\mathbf{p}|)\, dE_1\, d\phi_1 \qquad \text{(D-2)}$$

Special cases: (a) When $m_2 = \infty$ ($E_2 = \infty$, $E = \infty$):

$$\text{Density of states} = (2\pi)^{-3}\, E_1 |p_1|\, d\Omega_1 \qquad \text{(D-3)}$$

(b) In center-of-mass system $\mathbf{p} = 0$:

$$\text{Density of states} = (2\pi)^{-3} [E_1 E_2\, d\Omega_1 / (E_1 + E_2)] \qquad \text{(D-4)}$$

When a system disintegrates into a three-particle final state,

$$\text{Density of states} = (2\pi)^{-6}\, E_3 E_2 \frac{p_2^{\,3}\, p_1^{\,2}\, dp_1\, d\Omega_1\, d\Omega_2}{p_2^{\,2}(E - E_1) - E_2 \mathbf{p}_2 \cdot (\mathbf{p} - \mathbf{p}_1)} \tag{D-5}$$

Special case: When $m_3 = \infty$:

$$\text{Density of states} = (2\pi)^{-6}\, E_2\, |\mathbf{p}_2|\, d\Omega_2\, p_1^{\,2}\, dp_1\, d\Omega_1 \tag{D-6}$$

The Compton effect has a two-particle final state: taking particle 1 to be photon 2 and particle 2 to be electron 2, from Eq. (D-1),

$$\text{Density of states} = (2\pi)^{-3}\, \omega_2 E_2 \frac{\omega_2^{\,3}\, d\Omega_\omega}{(m + \omega_1)\omega_2^{\,2} - \omega_2(\omega_1 \omega_2 \cos\theta)}$$

COMPTON RADIATION

Calculation of $|M|^2$.

Using the Compton relation Eq. (19-7) to eliminate θ, this becomes

$$\text{Density of states} = (2\pi)^{-3}\, (E_2 \omega_2^{\,3}\, d\Omega_\omega / m\omega_1)$$

The probability of transition per second is given by

$$\text{Trans. prob./sec} = \sigma c = (2\pi / 2E_1 2E_2 2\omega_1 2\omega_2)\, |M|^2 \times (2\pi)^{-3} (E_2 \omega_2^{\,3}\, d\Omega_\omega / m\omega_1)$$

or

$$\sigma = [\omega_2^{\,2}\, d\Omega_\omega / (2\pi)^2 16 m^2 \omega_1^{\,2}]\, |M|^2$$

In working out the matrix element M, there are two ways in which the scattering can happen: (R) the incoming photon is absorbed by the electron and then the electron emits the outgoing photon; (S) the electron emits a photon and subsequently absorbs the incident photon.

These two processes are shown diagrammatically in Fig. 19-2.

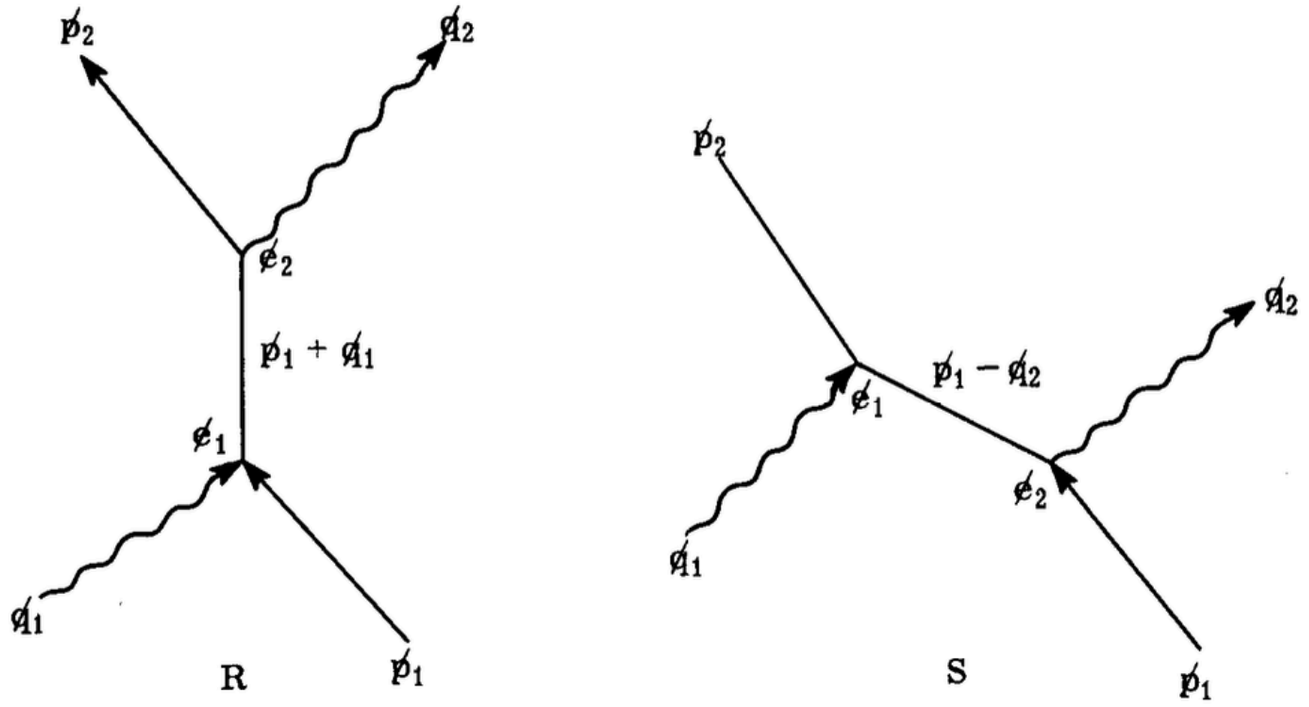


FIG. 19-2

In the momentum representation, the matrix element M for the first process R is

$$i[-i(4\pi e^2)^{1/2}]^2 \{\tilde{u}_2 \not{e}_2 [1/(\not{p}_1 + \not{q}_1 - m)] \not{e}_1 u_1\}$$

Reading from right to left the factors in the matrix element are interpreted as follows: (a) The initial electron enters with amplitude u_1; (b) the electron is first scattered by a potential (i.e., absorbs a photon); (c) having received momentum $\not{q}_1$ from the potential the electron travels as a free electron with momentum $\not{p}_1 + \not{q}_1$; (d) the electron emits a photon of polarization $\not{e}_2$; ; and (e) we now ask for the amplitude, that the electron is in a state u_2.

You should now be able to write down the matrix element for the second process S.

$$-i4\pi e^2\{\tilde{u}_2 \not{\epsilon}_1[1/(\not{p}_1 - \not{k}_2 - m)]\not{\epsilon}_2 u_1\} = -i4\pi e^2(\tilde{u}_2 S u_1)$$

The total matrix element is the sum of these two.

We can then rationalize these matrix elements and, using the table of matrix elements (Table 13-1) work out $|M|^2$.

The complete matrix element is the sum of these, so that the cross section becomes

$$\sigma = (e^4/4m)(\omega_2^2/\omega_1^2)\, d\Omega_2\, |\tilde{u}_2(R+S)u_1|^2$$

The problem now is actually to compute the matrix elements for R and S.

First R will be considered.

Using the identity

$$1/(\not{p} - m) = (\not{p} + m)/(p^2 - m^2)$$

the matrices may be removed from the denominator of R giving

$$R = \frac{\not{\epsilon}_2(\not{p}_1 + \not{k}_1 + m)\not{\epsilon}_1}{(\not{p}_1 + \not{k}_1) - m} = \frac{\not{\epsilon}_2(\not{p}_1 + \not{k}_1 + m)\not{\epsilon}_1}{2m\omega_1}$$

The denominator is seen to be $2m\omega_1$ from the following relations:

$$(\not{p}_1 + \not{q}_1)^2 - m^2 = p_1^2 + 2p_1 \cdot q_1 + q_1^2 - m^2$$

$$p_1^2 = m^2$$

$$q_1^2 = 0$$

$$2p_1 \cdot q_1 = 2m\omega_1$$

The matrix elements for the various spin and polarization combinations can be calculated straightforwardly from this point.

But certain preliminary manipulations will reduce the labor involved.

Using the identity

$$\not{a}\not{b} = 2a \cdot b - \not{b}\not{a}$$

it is seen that

$$\not{e}_2\not{p}_1\not{e}_1 = \not{e}_2(2p_1 \cdot e_1) - \not{e}_2\not{e}_1\not{p}_1$$

But p_1 has only a time component and e_1 only a space component so $p_1 \cdot e_1 = 0$.

Recalling that $\not{p}_1 u_1 = m u_1$, it is seen that

$$\tilde{u}_2\not{e}_2\not{p}_1\not{e}_1 u_1 = -\tilde{u}_2\not{e}_2\not{e}_1\not{p}_1 u_1 = -(\tilde{u}_2\not{e}_2\not{e}_1 u_1)m$$

and this is the matrix element of the first term of R.

the

It is also the negative of the matrix element of the last term of R, so R may be replaced by the equivalent

$$R = \not{e}_2 \not{A}_1 \not{e}_1 / 2m\omega_1$$

By an exactly similar manipulation, the S matrix is equivalent to

$$S = \not{e}_1 \not{A}_2 \not{e}_2 / 2m\omega_2$$

Substituting $\not{A}_1 = \omega_1(\gamma_t - \gamma_x)$ and $\not{A}_2 = \omega_2(\gamma_t - \gamma_x \cos\theta - \gamma_y \sin\theta)$ and transposing the 2m factor, the complete matrix may be written

$$2m(R+S) = \not{e}_2(\gamma_t - \gamma_x)\not{e}_1 + \not{e}_1(\gamma_t - \gamma_x \cos\theta - \gamma_y \sin\theta)\not{e}_2$$

A still more useful form is obtained by noting that $\not{e}_1$ anticommutes with q_1 $(e_1 \cdot q_1 = 0)$ and $\not{e}_2$ with q_2 and that $\not{e}_2\not{e}_1 = 2e_2 \cdot e_1 - \not{e}_1\not{e}_2$.

Thus,

$$2m(R+S) = -\not{e}_2\not{e}_1(\gamma_t - \gamma_x) - \not{e}_1\not{e}_2(\gamma_t - \gamma_x + \gamma_x - \gamma_x \cos\theta - \gamma_y \sin\theta)$$

$$= -2(e_2 \cdot e_1)(\gamma_t - \gamma_x) - \not{e}_1\not{e}_2[\gamma_x(1 - \cos\theta) - \gamma_y \sin\theta]$$

Using this form of the matrix, the matrix elements may be computed easily.

For example, consider the case for polarization: $\not{e}_1 = \gamma_z$, $\not{e}_2 = \gamma_y \cos\theta - \gamma_x \sin\theta$.

This corresponds to cases (A) and (B') of Part 19 and will be denoted by (AB').

The matrix is

$$2m(R+S) = -\gamma_z(\gamma_y \cos\theta - \gamma_x \sin\theta)[\gamma_x(1-\cos\theta) - \gamma_y \sin\theta]$$

since $e_2 \cdot e_1 = 0$.

Expanded this becomes

$$2m(R+S) = -\gamma_z[\gamma_y\gamma_x \cos\theta(1 \cos\theta) + \cos\theta \sin\theta + \sin\theta(1-\cos\theta)$$

$$+ \gamma_x\gamma_y \sin^2\theta]$$

$$= -\gamma_z(\gamma_x\gamma_y - \gamma_x\gamma_y\cos\theta + \sin\theta) = -\gamma_x\gamma_y\gamma_z(1-\cos\theta)$$

$$- \gamma_z \sin\theta$$

where the anticommutation of the γ's has been used.

In the case of spin-up for the incoming particle and spin down for the outgoing particle ($s_1 = -1$, $s_2 = -1$) the matrix elements

$$-2m(F_1F_2)^{1/2}(\tilde{u}_2\gamma_x\gamma_y\gamma_z u_1) = -iF_2p_{1+} - iF_1p_{2+}$$

$$-2m(F_1F_2)^{1/2}(\tilde{u}_2\gamma_z u_1) = +p_{1+}F_2 - p_{2+}F_1$$

may be found by reference to Table 13-1.

But note that in this problem $p_{1+} = p_{x1} + ip_{y1} = 0$ since particle 1 is at rest.

Hence the final matrix element for this case, polarization (AB'), spin $s_1 = +1$, $s_2 = -1$, is

$$2m\,(F_1F_2)^{1/2}\,(\tilde{u}_2(R+S)u_1) = -(1-\cos\theta)\,i\,F_1p_{2+} - \sin\theta\;\; p_{2+}\,F_1$$

The results for the other combinations of polarization and spin are obtained in the same manner and will only be presented in tabular form (Table 20-1 on next slide).

For any one of the polarization cases listed, $|M|^2$ is the sum of the square amplitudes of the matrix elements for outgoing spin states averaged over incoming spin states.

But this is seen to be simply the square magnitude of the nonzero matrix element listed under the appropriate polarization case.

For example, in case (AA'),

$$|M|^2 = |\tilde{u}_2(R+S)u_1|^2 = (1/4m^2F_1F_2)\,|2\,F_2F_1 - (1+\cos\theta)\,F_1p_{2-}$$

$$- i\,\sin\theta\;\; F_1p_{2+}$$

By employing the relation

$$p_{2-} = p_{1-} + q_{1-} - q_{2-} = q_{1-} - q_{2-} = \omega_1 - \omega_2\cos\theta + i\omega_2\sin\theta$$

and

$$(m/\omega_2) - (m/\omega_1) = 1 - \cos\theta$$

TABLE 20-1

Polarization		AA′	AB′	BA′	BB′
ϵ_1		γ_z	γ_z	γ_y	γ_y
ϵ_2		γ_z	$\gamma_y \cos\theta - \gamma_x \sin\theta$	γ_z	$\gamma_y \cos\theta - \gamma_x \sin\theta$
Matrix $2m(R+S)$		$2\gamma_t - \gamma_x(1+\cos\theta)$ $-\gamma_y \sin\theta$	$-\gamma_x\gamma_y\gamma_z(1-\cos\theta)$ $-\gamma_z \sin\theta$	$-\gamma_x\gamma_y\gamma_z(1-\cos\theta)$ $+\gamma_z \sin\theta$	$2\cos\theta\,\gamma_t - \gamma_x(1+\cos\theta)$ $-\gamma_y \sin\theta$
Matrix elements $\frac{\tilde{}}{2m\sqrt{F_1F_2}}(\tilde{u}_2(R+S)u_1)$	$s_1 = +1$	$+2F_2F_1 - (1+\cos\theta)F_1p_{2-}$	0	0	$2\cos F_2F_1 - (1+\cos\theta)F_1p_{2-}$
	$s_2 = +1$	$-i\sin\theta\ F_1p_{2-}$	0	0	$-i\sin\theta\ F_1p_{2-}$
	$s_1 = +1$	0	$-i(1-\cos\theta)\ F_1p_{2+}$	$-i(1-\cos\theta)F_1p_{2+}$	0
	$s_2 = -1$	0	$-\sin\theta\ F_1p_{2+}$	$+\sin\theta\ F_1p_{2+}$	0

Note: The matrix elements for $\begin{pmatrix} s_1 = -1 \\ s_2 = -1 \end{pmatrix}$ are the complex conjugates of those above for $\begin{pmatrix} s_1 = +1 \\ s_2 = +1 \end{pmatrix}$, and for $\begin{pmatrix} s_1 = -1 \\ s_2 = +1 \end{pmatrix}$ they are the complex conjugates of those for $\begin{pmatrix} s_1 = +1 \\ s_2 = -1 \end{pmatrix}$ above.

the square magnitudes of the matrix elements for the various cases reduce, after considerable amount of algebra, to the expressions given in Table 20-2.

TABLE 20-2

Polarization	$\|M\|^2$
AA′	$[(\omega_1 - \omega_2)^2/\omega_1\omega_2] + 4$
AB′	$[(\omega_1 - \omega_2)^2/\omega_1\omega_2]$
BA′	$[(\omega_1 - \omega_2)^2/\omega_1\omega_2]$
BB′	$[(\omega_1 - \omega_2)^2/\omega_1\omega_2] + 4\cos^2\theta$

It is clear that all four of these formulas may be written simultaneously in the form

$$|M|^2 = [(\omega_1 - \omega_2)^2/\omega_1\omega_2] + 4(e_1 \cdot e_2)^2$$

Note that these formulas are not adequate for circular polarization.

That is $\not{e}_1$ were, for example, $1/\sqrt{2}\ (i\gamma_z + \gamma_y)$, it is seen that because of the phasing represented by the imaginary part of $\not{e}_1$, all the calculations must be carried out before squaring the matrix elements in order to get the proper interference.

Finally the cross section for scattering with prescribed plane polarization of the incoming and outgoing photons is

$$\sigma = (e^4/4m^2)(\omega_2^2/\omega_1^2)\ d\Omega\,\omega_2\ [(\omega_2/\omega_1) + (\omega_1/\omega_2) - 2 + 4(e_1 \cdot e_2)^2]$$

This is the Klein-Nishina formula for polarized light.

For unpolarized light this cross section must be averaged over all polarizations.

It is noted that diagram cases such as Fig. 20-1 have been included in the previous derivation as a result of the generality in the transformation of $K_+(2,1)$ to the momentum representation.

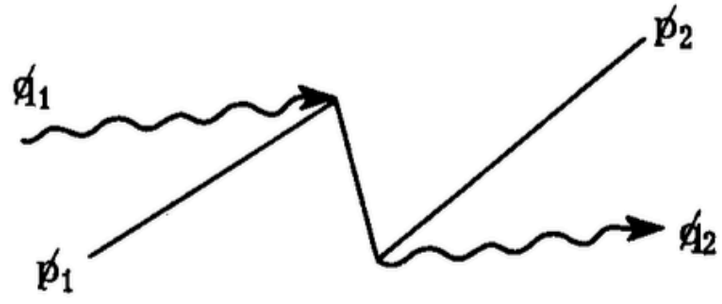


FIG. 20-1

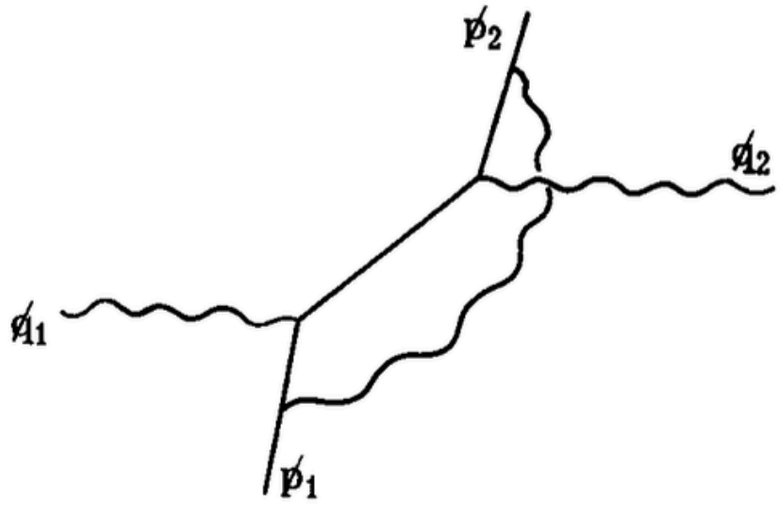


FIG. 20-2

In fact, all diagram cases have been included except higher-order effects to be discussed later.

They correspond to emission and reabsorption of a third photon by the electron, such as in Fig. 20-2.

Part 21 Discussion of the Klein-Nishina Formula.

In the 'Thompson limit," $\omega_1 << m$, the electron picks up very little energy in recoil, and $\omega_1 \approx \omega_2$.T

This can be seen from the relation

$$m\omega_1 - m\omega_2 = \omega_1 \omega_2 (1 - \cos\theta) \qquad (21\text{-}1)$$

In this limit, the Klein-Nishina formula gives

$$\sigma = (e^4/m^2)(\mathbf{e}_1 \cdot \mathbf{e}_2)^2 \, d\Omega_\omega \tag{21-2}$$

which is the Rayleigh-Thompson scattering cross section.

Note that ω is still very large compared to the eigenvalues of an atom, in accordance with our original assumptions for Compton scattering.

The same result is obtained by a classical picture.

Under the action of the electric field of the photon $\mathbf{E} = E_0\mathbf{e}_1\exp(i\omega t)$, the electron is given the acceleration

$$\mathbf{a} = (e/m)E_0\,\mathbf{e}_1 \exp(i\omega t)$$

Classically, an accelerated charge radiates to give the scattered radiation

$$\mathbf{E}_s = -\frac{e}{R} \quad \text{(retarded acceleration projected on plane } \perp \text{ to line of sight)}$$

The scattered radiation polarized in the direction $\mathbf{e}_2$ is determined by the component of the acceleration in this direction.

The intensity of the scattered radiation of polarization $\mathbf{e}_2$ is then (times R^2 per unit solid angle and per unit incident intensity

$$I = (e^4/m^2)(\mathbf{e}_1 \cdot \mathbf{e}_2)^2 \tag{21-2′}$$

The customary $\hbar$'s and c's may be replaced in Eq. (21-1) as follows (σ is an area or length squared):

$$e^4 = (e^2)^2 = (e^2/\hbar c)^2$$

$$m^2 = (mc/\hbar)^2 = \text{length squared}$$

$$e^4/m^2 = (e^2/mc^2)^2 = r_0^2 \approx 8 \times 10^{-26}\ \text{cm}^2$$

Averaging over Polarizations. It is often desired to have the scattering cross section for a beam regardless of the incoming or outgoing polarization.

This can be obtained by summing the probabilities over the polarizations of the outgoing beam and averaging over the incoming beam.

Thus, suppose the incoming beam has polarization of type A.

The probabilities (or cross sections) for the two possible types of outgoing polarization, A' and B' can be symbolized as AA' and AB'.

The total probability for scattering a photon of either polarization is AA' + AB'.

Then suppose the incoming beam is equally likely to be polarized as type A or type B.

The resulting probability can be obtained as the sum 1/2 (probability if type A) +1/2 (probability if type B).

This is the situation for unpolarized incoming beam, and gives

$$\sigma(\text{averaged over polarizations}) = (1/2)(AA' + AB') + (1/2)(BA' + BB')$$
$$= \frac{e^4}{2m^2}\left(\frac{\omega_2}{\omega_1}\right)^2 d\Omega_{\omega_2}\left(\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} - \sin^2\theta\right) \tag{21-3}$$

If, on the other hand, the polarization of the outgoing beam is measured (still with an unpolarized incoming beam), its dependence on frequency and scattering angle is given by the ratio

$$\frac{\text{Probability of polarization type } A'}{\text{Probability of polarization type } B'} = \frac{(1/2)[AA' + BA']}{(1/2)[AB' + BB']}$$

$$= \frac{(\omega_2/\omega_1) + (\omega_1/\omega_2)}{(\omega_2/\omega_1) + (\omega_1/\omega_2) - 2\sin^2\theta}$$

The forward radiation ($\theta = 0$) remains unpolarized, but a certain degree of polarization will be found in light scattered through any nonzero angle.

In the low-frequency limit ($\omega_1 \approx \omega_2$), the polarization is complete at $\theta = \pi/2$.

Thus an unpolarized beam becomes plane-polarized when scattered through 90◦.

Total Scattering Cross Section. If the cross section (averaged over polarizations) given in Eq. (21-3) is integrated over the solid angle

$$d\Omega = 2\pi\, d(\cos\theta) = (2\pi m/\omega_2^2)\, d\omega_2$$

the total cross section for scattering through any angle is obtained.

So, from Eq. (21-1),

$$\cos\theta = 1 - m/\omega_2 + m/\omega_1 \qquad (21\text{-}1')$$

and the variable ω_2 goes between the limits $m\omega_1/(2\omega_1 + m)$ and ω_1 as $\cos\theta$ goes from - 1 to + 1.

Equation (21-3) can be written

$$d\sigma_T = (e^4/2m^2)(2\pi/\omega_1^2)m\, d\omega_2(\omega_2/\omega_1 + \omega_1/\omega_2 - 2m/\omega_2 + 2m/\omega_1$$

$$+ m^2/\omega_1^2 + m^2/\omega_2^2 - 2m^2/\omega_1\omega_2)$$

where the last five terms replace $-\sin2\theta = \cos2\theta - 1$ using Eq. (21-1')

Simple integrations yield

$$\sigma_T = \pi e^4/m^2[(m/\omega_1 - 2m^2/\omega_1^2 - 2m^3/\omega_1^3)\log(2\omega_1/m + 1)$$

$$+ m/2\omega_1 + 4m^2/\omega_1^2 - m^3/2\omega_1(2\omega_1 + m)^2]$$

In the high-frequency limit ($\omega_1 \to \infty$)

$$\sigma_T \sim (1/\omega_1)\log\omega_1 \to 0$$

Thus Compton scattering is a negligible effect at high frequencies, where pair production becomes the important effect.

TWO-PHOTON PAIR ANNIHILATION

From the quantum-electrodynamical point of view, another phenomenon completely analogous to Compton scattering is two-photon pair annihilation.

Two photons are necessary (in the outgoing radiation) to maintain conservation of momentum and energy when pair annihilation takes place in the absence of an external potential.

The interaction can be diagrammed as shown in Fig. 21-1.

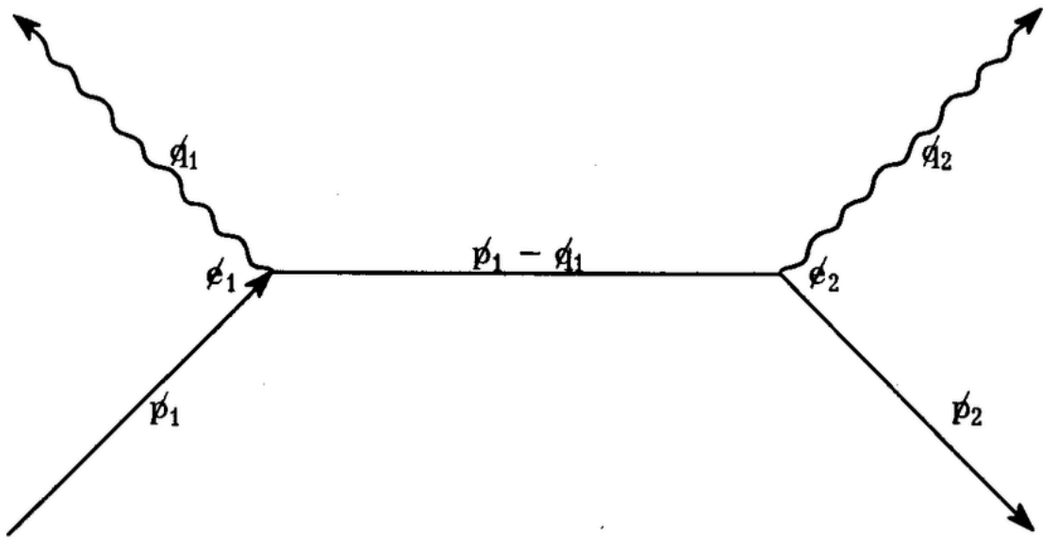


FIG. 21-1

This figure should be compared to that for Compton scattering (Part 20).

The only differences are that the direction of photon q_1 is reversed, and, since particle 2 is a positron, p_2 = - (momentum of positron).

So write

$$p_1 = (E_- \gamma_t - \mathbf{p}_- \cdot \boldsymbol{\gamma})$$

$$p_2 = -(E_+ \gamma_t - \mathbf{p}_+ \cdot \boldsymbol{\gamma})$$

where the energies E_- and E_+ of the electron and positron are both positive numbers.

The conservation law gives

$$\not{p}_2 = \not{p}_1 - \not{q}_1 - \not{q}_2 \tag{21-4}$$

(just as for Compton scattering, but the direction of $\not{q}_1$ reversed), so the matrix element for this interaction is

$$M_1 = -i\,4\pi e^2 (\tilde{u}_2 \not{\epsilon}_2 (\not{p}_1 - \not{q}_1 - m)^{-1} \not{\epsilon}_1 u_1)$$

The second possibility, indistinguishable from the first by any measurement, is obtained from the first by interchanging the two photons (see Fig. 21-2); again note similarity to Compton scattering.

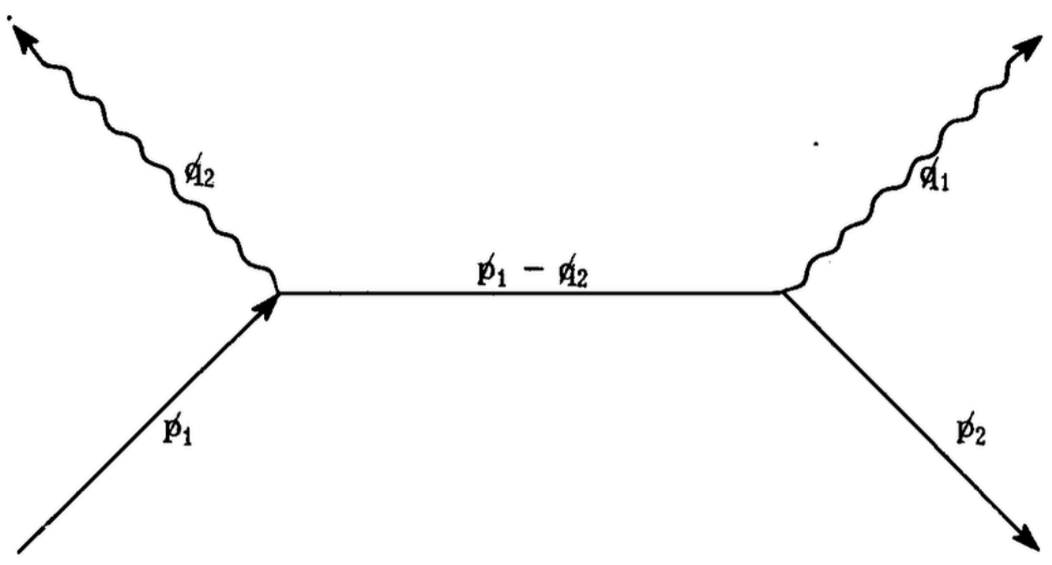


FIG. 21-2

Immediately, the matrix element is

$$M_2 = -i\,4\pi e^2 (\tilde{u}_2 \not{\epsilon}_1 (\not{p}_1 - \not{q}_2 - m)^{-1} \not{\epsilon}_2 u_1)$$

The sum of the two matrix elements and the density of final states gives the cross section

$$\sigma \cdot (\text{velocity of positron}) = 2\pi/(2E_- \cdot 2E_+ \cdot 2\omega_1 \cdot 2\omega_2) \cdot |M_1 + M_2|^2$$

$$\times (\text{density of states})$$

in a system where the electron is at rest and the positron is moving.

The density of final states is

$$\omega_1\omega_2/(2\pi)^3\ \omega_1{}^2\ d\Omega_1/(\omega_2\omega_1 - Q_2 \cdot Q_1)$$

Since particle 2 is a positron, $\not{p}_2 = -\not{p}_+$, so the conservation law, Eq. (21-4), gives

$$\not{p}_1 + \not{p}_+ = \not{q}_1 + \not{q}_2$$

Then

$$m^2 + 2(p_1 \cdot p_+) + m^2 = 0 + 2q_1 \cdot q_2 + 0$$

This reduces to

$$2m^2 + 2mE_+ = 2\omega_1\omega_2 - 2Q_1 \cdot Q_2$$

From a comparison of the diagrams, it is clear that the matrix elements for pair annihilation are the same as the matrix elements for the Compton effect if the sign of $\not{q}_1$ is changed.

In the cross section, this amounts changing the sign of ω_1.

Then the cross section is

$$\sigma = e^4 \omega_1^2 \, d\Omega_1/[4m^2(E_+ + m)|p_+|][(\omega_2/\omega_1) + (\omega_1/\omega_2) + 2 - 4(e \cdot e_2)^2]$$

in analogy with the Klein-Nishina formula.

Part 22 POSITRON ANNIHILATION FROM REST

The formula for positron-electron annihilation derived in Part 21 diverges as the positron velocity approaches zero ($\sigma \sim 1/v$; this is true for other cross sections when a process involves absorption of the incoming particle, and is the well-known 1/v law).

To calculate the positron lifetime in an electron density p (recall that the preceding cross section was for an electron density of one per cubic centimeter) as $v_+ \rightarrow 0$, we use

$$\text{Trans. prob./sec} = \sigma v_+ \rho$$

plus the fact that, as $v_+ \rightarrow 0$, $E_+ \rightarrow m$ and $\omega_1 \rightarrow \omega_2 \rightarrow m$ (when the electron and positron are both approximately at rest, momentum and energy can be conserved only with two photons of momenta equal in magnitude but opposite in direction).

Thus

$$\text{Trans. prob./sec} = \sigma v_+ \rho = (e^4/2m^2) \rho \, d\Omega \, (\sin^2 \theta) \qquad (22\text{-}1)$$

where θ = angle between directions of polarization of two photons ($\cos\theta = e_1 \bullet e_2$).

The $\sin^2\theta$ dependence indicates that the two photons have their polarizations at right angles.

To get the probability of transition per second for any photon direction and any polarization, it is necessary to sum over solid angle ($\int d\Omega = 4\pi$) and average over polarizations ($\sin^2\theta = 1/2$), giving

$$\text{Total trans. prob./sec} = 1/\tau = (\pi e^4/m^2)\,\rho = \pi(e^2/mc^2)^2\,c\rho = \pi r_0^{\,2}\,c\rho \qquad (22\text{-}2)$$

factors of c and $\hbar$ reinserted where required), where r_0 = classical electron radius, and τ = mean lifetime.

BREMSSTRAHLUNG

When an electron passes through the Coulomb field of a nucleus it is deflected.

Associated with this deflection is an acceleration which, according to the classical theory, results in radiation.

According to quantum electrodynamics, there is a certain probability that the incident electron will make a transition to a different electron state with a photon emitted, while in the field of the nucleus.

Interaction with the field of the nucleus is necessary to satisfy conservation of energy and momentum.

That is, the electron cannot emit a photon and make a transition to a different electron state while traveling along in a vacuum.

Figure 22-1 shows the process and defines angles that arise later.

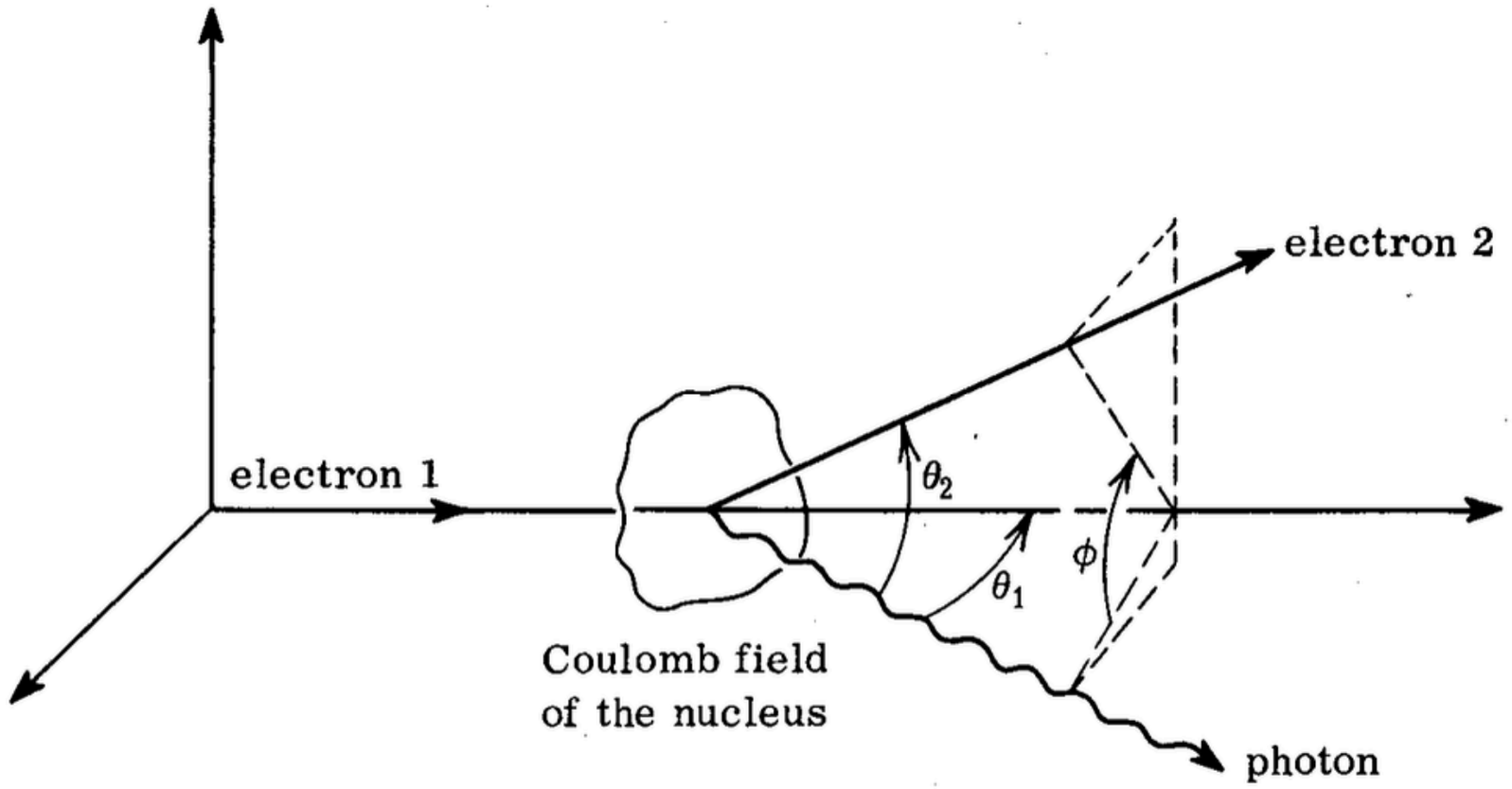


FIG. 22-1

The Coulomb potential of the nucleus will be considered to act only once (Born approximation).

The validity of this approximation was discussed in Part 16.

There are two (indistinguishable) orders in which the bremsstrahlung process can occur: (a) the electron interacts with the Coulomb field and subsequently emits a photon, or (b) the electron first emits a photon and then interacts with the Coulomb field.

The diagrams for these processes are shown in Fig. 22-2.

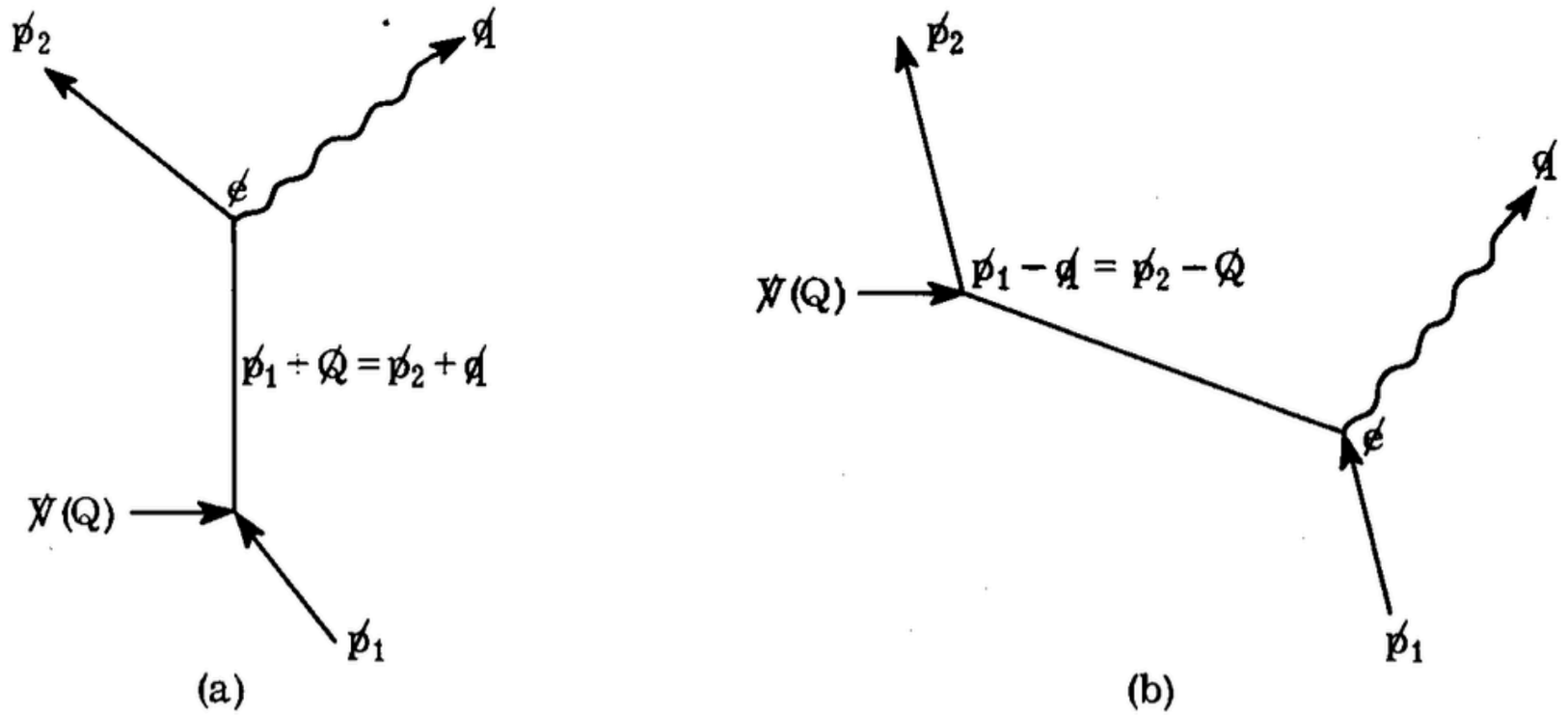


FIG. 22-2

The interaction with the nucleus gives momentum $\not Q$ to the electron.

Conservation of energy and momentum requires

$$\not p_1 + \not Q = \not p_2 + \not q \qquad \text{or} \qquad \not Q = \not p_2 - \not p_1 + \not q$$

In Part 18 it was shown that the Fourier transform of the Coulomb potential was proportional to $\delta(Q_4)$, since the potential is independent of time.

This means that only transitions for which $Q_4 = 0$ occur, or energy must be conserved among the incident electron, final electron, and photon.

Thus $E_1 = E_2 + \omega$.

The transition probability is given by

$$\text{Trans. prob./sec} = \sigma v_1 = (2\pi/2E_1\,2E_2\,2\omega)\,|\mathfrak{M}|^2 \times D$$

Since the nucleus is to be considered infinitely heavy,

$$D = (2\pi)^{-6}\,E_2 p_2\,d\Omega_2 \omega^2\,d\omega\,d\Omega_\omega$$

Notice that there is a spectrum of photons; that is, the photon energy is not determined (as it was in the Compton effect, for example.

Letting $\mathfrak{M} = (\tilde{u}_2 M u_1)$,

$$M = (-i)(4\pi e^2)^{1/2}\left[\not{\epsilon}\,\frac{1}{p_1 + Q - m}\,\not{A}(Q) + \not{A}(Q)\,\frac{1}{p_2 - Q - m}\,\not{\epsilon}\right] \qquad (22\text{-}3)$$

where the first term comes from Fig. 22-2a and the second term from Fig. 22-2b.

The explanation of the factors in the first term, for example, is, reading from right to left, that an electron initially in state uy is scattered by the Coulomb potential acquiring an additional momentum $\not{Q}$, the electron moves as a free particle with momentum by $\not{p}_1 + \not{Q}$ until it emits a photon of polarizatıon $\not{\epsilon}$.

We then ask: Is the electron in state u_2?

For the Coulomb potential

$$\not{V}(Q) = (4\pi Ze^2/Q^2)\,\delta(Q_4)\,\gamma_t = v(Q)\,\delta(Q_4)\,\gamma_t$$

(see Momentum Representation, Part 18) in a coordinate system in which the nucleus does not move.

[For potential other than Coulomb, use appropriate v(Q), the Fourier transform of the space dependence of the potential.]

Rationalizing the denominator of the matrix,

$$\begin{aligned}(\not p_1 + \not Q - m)(\not p_1 + \not Q + m) &= p_1^2 + 2p_1\cdot Q + Q^2 - m^2 = 2p_1\cdot Q + Q^2\\ &= -2\mathbf{p}_1\cdot\mathbf{Q} + Q^2\\ &= 2\mathbf{p}_1\cdot\mathbf{Q} - \mathbf{Q}^2 \qquad Q_4 = 0\end{aligned}$$

$$M = (-i)(4\pi e^2)^{1/2}\, v(Q)\left[\not\epsilon\,\frac{\not p_1 + \not Q + m}{-2\mathbf{p}_1\cdot\mathbf{Q} - \mathbf{Q}^2}\,\gamma_t + \gamma_t\,\frac{\not p_2 - \not Q + m}{2\mathbf{p}_2\cdot\mathbf{Q} - \mathbf{Q}^2}\,\not\epsilon\right] \tag{22-4}$$

The outgoing photon can be polarized in either of two directions, and the incoming and outgoing electron each have two possible spin states.

The various matrix elements can be worked out using Table 13-1 exactly as was done in deriving the Klein-Nishina cross section in Part 20 - nothing new is involved, so we omit the details.

After (1) summing over photon polarizations, (2) summing over outgoing electron spin states, and (3) averaging over incoming electron spin states, the following differential cross section is obtained:

$$d\sigma = \frac{1}{2\pi}\left(\frac{Ze^2}{Q^2}\right)^2 e^2 \frac{d\omega}{\omega}\frac{p_2}{p_1}\sin\theta_2\, d\theta_2 \sin\theta_1\, d\theta_1\, d\phi$$

$$\times\left\{\frac{p_2^2\sin^2\theta_2\,(4E_1^2 - Q^2)}{(E_2 - p_2\cos\theta_2)^2} + \frac{p_1^2\sin^2\theta_1\,(4E_2^2 - Q^2)}{(E_1 - p_1\cos\theta_1)^2}\right.$$

$$\left. - \frac{2p_1p_2\sin\theta_1\sin\theta_2\cos\phi\,(4E_1E_2 - Q^2 + 2\omega^2) - 2\omega^2(p_2^2\sin^2\theta_2 + p_1^2\sin^2\theta_1)}{(E_2 - p_2\cos\theta_2)(E_1 - p_1\cos\theta_1)}\right\} \qquad (22\text{-}5)$$

An approximate expression with a simple interpretation in terms of the Coulomb elastic scattering cross section can be obtained when the photon energy is small (small compared to rest mass of electron but large compared to electron binding energies).

Writing the matrix (22-3) in terms of $\not{q}$ instead of $\not{Q}$,

$$M = (-i)(4\pi e^2)^{1/2}\left[\not{\epsilon}\,\frac{1}{\not{p}_2 + \not{q} - m}\,\not{A}(Q) + \not{A}(Q)\,\frac{1}{\not{p}_1 - \not{q} - m}\,\not{\epsilon}\right]$$

$$= (-i)(4\pi e^2)^{1/2}\left[\not{\epsilon}\,\frac{\not{p}_2 - \not{q} + m}{+2p_2\cdot q}\,\not{A}(Q) + \not{A}(Q)\,\frac{\not{p}_1 - \not{q} + m}{-2p_1\cdot q}\,\not{\epsilon}\right]$$

using the relationships $\not{\epsilon}\not{p}_2 = -\not{p}_2\not{\epsilon} + 2e\cdot p_2$, $\not{p}_1\not{\epsilon} = -\not{\epsilon}\not{p}_1 + 2e\cdot p_1$, and neglecting $\not{q}$ in the numerator, since it is small, this becomes

$$M \approx (-i)(4\pi e^2)^{1/2}\, v(Q)\left[\frac{-\not p_2 \not e \gamma_t + 2e\cdot p_2 \gamma_t + m\not e\gamma_t}{2p_2\cdot q} + \frac{-\gamma_t \not e \not p_1 + 2p_1\cdot e\gamma_t + m\not e\gamma_t}{-2p_1\cdot q}\right]\delta(Q_4)$$

$$= (-i)(4\pi e^2)^{1/2}\, v(Q)\left[\frac{e\cdot p_1}{q\cdot p_1} - \frac{e\cdot p_2}{q\cdot p_2}\right]\gamma_t\,\delta(Q_4)$$

where use is made of the fact that the matrix element of M between states u_2 and u_1 is to be calculated and $\tilde{u}_2 \not p_2 = \tilde{u}_2 m$, $\not p_1 u_1 = m u_1$.

The cross section for photon emission can then be written

$$d\sigma = \frac{1}{v}\left[\frac{2\pi}{2E_1\,2E_2}|v(Q)|^2\,\frac{E_2 p_2\, d\Omega_2}{(2\pi)^3}\right]\left[\frac{e^2\, d\omega\cdot d\Omega_\omega}{\pi\omega}\left(\frac{p_2\cdot e}{p_2\cdot\frac{q}{\omega}} - \frac{p_1\cdot e}{p_1\cdot\frac{q}{\omega}}\right)^2\right]$$

The first bracket is the probability of transition for elastic scattering (see Part 16), so the last bracket may be interpreted as the probability of photon emission in frequency interval $d\omega$ and solid angle $d\Omega_\omega$ if there is elastic scattering from momentum p_1 to p_2.

PAIR PRODUCTION

It is easily shown that a single photon of energy greater than 2m cannot create an electron positron pair without the presence of some other means of conserving momentum and energy.

Two photons could get together and create a pair, but the photon density is so low that this process is extremely unlikely.

A photon can, however, create a pair with the aid of a field, such as that of a nucleus, to which it can impart some momentum.

As with bremsstrahlung, there are two indistinguishable ways in which this can happen: (a) The incoming photon creates a pair and subsequently the electron interacts with the field of the nucleus; or (b) the photon creates a pair and the positron interacts with the field of the nucleus.

The diagrams for these alternatives are shown in Fig. 22-3.

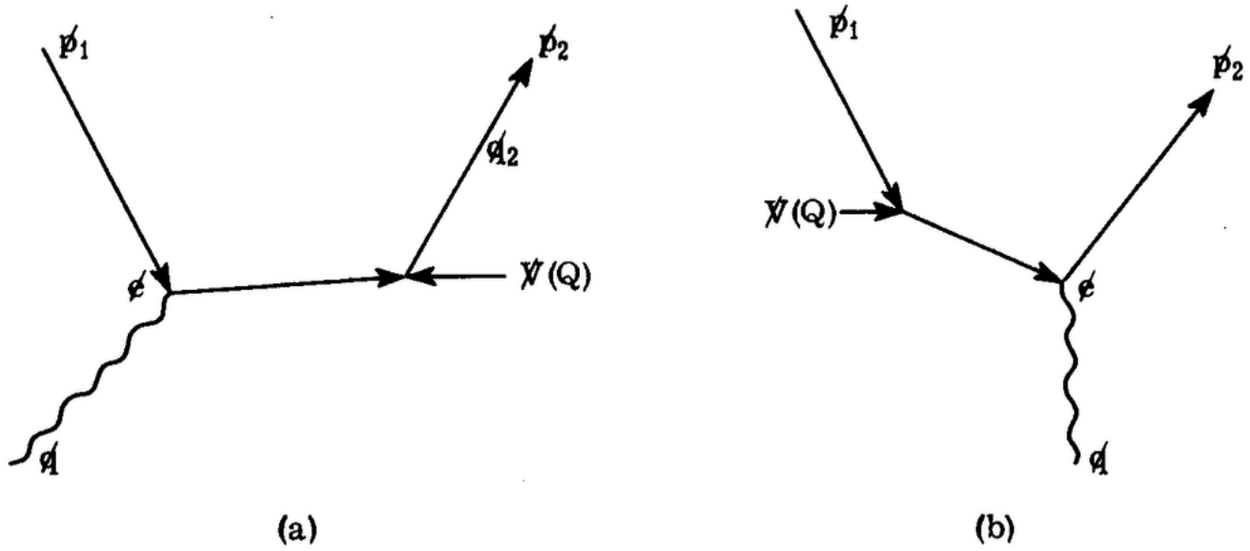


FIG. 22-3

The arrows in the diagram indicate that $\not{p}_1$ is the positron momentum and $\not{p}_2$ is the electron momentum.

Notice that, with respect to the directions that the arrows point and without regard to direction of increasing time), these diagrams look exactly like those for the bremsstrahlung process: Starting with $\not{p}_1$, in case (a), the particle is first scattered by the Coulomb potential and then by the photon; in case (b) the order of the events is reversed.

The difference between pair production and bremsstrahlung, when the direction of time is taken into account, is (1) $\not{p}_1$ is a positron state (an electron traveling backward in time), and (2) the photon $\not{q}$ is absorbed rather than emitted.

As a result, the bremsstrahlung matrix elements can be used for this process if $\not{p}_1$ is replaced by $-\not{p}_+$ and $\not{q}$ by $-\not{q}$.

The $\not{p}_+$ is then the positron momentum and $\not{q}$ is the momentum of the absorbed photon.

The density of final states is different, of course, since the particles in the final state are now a positron and electron.

Thus

$$d\sigma = (1/2\pi)(Ze^2/Q^2)^2\, e^2\, (p_+ p_-\, \sin\theta_+\, d\theta_+\, \sin\theta_-\, d\theta_-\, d\phi/\omega^3) \times \{\ \} \tag{22-6}$$

where the braces are the same as for bremsstrahlung, Eq. (22-5), except for the following substitutions:

p_- for p_2	$-\theta_-$ for θ_2	E_- for $\overline{E}_+$
$-p_+$ for p_1	$-\theta_+$ for θ_-	$-E_+$ for E_1
		$-\omega$ for ω

Figure 22-4 defines the angles (ϕ = angle between electron-photon plane and the positron-photon plane).

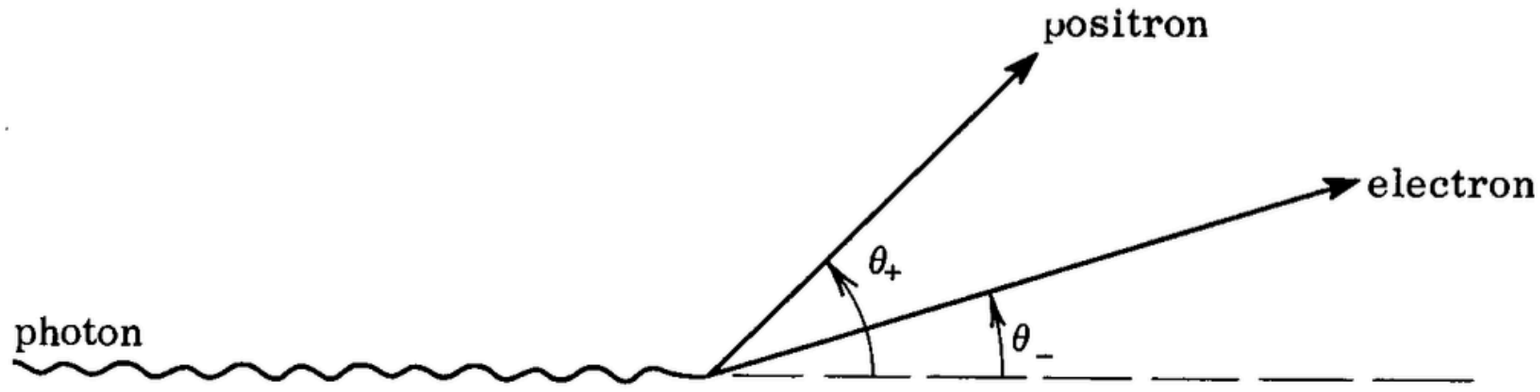


FIG. 22-4

Part 23 A METHOD OF SUMMING MATRIX ELEMENTS OVER SPIN STATES

By using current methods of computing cross sections, one first arrives at a cross section for "polarized" electrons, that is, electrons with definite incoming and outgoing spin states.

In practice it is common that the incoming beam will be "unpolarized" and the spins of the outgoing particles will be unobserved.

In this case, one needs the cross section obtained from that for "polarized" electrons by summing probabilities over final spin states and averaging this sum over initial spin states.

This is the correct process since the final spin states do not interfere and there is equal probability of initial spin in either direction.

Formally, if

$$\sigma \sim |(\tilde{u}_2 M u_1)|^2$$

one needs

$$\sigma \sim \frac{1}{2} \sum_{\text{spins } 1} \sum_{\text{spins } 2} |(\tilde{u}_2 M u_1)|^2 \qquad (23\text{-}1)$$

where $\sum_{\text{spins } 2}$ means the sum over final spin states for only one sign of the energy, that is over only two of the four possible eigenstates.

Similarly, $\sum_{\text{spins } 1}$ is the sum over initial spins for one sign of the energy.

The purpose now is to develop a simple method for obtaining these sums.

In accordance with the usual rule for matrix multiplication, the following is true:

$$\sum_{\text{all } u_1} (\tilde{u}_2 A u_1)(\tilde{u}_1 B u_2) = 2m(\tilde{u}_2 A B u_2) \qquad (23\text{-}2)$$

where A and B are any operators or matrices, the 2m factor on the right arises from the normalization $\tilde{u}u = 2m$, and the sum is over *all* eigenstates represented by u_y.

But the states u, which we want in Eq. (23-1) are not all states, just those satisfying $\not{p}_1 u_1 = m u_1$.

That is, they belong to the eigenvalue +m of the operator $\not{p}_1$.

Since $\not{p}_1^2 = m^2$, $\not{p}_1$ also has the eigenvalue -m, that is, there are two more solutions of which, together with the two we wish in Eq. (23-1) bring the total to four.

Let us call the latter "negative eigenvalue" states.

Now, if in Eg. (23-2) the matrix elements of B were zero in negative eigenvalue states, this would be the same as $\sum_{\text{spins } 1}$ that is, just over positive eigenvalue states.

So consider

$$\sum_{\text{all } u_1} (\tilde{u}_2 A u_1)(\tilde{u}_1(\not{p}_1 + m)B u_2) = (\tilde{u}_2 A(\not{p}_1 + m)B u_2)2m$$

But

$$\tilde{u}_1(\not{p}_1 + m) = 0 \qquad \text{for negative eigenvalue states}$$

$$= \tilde{u}_1(2m) \qquad \text{for positive eigenvalue states}$$

so the preceding sum also equals

$$\sum_{\text{spins } 1} (\tilde{u}_2 A u_1)2m(\tilde{u}_1 B u_2)$$

Cancelling the 2m factors, this gives

$$\sum_{\text{spins } 1} (\tilde{u}_2 A u_1)(\tilde{u}_1 B u_2) = (\tilde{u}_2 A(\not{p}_1 + m)B u_2)$$

$(\not{p}_1 + m)$ is called a projection operator for obvious reasons.

Similarly it follows that

$$\sum_{\text{spins } 2} (\tilde{u}_2 X u_2) = \sum_{\text{all } u_2} (1/2m)(\tilde{u}_2(\not{p}_2 + m)X u_2)$$

where X is again any matrix.

Remembering the normalization $\tilde{u}_2 u_2 = 2m$, it is seen that the last sum is just the trace or spur of the matrix $(\not{p}_2 + m)X$.

Note that the order of X and $\not{p}_2 + m$ is immaterial.

Finally, when one wants

$$\sum_{\text{spins } 1} \sum_{\text{spins } 2} |(\tilde{u}_2 M u_1)|^2$$

collection and specialization of the previous results is seen to give

$$\begin{aligned} 1/2 \sum_{\text{spins } 1} \sum_{\text{spins } 2} |\tilde{u}_2 M u_1|^2 &= 1/2 \sum_{\text{spins } 1} \sum_{\text{spins } 2} (\tilde{u}_2 M u_1)(\tilde{u}_1 \tilde{M} u_2) \\ &= 1/2 \, \mathrm{Sp}[(\not{p}_2 + m) M (\not{p}_1 + m) \tilde{M}] \end{aligned} \tag{23-3}$$

where the last notation means the spur of the matrix in the brackets.

It is true whether $\not{p}_1$, $\not{p}_2$ represent electrons or positrons.

The following list of the spurs of several frequently encountered matrices may be verified easily

$$\mathrm{Sp}[1] = 4 \qquad \mathrm{Sp}[\gamma_\mu] = 0 \qquad \mathrm{Sp}[xy] = \mathrm{Sp}[yx]$$

$$\mathrm{Sp}[x+y] = \mathrm{Sp}[x] + \mathrm{Sp}[y]$$

$$\mathrm{Sp}[\gamma_\nu \gamma_\mu] = 0 \qquad \text{if } \mu \neq \nu$$

$$= +4 \qquad \text{if } \mu = \nu = 4$$

$$= -4 \qquad \text{if } \mu = \nu = 1, 2, 3$$

$$\mathrm{Sp}[\not{a}\not{b}] = 1/2\ \mathrm{Sp}[\not{a}\not{b} + \not{b}\not{a}] = \mathrm{Sp}[a \cdot b] = 4\ a \cdot b$$

$$\mathrm{Sp}[\not{a}\not{b}\not{c}] = 0$$

It is also true that the spur of the product of any odd number of daggered operators is zero.

$$\mathrm{Sp}[(\not{p}_1 + m_1)(\not{p}_2 - m_2)] = \mathrm{Sp}[\not{p}_1\not{p}_2] + \mathrm{Sp}[m_1\not{p}_2 - \not{p}_1 m_2 - m_1 m_2]$$

$$= 4(p_1 \cdot p_2 - m_1 m_2) \tag{23-4}$$

$$\mathrm{Sp}[(\not{p}_1 + m_1)(\not{p}_2 - m_2)(\not{p}_3 + m_3)(\not{p}_4 - m_4)]$$

$$= 4(p_1 \cdot p_2 - m_1 m_2)(p_3 \cdot p_4 - m_3 m_4) - 4(p_1 \cdot p_3 - m_1 m_3)$$

$$\times (p_2 \cdot p_4 - m_2 m_4) + 4(p_1 \cdot p_4 - m_1 m_4)(p_2 \cdot p_3 - m_2 m_3) \tag{23-5}$$

As an example, the case of Coulomb scattering will be "treated" using this technique.

The cross section for polarized electrons was previously found to be

$$\sigma = (Z^2e^4/Q^4)\,|(\tilde{u}_2\gamma_t u_1)|^2$$

Therefore, since $\tilde{\gamma}_t = \gamma_t$, the cross section for unpolarized electrons is, by Eq. (23-3),

$$\sigma_{unpol} = 1/2\ (Z^2e^4/Q^4)\ Sp[(\not{p}_2 + m)\gamma_t\,(\not{p}_1 + m)\,\gamma_t]$$

The spur can be evaluated immediately from Eq. (23-5) with $m_2 = m_4 = 0$ and $\not{p}_2 = \not{p}_4 = \gamma_t$.

Another way is: Since $\gamma_t\not{p}_1 = 2E_1 - \not{p}_1\gamma_t$, it is seen that

$$(\not{p}_2 + m)\,\gamma_t\,(\not{p}_1 + m)\,\gamma_t = (\not{p}_2 + m)(2E_1\gamma_t - \not{p}_1 + m)$$

Using a few of the formulas listed previously, the spur of this matrix is seen to be

$$-4p_1 \cdot p_2 + 8E_1E_2 + 4m^2$$

But $p_1 \cdot p_2 = E_1E_2 - \mathbf{p}_1 \cdot \mathbf{p}_2$, $\mathbf{p}_1 \cdot \mathbf{p}_2 = p^2 \cos\theta$, and $E_1 = E_2$, so this is

$$4E^2 + 4m^2 + 4p^2 \cos\theta$$

Also $m^2 = E^2 - p^2$, so that finally the cross section becomes

$$\sigma_{\text{unpol}} = 1/2\ (Z^2e^4/Q^4)\,[8E^2 + 4p^2\,(\cos\,\theta\,-\,1)]$$

$$= (4Z^2e^4/Q^4)\,E^2\,[1\,-\,v^2\,\sin^2\,(\theta/2)]$$

where $v^2 = p^2/E^2$.

This is the same cross section obtained previously by other methods.

EFFECTS OF SCREENING OF THE COULOMB FIELD IN ATOMS

The cross sections for the pair production and bremsstrahlung processes contained the factor $[V(Q)]^2$, where V(Q) is the momentum representation of the potential; that is,

$$\mathbf{V(Q)} = \int V(R)\ \exp\,(-i\mathbf{Q}\cdot\mathbf{R})\ d^3\mathbf{R}$$

which for a Coulomb potential is

$$\mathbf{V(Q)} = 4\pi Ze^2/Q^2$$

where **Q** is the momentum transferred to the nucleus or $\mathbf{p}_1$- $\mathbf{p}_2$ - **q**.

Clearly V(**Q**)) gets large as **Q** gets small.

The minimum value of **Q** occurs when all three momenta are lined up (Fig. 23-1):

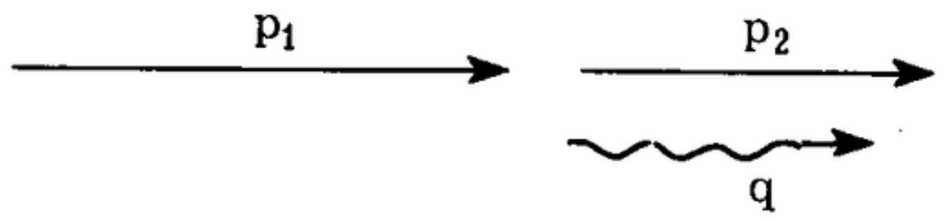


FIG. 23-1

$$|Q_{min}| = p_1 - p_2 - q$$

$$= |p_1| - |p_2| - (E_1 - E_2)$$

For very high energies $E \gg m$,

$$E - p \approx m^2/2E$$

so that in this case

$$Q_{min} = (m^2/2)[(1/E_2) - (1/E_1)] \approx m^2 q/2E_1 E_2$$

From this it is seen that $Q_{min} \to 0$ as $E_1 \to \infty$.

This shows clearly why the cross sections for pair production and bremsstrahlung go up with energy.

From the integral expression for V(Q) it is seen that the main contribution to the integral comes when $R \sim 1/Q$.

So as Q becomes small the important range of R gets large.

It is in this way that screening of the Coulomb field becomes effective.

The value of $1/Q_{min}$ for a contemplated process can be estimated from the foregoing formula.

The atomic radius is given roughly by $a_0 Z^{-1/2}$, where a_0 is the Bohr radius.

Thus if

$$R_{eff} = 1/Q_{min} > a_0 Z^{-1/3}$$

or, what is the same

$$E_1 E_2/q > 1/2\ (137)\ m Z^{-1/3}$$

then screening effect will be important, and vice versa for the opposite inequalities.

If from this estimate screening would appear to be important, one should use the screened Coulomb potential.

It gives the result

$$V(Q) = (4\pi e^2/Q^2)[Z - F(Q)]$$

where F(**Q**) is the atomic structure factor given by

$$F(Q) = \int n(R)\ \exp(-i\mathbf{Q}\cdot\mathbf{R})\ d^3R$$

and n(R) is the electron density as a function of R.

Part 24 Interaction of Several Electrons

Even though the Dirac equation describes the motion of one particle only, we can obtain the amplitude for the interaction of two or more particles from the principles of quantum electrodynamics (so long as nuclear forces are not involved).

For example, if we consider an energetic Compton scattering in which a third, weak photon is emitted as shown in the three diagrams of Fig. 24-1.

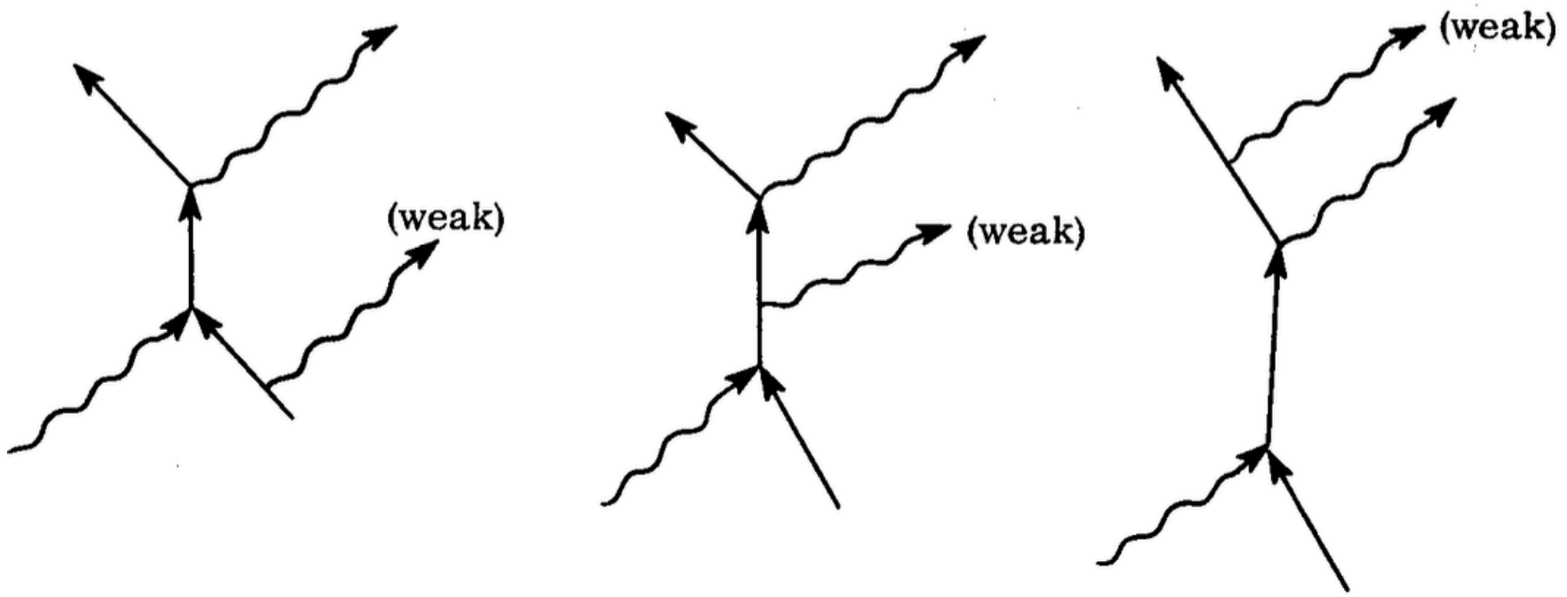


FIG. 24-1

then the cross section for this effect is given by Eq. (24-1), with the Klein-Nishina formula replacing σ_0. (Remember that q is assumed to be small.)

$$\sigma = \sigma_0 \, e^2 \, 4\pi \, d\Omega \, (d\omega/\pi\omega)[p_2 \cdot e/p_2 \cdot (q/\omega) - p_1 \cdot e/p_1 \, (q/\omega)]^2 \qquad (24\text{-}1)$$

Now to work out the details, let us first consider two electrons moving through a region where a potential is present and assume that they do not interact with one another (see Fig. 24-2).

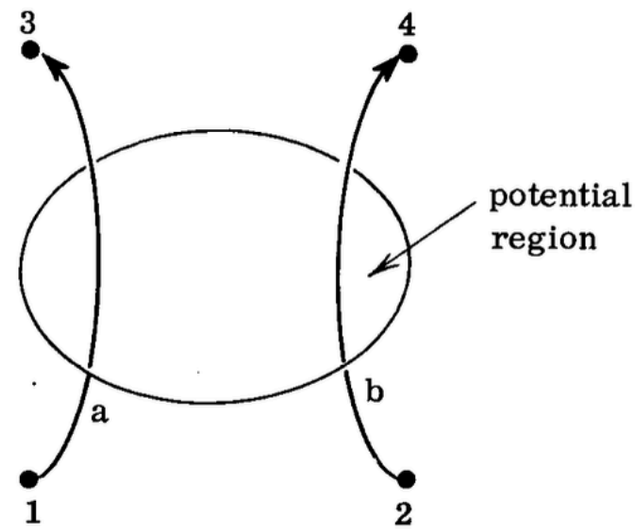


FIG. 24-2

The amplitude for electron a moving from 1 → 3, while electron b moves from 2 - 4 is given the symbol K(3,4;1,2).

If it is assumed that no interaction between electrons takes place, then K can be written as the product of kernels $K_+^{(a)}(3,1)\ K_+^{(b)}(4,2)$, where the superscript means that $K_+^{(a)}$ operates only on those variables describing particle a, and similarly for $K_+^{(b)}$.

A second type of interaction gives a result indistinguishable from the first by any measurement in accordance with the Pauli principle.

This differs from the first case by the interchange of particles between positions 3 and 4 (see Fig. 24-3).

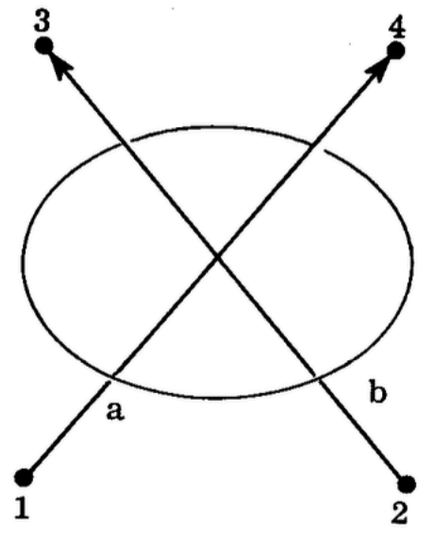


FIG. 24-3

Now the Pauli principle says that the wave function of a system composed of several electrons is such that the interchange of space variables for two particles results in a change of sign for the wave function.

Thus the amplitude (including both possibilities) is K = K, (a) (3,1) K,"b) (4,2) -1, (*)(4,1) K, (b)(3,2).

A similar situation arises in the following occurrence.

Initially, one electron moves into a region where a potential is present.

The potential creates a pair.

Finally one positron and two electrons emerge from the region.

There are two possibilities for this occurrence, as shown in Fig. 24-4.

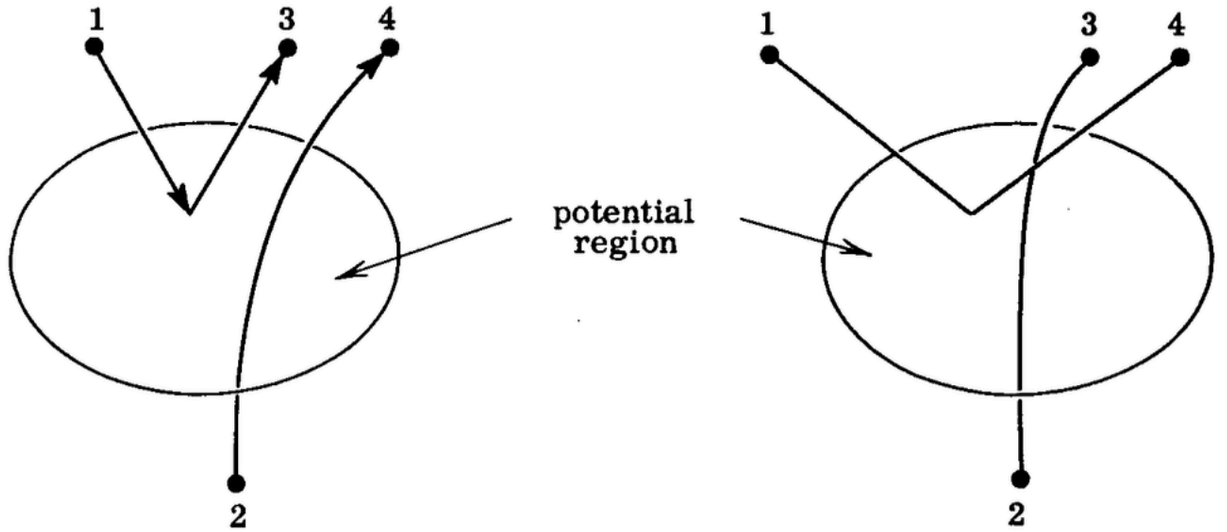


FIG. 24-4

Again, the total amplitude for the occurrence is the difference between the amplitudes for the two possibilities.

the the the the the the the

The probability of this occurrence, or the previous, or any other similar occurrence is given by the absolute square of the amplitude times the number P_V.

The P_V is actually the probability that a vacuum remains a vacuum; because of the possibility of pair production, it is not unity.

The P_V can be computed by making a table of the probabilities of starting with nothing and ending with various numbers of pairs, as is shown in Table 24-1.

TABLE 24-1

Final number of pairs	Probability
0	$P_V \, 1^2$
1	$P_V \, \lvert K_+(2,1)\rvert^2$
2	$P_V \, \lvert K_+(3,1)\, K_+(4,2) - K_+(4,1)\, K_+(3,2)\rvert^2$
3	etc.
etc.	

the

The sum of all these probabilities must equal unity, and P_V is determined from this equation.

The magnitude of P_V depends on the potential present.

So the "probabilities" taken as merely the squares of amplitudes (that is, omitting the P_V factor) are actually relative probabilities for various occurrences in a given potential.

Use of $\delta_+(s^2)$. For the present, the existence of more than one possibility for an occurrence (the Pauli principle) will be neglected.

The total amplitude can always be derived from one by interchanging the proper space variables, making the corresponding changes in sign, and summing all the amplitudes so obtained.

The nonrelativistic Born approximation to the amplitude for an interaction is

$$K(3,4;\,1,2) = K^{(0)} + K^{(1)}$$

where, from earlier notes,

$$K^{(0)} = K_0^{(a)}(3,1)\,K_0^{(b)}(4,2)$$

and

$$K^{(1)} = -i \int K^{(0)}(3,4;5,6)\,V(5,6)\,K^{(0)}(5,6;\,1,2)\,d^3X_5\,d^3X_6\,dt_5$$

Note that $t_5 = t_6$ since a nonrelativistic interaction affects both particles simultaneously.

The potential for the interaction is the Coulomb potential

$$V(5,6) = e^2/r_{5,6}$$

Separate variables may be used for t_5 and t_6, if the function $\delta(t_5 - t_6)$ is included as a factor.

Then

$$K^{(1)} = -i \int\int K_0(3,5)\, K_0(4,6)(e^2/r_{5,6})\delta(t_5 - t_6)\, K_0(5,1)\, K_0(6,2)$$

$$\times\, d\tau_5\, d\tau_6$$

where the differential $d\tau$ includes both space and time variables.

It is conceivable that the relativistic kernel could be obtained by substituting K_+ for K_0, and introducing the idea of a retarded potential by replacing $\delta(t_5 - t_6)$ by $\delta(t_5 - t_6 - r_{5,6})$.

However this δ function is not quite right.

Its Fourier transform contains both positive and negative frequencies, whereas a photon has only positive energy.

Thus

$$\delta(X) = \int_{-\infty}^{\infty} \exp(-i\omega X)\, d\omega/2\pi$$

To correct this, define the function

$$\delta_+(X) = \int_0^{\infty} \exp(-i\omega X)\, d\omega/\pi$$

which contains only positive energy.

The value of the function is determined by the integral.

Thus,

$$\delta_+(X) = \lim_{\epsilon \to 0} (1/\pi i)(X - i\epsilon)$$

$$= \delta(X) + (1/\pi i)(\text{principal value } 1/X)$$

Abbreviating $t_5 - t_6 = t$ and $r_{5,6} = r$, and taking account of the fact that both $t_5 \geq t_6$ and $t_5 \leq t_6$ are possible, the retarded potential is

$$V(5,6) = (e^2/2r)[\delta_+(t-r) + \delta_+(-t-r)]$$

It is straightforward to show that

$$(1/2r)[\delta_+(t-r) + \delta_+(-t-r)] = \delta_+(t^2 - r^2)$$

Defining $t^2 - r^2$ as $s_{5,6}^2$, a relativistic invariant, the potential is $e^2\delta(S_{5,6}^2)$.

Another term which must be included is the magnetic interaction, proportional to $-V_a \cdot V_b$.

In the notation used for the Dirac equation, this product is $-\alpha_a \cdot \alpha_b$.

It will be found convenient to express this in the equivalent form $-(\beta\alpha)_a \cdot (\beta\alpha)_b$, and in this notation the retarded Coulomb potential is proportional to $\beta_a\beta_b$.

These β's come from the use of the relativistic kernel.

Thus the complete potential for the interaction becomes

$$e^2\delta_+(s_{5,6}{}^2)[\beta_a\cdot\beta_b-(\beta\alpha)_a\cdot(\beta\alpha)_b]=e^2\delta(s_{5,6}{}^2)\,\gamma_\mu{}^{(a)}\gamma_\mu{}^{(b)}$$

and then the first-order kernel is

$$\begin{aligned}K^{(1)}(3,4;1,2) &= -ie^2\int\int K_+{}^{(a)}(3,5)\;K_+{}^{(b)}(4,6)\,\gamma_\mu{}^{(a)}\gamma_\mu{}^{(b)}\\ &\qquad\times\delta_+(s_{5,6}{}^2)\;K_+{}^{(a)}(5,1)\;K_+{}^{(b)}(6,2)\;d\tau_5\,d\tau_6\\ &= -ie^2\int\int[K_+(3,5)\,\gamma_\mu\,K_+(5,1)]_a\;\delta_+(s_{5,6}{}^2)\\ &\qquad\times[K_+(4,6)\,\gamma_\mu\,K_+(6,2)]_b\;d\tau_5\;d\tau_6\end{aligned}\tag{24-2}$$

Here the superscript γ_μ indicates on which set of variables the matrix operates, just as for the superscripts on K_+.

The occurrence represented by this kernel can be diagrammed as in Fig. 24-5.

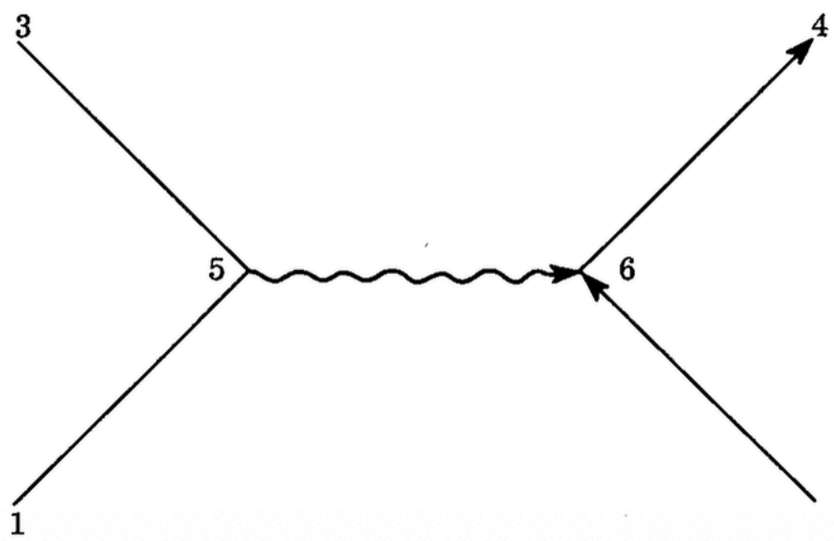


FIG. 24-5

This represents the exchange of a virtual photon between the electrons.

The virtual photon can be polarized in any one of the four directions, t, x, y, z.

Summation over these four possibilities is indicated by the repeated index of $\gamma_\mu\gamma_\mu$.

The integral expression for the kernel, Eq. (24-2), implies that the amplitude for a photon to go from $5 \rightarrow 6$ (or from $6 \rightarrow 5$ depending on timing) is $\delta(s_{5,6}^2)$.

Equation (24-2) can be taken as another statement of the fundamental laws of quantum electrodynamics.

One can also show that

$$\delta_+(s^2) = -4\pi \int [\exp(-ik \cdot X)]\, d^4k/(k^2 + i\epsilon)(2\pi)^4$$

Thus, in momentum space,

$$\delta_+(s^2) \rightarrow -4\pi/k^2$$

Part 25 DERIVATION OF THE "RULES" OF QUANTUM ELECTRODYNAMICS

From the results of the last part, it is evident that the laws of electrodynamics could be stated as follows: (1) The amplitude to emit (or absorb) a photon is $e\gamma_\mu$, and (2) the amplitude for a photon to go from 1 to 2 is $\delta_+(s_{1,2}^2)$, where

$$\delta_+(s_{1,2}^2) = -4\pi \int e^{\frac{-ik\cdot(x_2-x_1)}{k^2+i\epsilon}} \frac{d^4k}{(2\pi)^4}$$

$$= -4\pi/(k^2+i\epsilon) \qquad (25\text{-}1)$$

in momentum representation.

It is interesting to note that $\delta_+(s_{1,2}^2)$ is the same as $I_+(s_{1,2}^2)$, the quantity appearing in the derivation of the propagation kernel of a free particle, with m, the particle mass, set equal to zero.

A more direct connection with the Maxwell equations can be seen by writing the wave equation, $\Box^2 A_\mu = 4\pi J_\mu$ in momentum representation,

$$-k^2 a_\mu = 4\pi j_\mu \qquad \text{or} \qquad a_\mu = -(4\pi/k^2)\, j_\mu \qquad (25\text{-}2)$$

We now consider the connection with the "rules" of quantum electrodynamics given in the second part.

The amplitude for a to emit a photon which b absorbs will now be calculated according to those rules (see Fig 25-1).

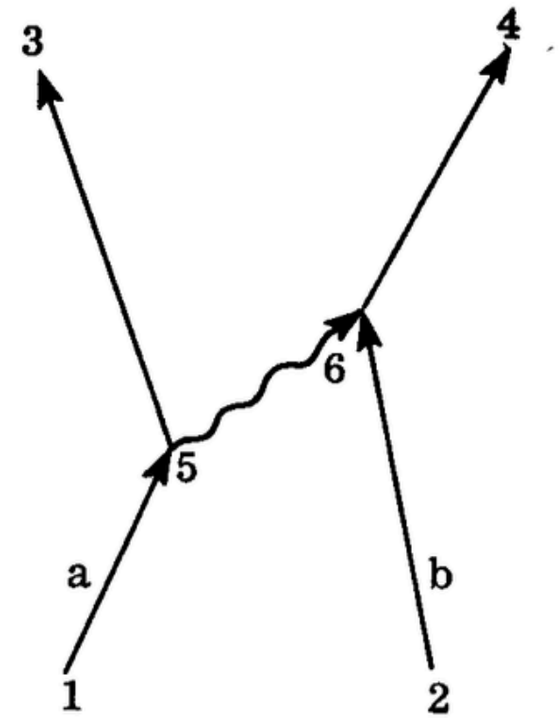


FIG. 25-1

The amplitude that electron a goes from 1 to 5, emits a photon of polarization $\not{\epsilon}$ and direction **K**, then goes from 5 to 3 is given by

$$[K_+(3,5)\,\not{\epsilon}\,\sqrt{(4\pi e^2/2K)}\,\exp(-i\mathbf{K}\cdot\mathbf{r}_5)\,\exp(iKt_5)\,K_+(5,1)]_a$$

whereas the amplitude that b goes from 2 to 6, absorbs a photon of polarization $\not{\epsilon}$ and direction **K** at 6, then goes from 6 to 4 is given by

$$[K_+(4,6)\,\not{\epsilon}\,\sqrt{(4\pi e^2/2K)}\,\exp(i\mathbf{K}\cdot\mathbf{r}_6)\,\exp(-iKt_6)\,K_+(6,2)]_b$$

The amplitude that both these processes occur, which is equivalent to b absorbing a's photon if $t_6 > t_5$ is just the product of the individual amplitudes.

If a absorbs b's photon, the signs of all the exponentials in the preceding amplitudes are changed and t_6 must be less than t_5.

To obtain the amplitude that any photon is exchanged between a and b, it is necessary to integrate over photon direction, sum over possible photon polarizations, and integrate over t_5 and t_6, subject to the aforementioned restrictions.

In summing over polarizations, $\not{\epsilon}$ will be replaced by γ_μ and a summation over μ will be taken.

This amounts to summing over four directions of polarization, something that will be explained later.

Thus

$$\left\{\begin{matrix}\text{Amp. for}\\ \text{photon}\\ a \rightarrow b\end{matrix}\right\} = 4\pi e^2 \sum_\mu \int \exp[-i\mathbf{K}\cdot(\mathbf{r}_5-\mathbf{r}_6)]\ \exp[iK(t_5-t_6)]$$

$$\times [K_+(3,5)\gamma_\mu K_+(5,1)]_a\ [K_+(4,6)\gamma_\mu K_+(6,2)]_b$$

$$\times\ (1/2K)[d^3K/(2\pi)^3]\ dt_5\ dt_6 \qquad t_6 > t_5$$

$$= 4\pi e^2 \sum_\mu \int \exp[i\mathbf{K}\cdot(\mathbf{r}_5-\mathbf{r}_6)] \exp[-iK(t_5-t_6)]$$

$$\times [K_+(3,5)\gamma_\mu K_+(5,1)]_a [K_+(4,6)\gamma_\mu K_+(6,2)]_b$$

$$\times\ (1/2K)[d^3K/(2\pi)^3]\ dt_5\ dt_6 \qquad t_6 < t_5 \qquad (25\text{-}3)$$

Comparing this with the result of the last part, it must be that

$$\delta_+(s_{5,6}^2) = 4\pi \int \exp[-i\mathbf{K}\cdot(\mathbf{r}_5 - \mathbf{r}_6)] \exp[iK(t_5 - t_6)\,(1/2K)$$

$$\times\ [d^3K/(2\pi)^3] \qquad t_6 > t_5$$

$$= 4\pi \int \exp[i\mathbf{K}\cdot(\mathbf{r}_5 - \mathbf{r}_6)]\ \exp[-iK(t_5 - t_6)]\ (1/2K)$$

$$\times\ [d^3K/(2\pi)^3] \qquad t_6 < t_5$$

This can be written in a form which makes the space-time symmetry evident by using the Fourier transform

$$\exp(-iK|t|) = \int_{-\infty}^{\infty} [2iK/(\omega^2 - K^2 + i\epsilon)]\ \exp(-i\omega t)\ d\omega/2\pi$$

so that the foregoing equation becomes

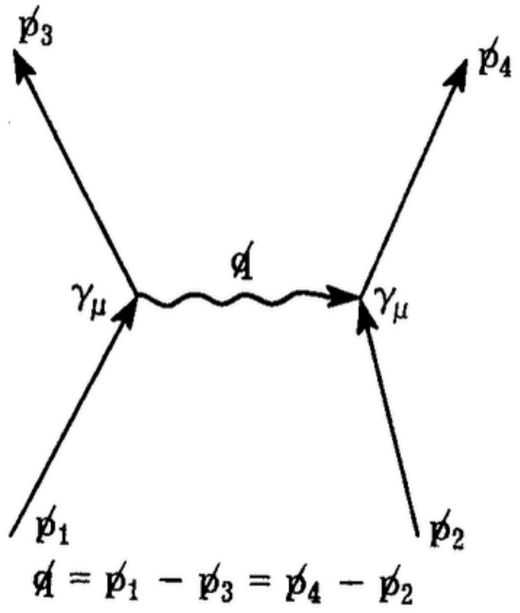
$$\delta_+(s_{5,6}^2) = -4\pi \int \frac{\exp[-ik\cdot(x_5 - x_6)]}{k_4^2 - \mathbf{K}\cdot\mathbf{K} + i\epsilon}\ \frac{d^4k}{(2\pi)^4} \tag{25-4}$$

and comparing this with the result of Part 24 establishes that the rules given in Part 2 are consistent with relativistic electrodynamics developed in the last part.

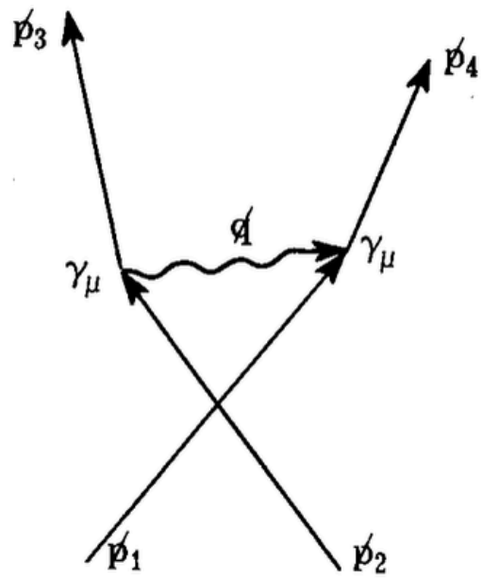
ELECTRON-ELECTRON SCATTERING

The theory will now be used to obtain the electron-electron scattering cross section.

The diagrams for the two indistinguishable processes are shown in Fig. 25-2.



R



S

FIG. 25-2

The amplitude expressed in momentum representation is obtained as follows: write Eq. (25-3) [with the aid of Eq. (25-4)] as

$$e^2 \sum_\mu \int [K_+(3,5)\gamma_\mu K_+(5,1)]_a \frac{-4\pi}{q^2} [K_+(4,6)\gamma_\mu K_+(6,2)]_b \frac{d^4q}{(2\pi)^4}$$

$$\times \, d\tau_5 \, d\tau_6$$

Since electron state 1 is a plane wave of momentum $\not p_1$ and electron state 3 is a plane wave of momentum $\not p_3$, it is clear that in momentum representation the spinor part of the first bracket will become $(\tilde{u}_3 \gamma_\mu u_1)$ and the spinor part of the second bracket will become $(\tilde{u}_4 \gamma_\mu u_2)$.

Integration over τ_5 and τ_6 produces the conservation laws given at the bottom of the diagrams.

Dropping the integration over q puts the photon propagation in momentum representation directly.

Thus the matrix element can be written

$$M = +i4\pi e^2 \left[\frac{(\tilde{u}_4 \gamma_\mu u_2)(\tilde{u}_3 \gamma_\mu u_1)}{(\not{p}_1 - \not{p}_3)^2} - \frac{(\tilde{u}_4 \gamma_\mu u_1)(\tilde{u}_3 \gamma_\mu u_2)}{(\not{p}_4 - \not{p}_2)^2} \right]$$

The first term comes from diagram R, the second from diagram S, and the summation over μ is implied.

In the center-of-mass system, the probability of transition per second is

$$\text{Trans. prob./sec} = \sigma v_1 = \frac{2\pi}{(2E)^4} |M|^2 \frac{E^2 p^3 \, d\Omega}{(2\pi)^3 2Ep^2}$$

(see Density of Final States, Part 19).

The method of Part 23 can be used to average over initial spin states and sum over final spin states.

For example, the sums over spin states that result from $\tilde{R}$ by R matrices and $\tilde{R}$ by S matrices plus R by $\tilde{S}$ matrices are

$$\tilde{R}R \rightarrow \frac{\mathrm{Sp}\left[(\not{p}_4 + m)\gamma_\mu (\not{p}_2 + m)\gamma_\nu\right] \mathrm{Sp}\left[(\not{p}_3 + m)\gamma_\mu (\not{p}_1 + m)\gamma_\nu\right]}{\left[(\not{p}_1 - \not{p}_3)^2\right]^2}$$

$$\tilde{R}S + R\tilde{S} \rightarrow - \frac{\mathrm{Sp}\left[(\not{p}_4 + m)\gamma_\nu (\not{p}_1 + m)\gamma_\mu (\not{p}_3 + m)\gamma_\nu (\not{p}_2 + m)\gamma_\mu\right]}{(\not{p}_1 - \not{p}_3)^2 (\not{p}_4 - \not{p}_1)^2}$$

By judicious use of the spur relations given in Part 23 the following differential cross section is obtained (alternatively, Table 13-1 could be used to calculate M directly):

$$d\sigma = \frac{2e^4 p\, d\Omega}{E^3}\left[\frac{4x^2 + 8x\cos\theta + 2(1-\cos^2\theta) + 4\cos\theta}{(1-\cos\theta)}\right.$$

$$\left. + \frac{4x^2 - 8x\cos\theta + 2(1-\cos^2\theta) - 4\cos\theta}{(1+\cos\theta)^2} - \frac{4(1+x)(x-3)}{(1-\cos\theta)(1+\cos\theta)}\right]$$

where $x = E^2/p^2$.

This is called Möller scattering (see Fig. 25-3).

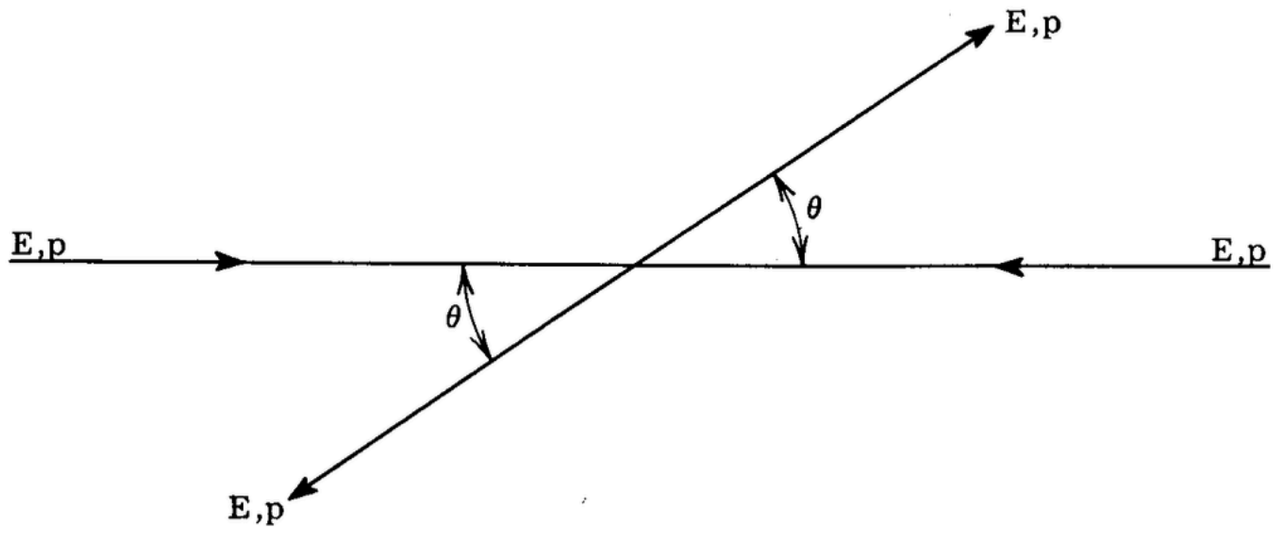


FIG. 25-3

Part 26 Discussion and Interpretation of Various “Correction" Terms

In many processes the behavior of electrons in the quantum-electrodynamic theory turns out to be the same as predicted by simpler theories save for small "correction" terms.

It is the purpose of the present part to point out and discuss a few such cases.

ELECTRON-ELECTRON INTERACTION

The simplest diagrams for the interaction are shown in Fig. 26-1.

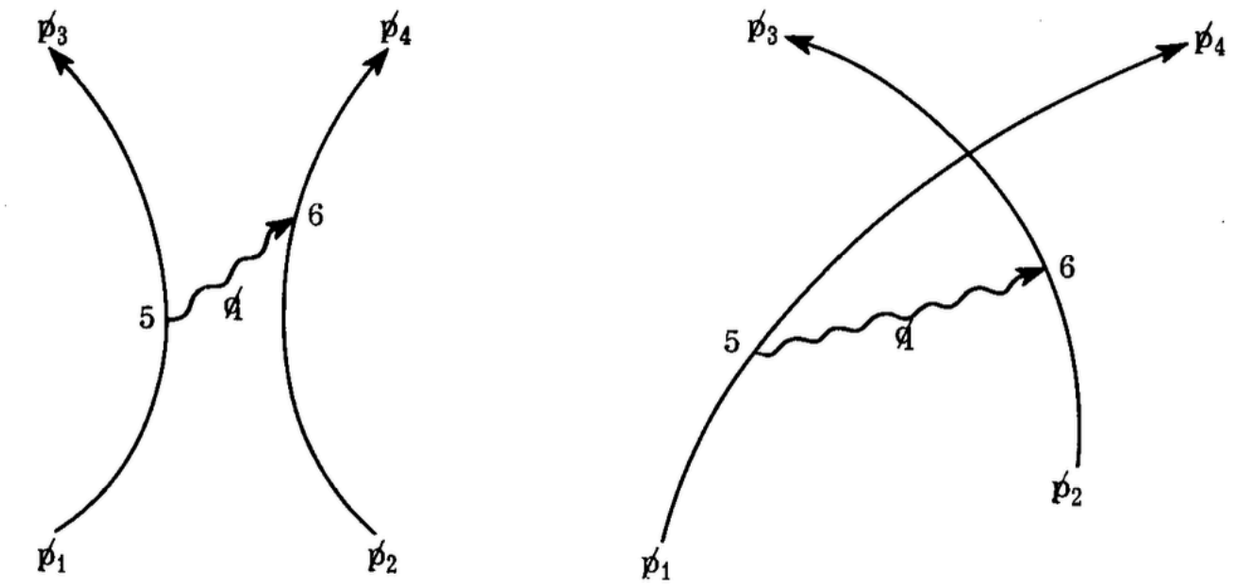


FIG. 26-1

The amplitude for the process has been found to be proportional, in momentum representation, to

$$(\tilde{u}_3 \gamma_\mu u_1)(\tilde{u}_4 \gamma_\mu u_2)/q^2$$

where q = (**Q**, q_4) and Q is the momentum exchanged by the two electrons.

Also, since 4 = py - p it follows that

$$(\tilde{u}_3 \not{q} u_1) = (\tilde{u}_3 (\not{p}_1 - \not{p}_3) u_1) = 0$$

From this identity it was deduced in the last part that the amplitude for the process as just given is equivalent to

$$[-(\tilde{u}_3 \gamma_t u_1)(\tilde{u}_4 \gamma_t u_2)/Q^2] - \sum_{1,2} (\tilde{u}_3 \gamma_{tr} u_1)(\tilde{u}_4 \gamma_{tr} u_2)/q^2$$

By taking the Fourier transform of the first term, it can be seen that it is the momentum representation of a pure, instantaneous Coulomb potential.

The second term then constitutes a correction to the simple Coulomb inter-action.

In it γ_{tr} denotes the γ 's for two directions transverse to the direction of **Q**.

For slow electrons, the correction to the Coulomb potential may be simplified and interpreted in a simple manner.

Note that in this case

$$\mathbf{Q} = \mathbf{p}_1 - \mathbf{p}_3$$

and

$$q_4 = E_1 - E_2 \approx [m + (p_1^2/2m)] - [m + (p_3^2/2m)] = (p_1^2 - p_3^2)/2m$$

$$= [(p_1 + p_3)/2m](p_1 - p_3) \approx v(p_1 - p_3)$$

so that $q_4^2 \sim v^2 Q^2$ and q^2 in the denominator can be replaced by $-Q^2$ with small error.

In the C.G. system, $q_4 = 0$ exactly.

The correction term becomes

$$+\sum_{1,2}(\tilde{u}_3\gamma_{tr}u_1)(\tilde{u}_4\gamma_{tr}u_2)/Q^2$$

but

$$(\tilde{u}_3\gamma_{tr}u_1) \equiv u_3^*\,\alpha_{tr}\,u_1$$

It is recalled that $u \equiv \begin{pmatrix} u_a \\ u_b \end{pmatrix}$, where u_a is the large part and u_p the small part and that in the nonrelativistic approximation

$$u_b \approx (1/2m)(\sigma \cdot \Pi)u_a$$

Also, since

$$\alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix}$$

it follows that (taken between positive energy states)

$$u_3^*\,\alpha_{tr}\,u_1 = \overbrace{u_{3a}^* u_{3b}^*}\begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix}_{tr}\begin{pmatrix} u_{1a} \\ u_{1b} \end{pmatrix} = (u_{3a}^*\,\sigma\,u_{1b} + u_{3b}^*\,\sigma\,u_{1a})_{tr}$$

$$= 1/2m\,[u_{3a}^*\,\sigma\,(\sigma\cdot\Pi_1) + (\sigma\cdot\Pi_3)\sigma\,u_{1a}]_{tr}$$

In free space $\Pi = p$, so the x component, for example, of the foregoing matrix is

$$\sigma_x(\sigma_x p_{1x} + \sigma_y p_{1y} + \sigma_z p_{1z}) + (\sigma_x p_{3x} + \sigma_y p_{3y} + \sigma_z p_{3z})\sigma_x$$

$$= (p_1 + p_3)_x + i\,[\sigma_z(p_1 - p_3)_y - \sigma_y(p_1 - p_3)_z]$$

where the commutation relations for the σ's have been used.

From this it is easily seen that the amplitude for the correction to the Coulomb potential may be written altogether in the form

$$\sum_{1,2} \frac{1}{Q^2} \left\{ u_{3a}^* \left[\frac{\mathbf{p}_1 + \mathbf{p}_3}{2m} - i \frac{\sigma \times (\mathbf{p}_1 - \mathbf{p}_3)}{2m} \right]_{tr} u_{1a} \right\}$$

$$\times \left\{ u_{4a} \left[\frac{\mathbf{p}_4 + \mathbf{p}_2}{2m} - i \frac{\sigma \times (\mathbf{p}_2 - \mathbf{p}_4)}{2m} \right]_{tr} u_{2a} \right\}$$

The first terms in each of the brackets represent currents due to motion of the electron transverse to Q and the second terms represent the transverse components of the magnetic dipole of each.

So altogether it appears that the correction arises from current-current, current-dipole, and dipole-dipole interactions between the electrons.

These interactions are expected even on the basis of classical theory and were described by Breit before quantum electrodynamics, hence are referred to as the Breit interaction.

Consider the dipole-dipole term arising in the correction factor.

Since $\mathbf{Q} = \mathbf{p}_1 - \mathbf{p}_3 = \mathbf{p}_2 - \mathbf{p}_4$ it is

$$\sum_{1,2} (\sigma_1 \times \mathbf{Q})_{tr} \, (\sigma_2 \times \mathbf{Q})_{tr} / Q^2$$

But since $\sigma \times \mathbf{Q}$ is zero when σ and $\mathbf{Q}$ have the same direction, the sum could as well be over all three directions and then it is equivalent to a dot product.

That is, this term of the correction is

$$(\sigma_1 \times Q) \cdot (\sigma_2 \times Q)/Q^2$$

By taking the Fourier transform this will be seen to be the momentum representation of the interaction between two dipoles as was stated.

Note that the approximation $q_4 \sim (v/c)Q$ used above applies only between positive energy states.

For, if one of the states represents a positron, then

$$q_4 = E_1 - E_2 \neq 0$$

$$= 2m$$

However, 2m is very large, so the correction is still small.

It is necessary to redo the analysis nevertheless.

ELECTRON-POSITRON INTERACTION

It would appear that, since the electron and positron are distinguishable, the Pauli principle would not require the interchange diagram, leaving as the only one Fig. 26-2.

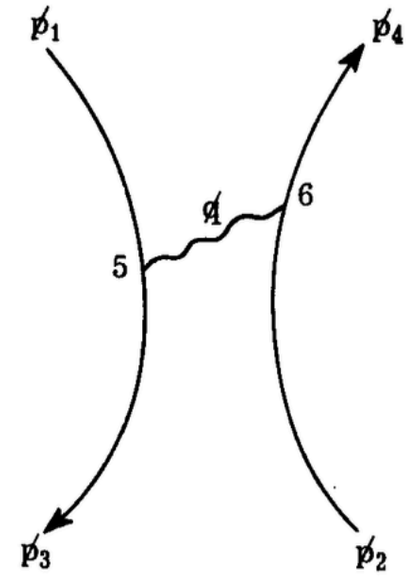


FIG. 26-2

But it is still possible by the same phenomenological reasoning to conceive of the diagram in Fig. 26-3, which would represent virtual annihilation of the electron and positron with the photon later creating a new pair.

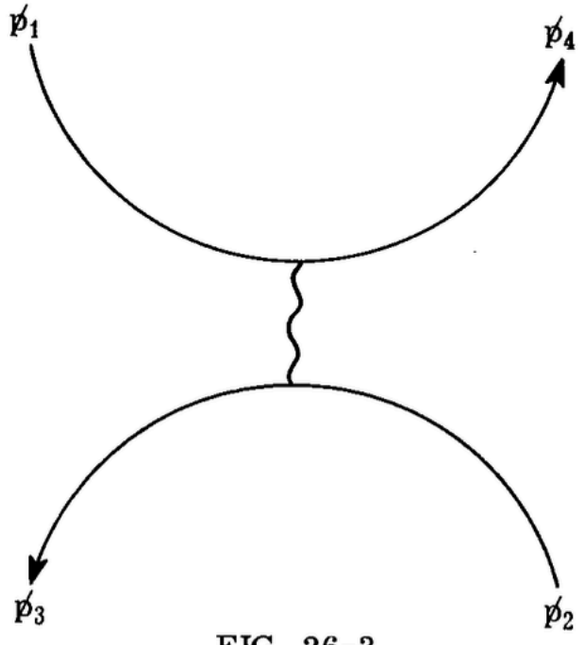


FIG. 26-3

It turns out that it is necessary to regard an electron-positron pair as existing part of the time in the form of a virtual photon in order to obtain agreement with experiment.

From the point of view that positrons are electrons moving backward in time, Fig. 26-3 differs from Fig. 26-2 only in the interchange of the “final" states $\not{p}_3$, $\not{p}_4$.

The Pauli principle extended to this case continues to operate; the amplitudes of the two diagrams must be subtracted, since they differ only in which outgoing (in the sense of the arrows) particle is which.

POSITRONIUM

An electron and positron can exist for a short time in a hydrogenlike bound state known as the atom *positronium*.

The ground state of positronium is an S state and may be singlet or triplet, depending on the spin arrangement.

The ^1S state can annihilate only in two photons, whereas the ^3S state decays only by three-photon annihilation.

The mean life for two-photon annihilation is $1/8 \times 10^9$ sec and for three photons it is $1/7 \times 10^9$ sec.

Figure 26-2 contributes the Coulomb potential holding the positronium together.

The correction term (Breit's interaction) arising from this same diagram contributes a dipole-dipole or spin-spin interaction that is different in the ^3S and ^1S states (the current-current and spin-current interactions are the same for both states).

Thus this amounts to a fine-structure separation of the ^3S and ^1S states which can be shown to be 4.8×10^{-4} ev.

In view of the fact that a photon has spin 1, and the 's state of positronium spin 0, conservation of angular momentum prohibits the process in Fig. 26-3 from occurring in the ^1S state.

It does occur in the ^3S state, however.

The term arising from this diagram is small and, therefore, constitutes another fine-structure splitting of the ^3S and ^1S levels.

It can be shown to amount to 3.7×10^{-4} ev in the same direction as the spin-spin splitting.

It is referred to as splitting due to the "new annihilation force."

In order to calculate the term arising from Fig. 26-3, one needs to compute

$$-(\tilde{u}_4 \gamma_\mu u_1)(\tilde{u}_3 \gamma_\mu u_2)/q^2$$

In this case $q^2 \approx 4m^2$ ($Q = 0$ in the C.G. system), and all matrix elements are 1 or 0 (regarding particles as essentially at rest in the positronium), so the result is just a number.

This means that taking the Fourier transform one gets a δ function of the relative coordinate of the electron and positron for the interaction in real space.

For this reason it is sometimes referred to as the "short-range" interaction of the electron and positron.

The combined fine-structure splitting due to the effects already outlined turns out to be represented by

$$(1/2)\,\alpha^2 \text{ Rydberg } (7/3)$$

where α is the fine-structure constant.

This amounts to 2.044×10^5 Mc, using frequency as a measure of energy.

There is still another correction, however, not yet mentioned, arising from diagrams, such as Fig. 26-4, where the electron or positron may emit and reabsorb its own photon.

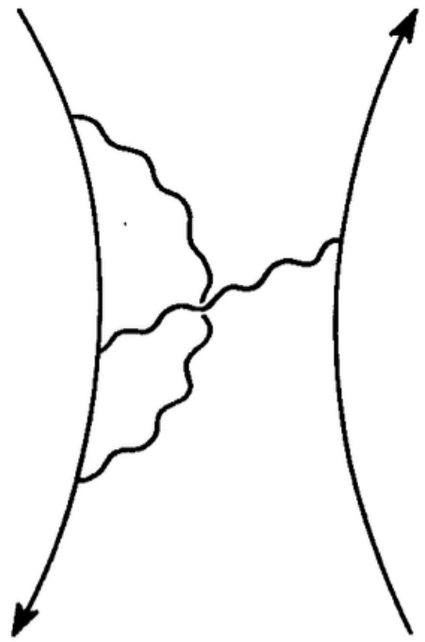


FIG. 26-4

Taking this into account, the fine-structure splitting in positronium is given by

$$(1/2)\ \alpha^2\ \text{Rydberg}\ \{(7/3) - [(32/9) + 2\ \ln\ 2]\ (\alpha/\pi)\}$$

having a value of 2.0337×10^5 Mc.

The experimental value for the positronium fine structure is $2.035 + 0.003$ Mc, so it is seen that this last correction, though of order a smaller than the main terms, is necessary to obtain agreement with experiment.

It is referred to both in positronium and in hydrogen as the Lamb-shift correction because of its experimental observation by Lamb as the source of the small splitting between the $^2S_{1/2}$ and $^2P_{1/2}$ levels in hydrogen.

In general, it comes under the heading of self-action of the electron, to be treated in more detail later.

TWO-PHOTON EXCHANGE BETWEEN ELECTRONS AND/OR POSITRONS

It is easy to imagine that processes, indicated by the diagrams in Fig. 26-5, may occur where two photons instead of one are exchanged.

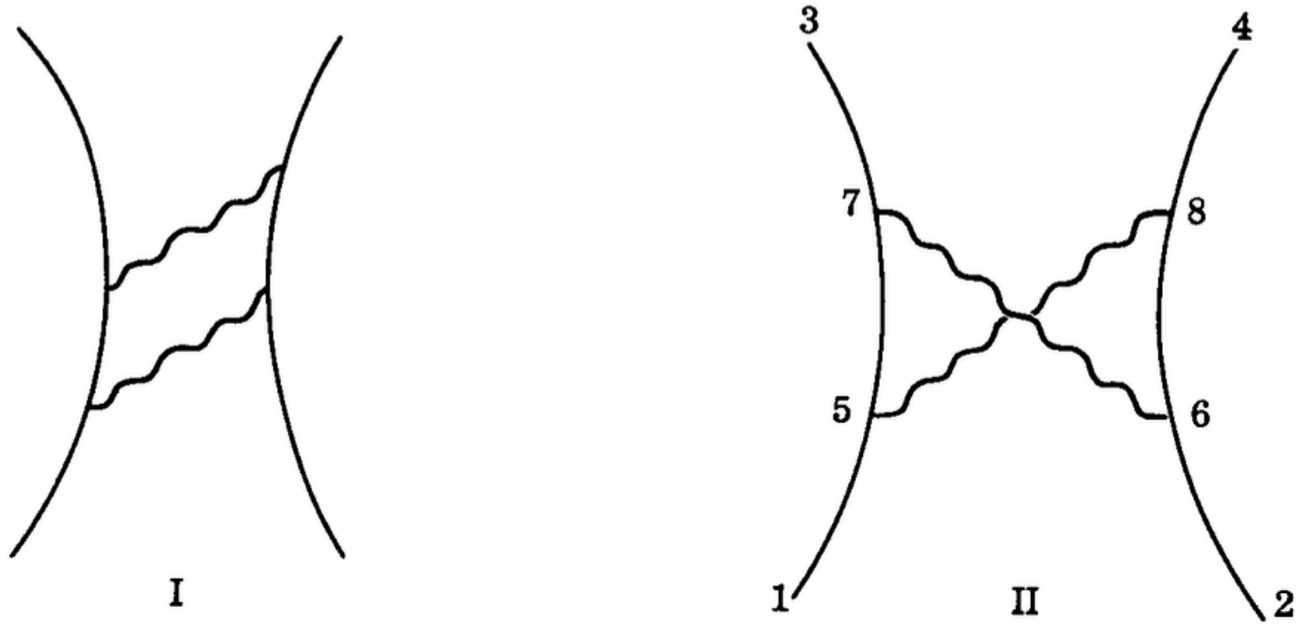


FIG. 26-5

Although it has not been necessary to consider such high-order processes to secure agreement with experiment, it may become necessary as experimental results improve.

The amplitudes for the processes may be written down easily but their calculation is difficult.

The amplitude for case II in space-time representation is, for example,

$$-e^4 \iiint [K_+(3,7)\gamma_\nu\ K_+(7,5)\gamma_\mu K_+(5,1)]\,[K_+(4,8)\gamma_\mu K_+(8,6)\gamma_\nu \times K_+(6,2)]\ \delta_+(s^2_{7,6})\,\delta_+(s^2_{5,8})\ d\tau_5\, d\tau_6\, d\tau_7\, d\tau_8$$

or in momentum representation it is

$$-(4\pi)^2 e^4 \int \left(\tilde{u}_3 \gamma_\nu \frac{1}{\not p_1 - \not k_1 - m}\gamma_\mu u_1\right)\left(\tilde{u}_4 \gamma_\mu \frac{1}{\not p_2 - \not k_2 - m}\gamma_\nu u_2\right) \times \frac{d^4 k_1}{k_1^2 k_2^2 (2\pi)^4}$$

where

$$\not p_2 - \not k_2 + \not k_1 = \not p_4 \qquad \text{or} \qquad \not k_2 = \not p_2 + \not k_1 - \not p_4$$

(see Fig. 26-6).

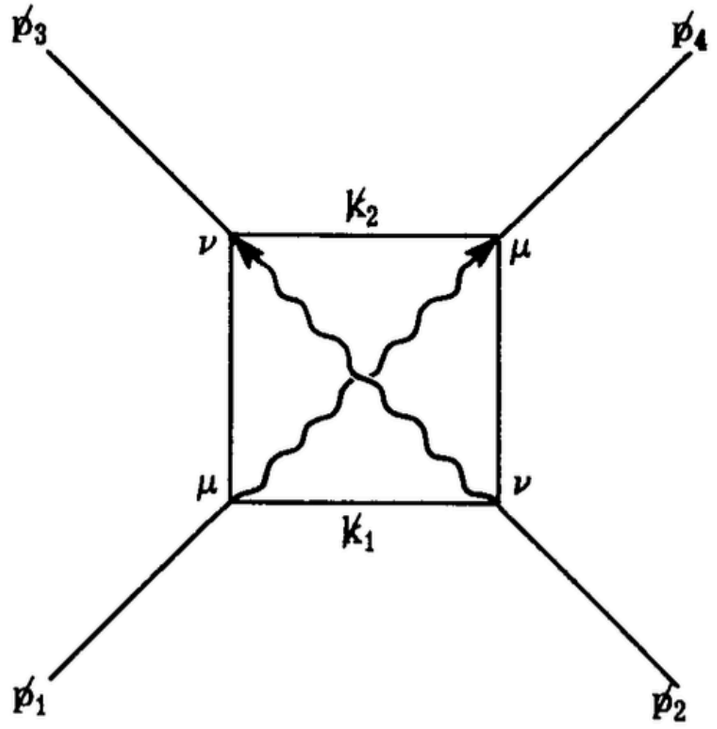


FIG. 26-6

Thus it is possible to determine k_1 and k_2 in terms of each other but not independently; that is, the momentum may be shared in any ratio between the two photons.

It is for this reason that the integral over k_1 arises in the expression for the amplitude.

Part 27 SELF-ENERGY OF THE ELECTRON

In Part 26 the following idea was introduced: An electron may emit and then absorb the same photon, as in Fig. 27-1.

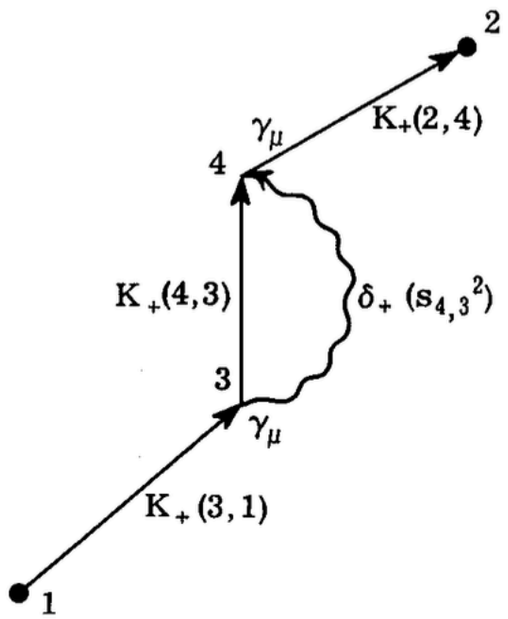


FIG. 27-1

Then the propagation kernel for a free electron moving from point 1 to point 2 should include terms representing this possibility. Including only a first-order term (only one photon is emitted and absorbed), the resulting kernel is

$$K(2,1) = K_+(2,1) - ie^2 \iint K_+(2,4)\,\gamma_\mu\,K_+(4,3)\,\gamma_\mu\,K_+(3,1)\,\delta_+(s_{4,3}^2) \times d\tau_4\,d\tau_3 \tag{27-1}$$

The correction term in this equation is written down by an inspection of the diagram, following the usual procedure for scattering processes.

In the present case, the initial and final momenta are identical.

Therefore the nondiagonal elements in the perturbation matrix will all be zero.

A diagonal element is one in which the resulting wave functions of a particle remain in the same eigenstate.

For time-independent perturbations, it was shown in the development of perturbation theory that the only effect on such wave functions is a change in phase, proportional to the time interval T over which the perturbation is applied.

The resulting wave function is

$$\exp(-iE_n T)\exp[-i(\Delta E)T] \qquad (27\text{-}2)$$

Since the perturbation effect (ΔE)T is small, the second exponential can be expanded as 1 - i(ΔE)T + ... and higher-order terms neglected.

It is the second term of this expansion which is represented by the integral on the right side of Eq. (27-1).

The representation is not yet an equality, since certain normalizing factors are different in the two expressions.

To obtain the correct equation proceed as follows: First, it is clear that the probability of the occurrence depends only on the interval in space and time between points 3 and 4, and not at all on the absolute values of the space and time variables.

So suppose a change of variable is made so that $d\tau_4$, represents the element of interval (in space and time) between 3 and 4.

Then write the integral in Eq. (27-1)

$$\iint \tilde{f}(4)\,\gamma_\mu\,K_+(4,3)\,\gamma_\mu\,\delta_+(s_{4,3}{}^2)\,f(3)\;d\tau_4\;\;d^3x_3\;\;dt_3 \qquad (27\text{-}3)$$

where it is clear that the operators K_+ and δ_+ depend only on the interval 3-4.

Second, expression (27-2) contains the time-dependent part of the wavefunction, exp ($-iE_nt$), because it was assumed that the wave functions used did not contain time factors.

In Eq. (27-3), f(3), f(4) do already include the time-dependent part, so it should be omitted in Eq. (27-2).

Third, the normalization of wave functions is different for the two approaches.

For the development that led to Eq. (27-2), the normalization

$$\int \Psi^*\Psi\;dv = 1$$

was used.

For the present development the normalization is

$$\int u^* u \, dv = (2E/cm^3) \cdot V \tag{27-4}$$

Thus, to establish an equality, expression (27-3) must be divided by the normalizing integral of Eq. (27-4).

The resulting expression is

$$-i \Delta ET = \frac{-ie^2 \iint \tilde{f}(4) \gamma_\mu K_+(4,3) \gamma_\mu \delta_+(s_{4,3}^2) f(3) \, d\tau_4 \, d^3x_3 \, dt_3}{2E \cdot V}$$

The integral over d^3x_3 gives a V which cancels with the denominator, and the integral over dt_3 gives a T which cancels with the left-hand side, so finally

$$2E \, \Delta E = +e^2 \int \tilde{u} \gamma_\mu K_+(4,3) \gamma_\mu \delta_+(s_{4,3}^2) u \, d\tau_4 \tag{27-5}$$

Note that the integral is relativistically invariant. Further, since p is the same before and after the perturbation and $E^2 = m^2 + p^2$, the change in E can be taken as a change in the mass of the electron, from

$$2E \, \Delta E = 2m \, \Delta m$$

Using this expression, and transforming to momentum space,

$$\Delta m = \frac{4\pi e^2}{2mi} \int \tilde{u} \left(\gamma_\mu \frac{1}{\not{p} - \not{k} - m} \gamma_\mu \right) u \, \frac{d^4k}{(2\pi)^4} \, \frac{1}{k^2} \tag{27-6}$$

The integrand may be rewritten from

$$\gamma_\mu \frac{1}{\not p - \not k - m} \gamma_\mu = \frac{\gamma_\mu (\not p - \not k + m) \gamma_\mu}{p^2 - 2p \cdot k + k^2 - m^2} = \frac{2m + 2\not k}{k^2 - 2p \cdot k}$$

using $\not p u = mu$ and the relations of Part 10.

Then Eq. (27-6) becomes

$$\Delta m = \frac{4\pi e^2}{i} \int \frac{2m + 2\not k}{k^2 - 2p \cdot k} \frac{d^4k}{(2\pi)^4} \frac{1}{k^2} \qquad (27\text{-}6')$$

This integral is divergent, and this fact presented a major obstacle to quantum electrodynamics for 20 years.

Its solution requires a change in the fundamental laws.

Thus suppose that the propagation kernel for a photon is the$1/k^2)c(k^2)$ instead of just $(1/k^2)$, where $c(k^2)$ is so chosen that $c(0) = 1$ and $c(k^2) \to 0$ as $k^2 \to \infty$.

In space representation the modification takes the form

$$\delta_+(s_{1,2}{}^2) \to f_+(s_{1,2}{}^2) = \int (1/k^2) c(k^2) \exp(-ik \cdot x)\, d^4k/(2\pi)^4 \qquad (27\text{-}7)$$

The new function f_+ differs significantly from δ_+, only for small intervals. T

This is clear from the fact that if the high-frequency components are removed from the Fourier expansion of a function, only the short-range details are modified.

In the present case the size of the interval over which the function is modified can be described roughly as follows: Consider a large number, λ^2, and suppose that so long as $k^2 << \lambda^2$, $c(k^2) \approx 1$.

Then (from the exponential term) differences will occur when the interval $s^2 \sim 1/\lambda^2$.

Call this value a^2, and the general behavior of f, is shown by Fig. 27-2.

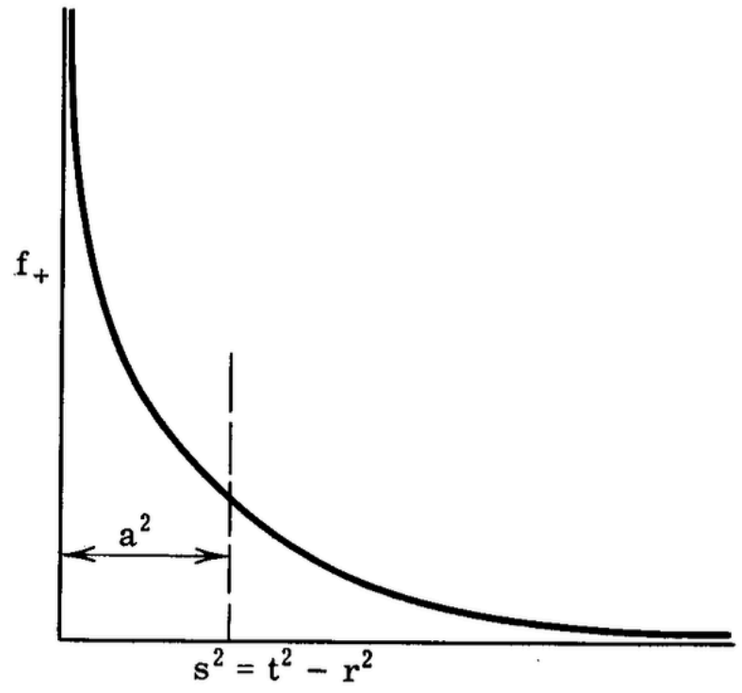


FIG. 27-2

Thus a^2 is sort of a "mean width" of f.

If $a^2 << 1$, as assumed, then when

$$t^2 - r^2 = a^2 \qquad t - r \approx a^2/2r \tag{27-8}$$

which is the size of the interval.

The significance of the form of $f_+(s^2)$ can be understood from the following.

The original function, $\delta_+(s^2)$ differs from zero only when $s^2 = t^2 - r^2 = 0$.

That is to say, an electromagnetic signal can reach a point at distance r only at a time t such that $t^2 - r^2 = 0$ or $t = r$ (i.e., the speed of light is 1).

This is no longer true for $f_+(s^2)$.

The departure is obtained by a measure of t - r.

But, by Eq. (27-8), for all values of $r >> a$ this measure is negligible.

Thus, depending on λ^2, the laws will be found unaffected over any practical distance.

Choosing $\lambda^2 >> m^2$, a practical (and general) representation of $c(k^2)$ is

$$c(k^2) = \int G(\lambda)\, d\lambda\, (-\lambda^2)(k^2 - \lambda^2)^{-1}$$

and the simple form is suggested,

$$c(k^2) = -\lambda^2/(k^2 - \lambda^2)$$

From this, obtain the propagation kernel as

$$1/k^2(-\lambda^2)(k^2-\lambda^2)^{-1} = 1/k^2 - 1/(k^2-\lambda^2)$$

The second term is that for the propagation of a photon of mass λ; however, the minus sign in front of the term has not been explained so far from this point of view.

A convenient representation for this kernel is the integral

$$-\int_0^{\lambda^2} dL/(k^2 - L)^2 \qquad (27\text{-}9)$$

Introducing this kernel into Eq. (27-6) in place of $1/k^2$ gives

$$\int \frac{2m + 2\not{k}}{k^2 - 2p\cdot k} \frac{d^4k}{k^2}\left(\frac{-\lambda^2}{k^2-\lambda^2}\right) \qquad (27\text{-}10)$$

which can be written as the sum of two integrals, which differ only by having m or k in the numerator, that is, m or k_σ (since $\not{k} = k_\sigma\gamma_\sigma$).

METHOD OF INTEGRATION OF INTEGRALS APPEARING IN QUANTUM ELECTRODYNAMICS

We shall need to do many integrals of a form similar to the preceding one.

A method has been worked out to do these fairly efficiently.

We now stop to describe this method of integration.

Everything will be based on the following two integrals due to Feynman:

$$\int_{-\infty}^{\infty} \frac{(1; k_\sigma)\, d^4k}{(2\pi)^4 (k^2 + i\epsilon - L)^3} = (32\pi^2 iL)^{-1} (1;0) \qquad (27\text{-}11)$$

$$\int_0^1 [ax + b(1-x)]^{-2}\, dx = 1/ab \qquad (27\text{-}12)$$

In Eq. (27-11), to write a little more compactly, we use the notation $(1; k_\sigma)$ to mean that either 1 or k_σ in the numerator, in which case, on the right and side the (1; 0) is 1 or 0, respectively.

To prove the first of these, note that, if k_σ is in the numerator, the integrand is an odd function.

Thus the integral is zero.

With 1 in the numerator, contour integration is employed.

Write the integral

$$\int\int\int_{-\infty}^{\infty}\int [\omega^2 + i\epsilon - (L + k^2)]^{-3}\, d\omega\, d^3\mathbf{k}$$

Then for $\epsilon << L + k^2$, there are poles at $\omega = [L + k^2)^{1/2} - i\,\epsilon]$, and contour integration of ω gives

$$\int_{-\infty}^{\infty}[\omega^2 + ie - (L+k^2)]^{-3}\, d\omega = (6\pi/16i)(L + k^2)^{-5/2}$$

Then the remaining integral is

$$\int\int_{-\infty}^{\infty}\int (L+k^2)^{-5/2}\, d^3\mathbf{k} = 4\pi \int_0^{\infty}(L+k^2)^{-5/2}\, k^2\, dk$$

$$= 4\pi\, [k^3/3L(L+k^2)^{3/2}]\Big|_0^{\infty} = 4\pi/3L$$

which proves Eq. (27-11).

If k-p is substituted for the variable of integration in Eq. (27-11), the result is

$$\int_{-\infty}^{\infty} \frac{(1;k_\sigma)\, d^4k}{(2\pi)^4(k^2 - 2p\cdot k - \Delta)^3} = [32\pi^2\, i(p^2 + \Delta)]^{-1}\,(1;p_\sigma) \qquad (27\text{-}13)$$

By differentiating both sides of Eq. (27-13) with respect to Δ or with respect to p_j, there follows directly

$$\int_{-\infty}^{\infty} \frac{(1;k_\sigma;k_\sigma k_j)d^4k}{(2\pi)^4\,(k^2 - 2p\cdot k - \Delta)^4} = -\frac{[1;p_\sigma;p_\sigma p_j - (1/2)\delta_{\sigma j}(p^2 + \Delta)]}{96\pi^2\, i(p^2 + \Delta)^2}$$

Further differentiations give directly successive integrals including more k factors in the numerator and higher powers of $(k^2 - 2\mathbf{p}\cdot\mathbf{k} - \Delta)$ in the denominator.

Part 28 SELF-ENERGY INTEGRAL WITH AN EXTERNAL POTENTIAL

In the last part it was found that the self-energy of the electron is equivalent to a change in mass

$$\Delta m = \frac{4\pi e^2}{2mi} \int \frac{\tilde{u}(2m + 2\not{k})\, u}{k^2 - 2p \cdot k} \left(\frac{-\lambda^2}{k^2 - \lambda^2} \right) \frac{1}{k^2} \frac{d^4 k}{(2\pi)^4} \tag{28-1}$$

and that this could also be expressed in terms of integrals,

$$I = -\int_0^{\lambda^2} dL \int \frac{(1;k_\sigma)}{(k^2 - 2p \cdot k)(k^2 - L)^2} \frac{d^4 k}{(2\pi)^4} \tag{28-2}$$

t was also found that

$$\int \frac{(1;k_\sigma)}{(k^2 - 2p \cdot k - \Delta)^3} = (32\pi^2 i)^{-1} (p^2 + \Delta)^{-1} \tag{28-3}$$

Using the definite integral

$$\frac{1}{ab^2} = \int_0^1 \frac{2(1 - x)\, dx}{[ax + b(1 - x)]^3} \tag{28-4}$$

the denominator of the integrand of Eq. (28-2) may be expressed as

$$\frac{1}{(k^2 - 2p \cdot k)(k^2 - L)^2} = \int_0^1 \frac{2(1 - x)\, dx}{[k^2 - 2xp \cdot k - L(1 - x)]^3}$$

so that Eq. (28-2) becomes

$$I = -\int_0^{\lambda^2} dL \int\int_0^1 \frac{d^4 k\, (1;k_\sigma)\, 2(1 - x)\, dx}{[k^2 - 2xp \cdot k - L(1 - x)]^3\, (2\pi)^4} \tag{28-5}$$

The integral over k can be done by using Eq. (28-3) with the substitutions xp for p and L(1 - x) for Δ, giving

$$I = -\int_0^{\lambda^2} dL \int_0^1 \frac{(1;p_\sigma)\,2(1-x)\,dx}{[32\pi^2 i]\,[x^2 p^2 + L(1-x)]} \qquad p^2 = m^2$$

The integral over L is elementary and gives

$$I = -2(32\pi^2 i)^{-1} \int_0^1 dx\,(1;xp_\sigma)\,\ln\,[(1-x)\lambda^2 + m^2x^2/m^2x^2]$$

When $\lambda^2 >> m^2$, it is legitimate to neglect m^2x^2 in the numerator [it is true that when $x \approx 1$, $(1 - x)^2\,\lambda^2$ is not much larger than m^2x^2, but the interval over which this is true is so small, for $\lambda^2 >> m^2$, that the error is small], so that, when the x integration is performed, using

$$\int_0^1 \ln\,[x^{-2}(1-x)]\,dx = 1 \qquad \int_0^1 x\,\ln\,[x^{-2}(1-x)]\,dx = -1/4$$

we get

$$I \approx -(32\pi^2 i)^{-1}\{\,2[\ln(\lambda^2/m^2) + 2];\, p_\sigma[\ln(\lambda^2/m^2) - 1/2]\}$$

$$\lambda >> m^2$$

The change in mass is [from Eq. (28-1)]

$$\Delta m = (4\pi^2/2mi)(-32\pi^2 i)^{-1}\,(\tilde{u}\,\{2m[2\,\ln\,(\lambda^2/m^2) + 2]$$

$$+ 2\not{p}\,[\ln\,(\lambda^2/m^2) - (1/2)]\}u)$$

Since $\not{p}u = mu$ and $(\tilde{u}u) = 2m$, this can be simplified to

$$\Delta m/m = (e^2/2\pi)\,[3 \ln (\lambda/m) + (3/4)] \qquad (28\text{-}6)$$

Now $(e^2/2\pi)$ is about 10^{-3}, so that even if λ is many times m, the fractional change in mass will not be large.

The interpretation of this result is as follows.

There is a shift in mass which depends on λ and hence cannot be determined theoretically.

One can imagine an experimental mass and a theoretical mass which are related by

$$m_{exp} = m_{th} + \Delta m \qquad (28\text{-}7)$$

All our measurements are of m_{exp}, that is, self-action is included, and m_{th}, the mass without self-action, cannot be determined.

More accurately stated,

$$\left\{\begin{array}{l}\text{A theory using } m_{th} \text{ and} \\ e^2/\hbar c \text{ self-action}\end{array}\right\} \quad \text{is equivalent to} \quad \left\{\begin{array}{l}\text{a theory using } m_{exp}\text{, } \textit{plus} \\ e^2/\hbar c \text{ self-action } \textit{minus} \\ \Delta m \text{ as computed for a} \\ \text{free particle}\end{array}\right\}$$

When the electron is free, the $e^2/\hbar c$ self-action term exactly cancels the Δm term and a theory using m_{exp} is exactly correct.

When the electron is not free, $e^2/\hbar c$ self-action is not quite equal to the Δm term and there is a small correction to a theory using m_{exp}.

This effect leads to the Lamb shift in the hydrogen atom, and, in order to calculate such effects, we shall now consider the effect of self-action on the scattering of an electron by an external potential.

SCATTERING IN AN EXTERNAL POTENTIAL

The diagram for scattering in an external potential is shown in Fig. 28-1, and the relationships for this process, excluding the possibility of self-action, are as follows:

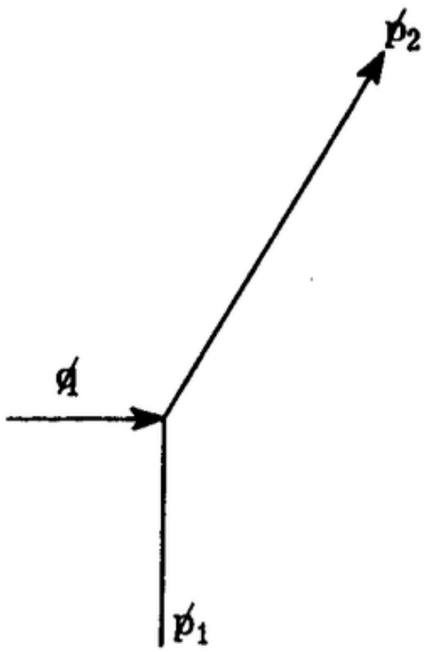


FIG. 28-1

Potential: $\not{A}(q) = \gamma_t (4\pi Ze/Q^2)\,\delta(q_4)$ **for Coulomb potential**
Matrix element: $M = -ie(\tilde{u}_2 \not{A} u_1)$
Conservation relation: $\not{p}_2 = \not{p}_1 + \not{q}$

First-order self-action will produce the diagrams shown in Fig. 28-2.

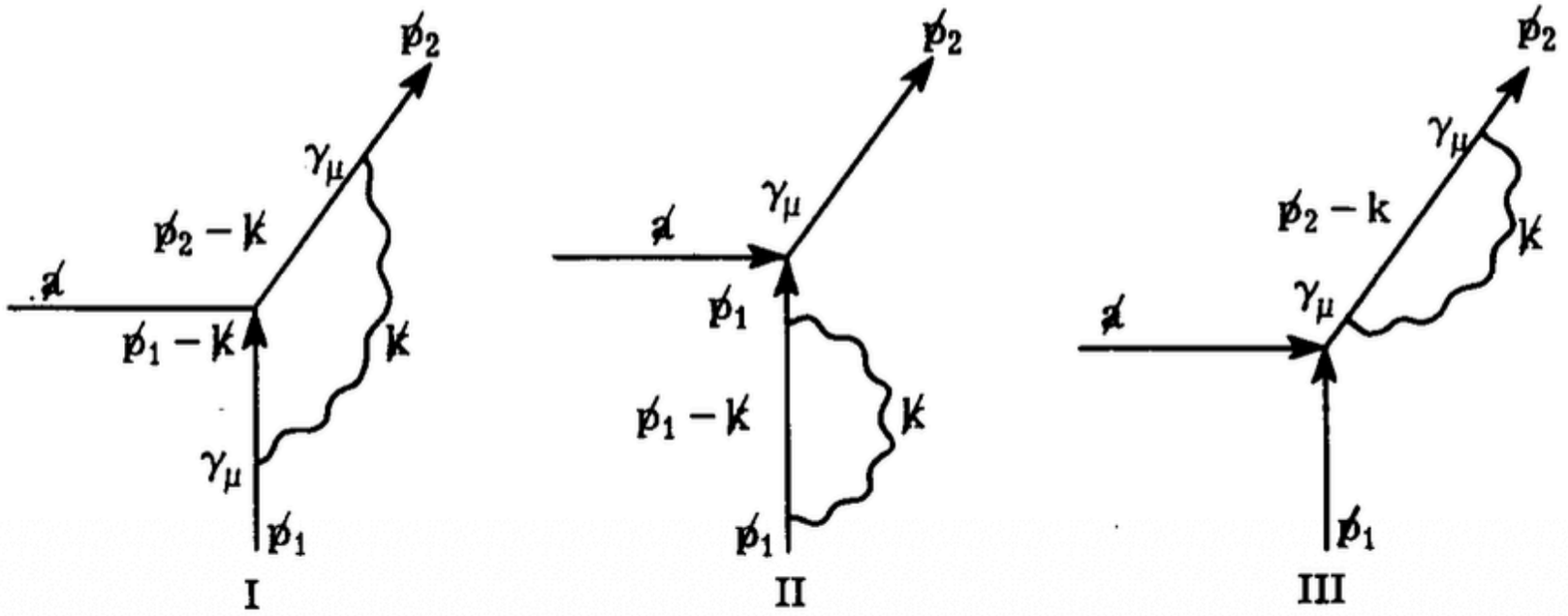


FIG. 28-2

The amplitude for process is obtained in the usual manner.

For example, diagram I gives

$$I_1 = \frac{4\pi e^2}{i}\int\left(\tilde{u}_2\gamma_\mu \frac{1}{\not p_2 - \not k - m} \not a \frac{1}{p_1 - \not k - m}\gamma_\mu u_1\right)(k^2)^{-1}(2\pi)^{-4}\, d^4k$$

Rationalizing the denominators and inserting the convergence factor, this becomes

$$I_1 = \frac{4\pi e^2}{i}\int \frac{(\tilde{u}_2\gamma_\mu[\not p_2 - \not k + m]\not a[\not p_1 - \not k + m]\gamma_\mu u_1)}{(k^2 - 2p_2\cdot k)(k^2 - 2p_1\cdot k)} \quad \frac{-\lambda^2}{k^2 - \lambda^2} \qquad (28\text{-}8)$$

$$\times (2\pi)^{-4}\frac{d^4k}{k^2}$$

This expression also happens to diverge for small photon momenta (k) (a result which has been called the "infrared catastrophe, " but which has a clear physical interpretation, discussed later).

Temporarily the k^2 under dk will be replaced by $(k^2 - \lambda^2_{min})$, where $\lambda^2_{min} << m^2$, to make the integral convergent.

This is equivalent to cutting off the integral somewhere near $k = \lambda_{min}$, and the physical interpretation is left to Parts 29 and 30.

To facilitate the integration over k, the following identity is used:

$$-\int_{\lambda^2_{min}}^{\lambda^2} (k^2 - L)^{-2}\, dL = \frac{1}{k^2 - \lambda^2_{min}} - \frac{1}{k^2 - \lambda^2}$$

$$= \frac{\lambda^2_{min} - \lambda^2}{(k^2 - \lambda^2_{min})(k^2 - \lambda^2)} \approx \frac{-\lambda^2}{k^2 - \lambda^2}$$

$$\times \frac{1}{k^2 - \lambda^2_{min}}$$

since $\lambda^2 >> m^2 >> \lambda^2_{min}$.

This substitution produces integrals of the form

$$-\int_{\lambda^2_{min}}^{\lambda^2} dL \int \frac{(1; k_\sigma; k_\sigma k_\tau)(2\pi)^{-4}\, d^4 k}{(k^2 - 2p_1 \cdot k)(k^2 - 2p_2 \cdot k)(k^2 - L)^2}$$

To evaluate these integrals, we make use of the identity

$$(ab)^{-1} = \int_0^1 dy/[ay + b(1 - y)]^2$$

so that

$$\frac{1}{(k^2 - 2p_1 \cdot k)(k^2 - 2p_2 \cdot k)} = \int_0^1 \frac{dy}{(k^2 - 2\not{p}_y \cdot k)^2}$$

where $\not{p}_y = y\not{p}_1 + (1 - y)\not{p}_2$.

Performing integrations in the order, k, L, y, and using the appropriate integrals in Eq. (28-6) gives as the matrix to be taken between states u_2 and u_1

$$M_1 = \frac{e^2}{2\pi}\left[2\left(\ln\frac{m}{\lambda_{min}} - 1\right)\left(1 - \frac{2\theta}{\tan 2\theta}\right) + \theta\tan\theta + \frac{4}{\tan 2\theta}\right.$$

$$\left. \times \int_0^\theta \alpha\tan\alpha\ d\alpha\right]\not{q}$$

$$+ \frac{e^2}{2\pi}\left[\frac{1}{4m}(\not{q}\not{q} - \not{q}\not{q})\frac{2\theta}{\sin 2\theta} + r\not{q}\right]$$

where $r = \ln(\lambda/m) + 9/4 - 2\ln(m/\lambda_{min})$ and $4m^2\sin^2\theta = q^2$.

It is shown in Part 30 that diagrams II and III (Fig. 28-2) produce a contribution $M_2 + M_3 = -(e^2/2\pi)r\not{q}$, which just cancels a similar term in M_g.

When q is small, $\theta \approx (q^2)^{1/2}/2m$, and the sum $M_1 + M_2 + M_3$ can be approximated by

$$M \approx \frac{e^2}{4\pi}\left[\frac{1}{2m}\,(\not{q}\not{a} - \not{a}\not{q}) + \frac{4q^2}{3m^2}\,\not{a}\left(\ln\frac{m}{\lambda_{min}} - \frac{3}{8}\right)\right] \tag{28-10}$$

The $(\not{q}\not{a} - \not{a}\not{q})$ can be written out

$$(\not{q}\not{a} - \not{a}\not{q}) = \gamma_\mu\gamma_\nu\,(q_\mu a_\nu - a_\mu q_\nu)$$

But q_μ is the gradient operator so this can be written, in coordinate representation,

$$\gamma_\mu\gamma_\nu\,(\nabla_\mu A_\nu - \nabla_\nu A_\mu) = +\,\gamma_\mu\gamma_\nu F_{\mu\nu}$$

[see Eq. (7-1)].

Reference earlier work shows that the effect of a particle's having an anomalous magnetic moment is to subtract a potential $\mu\gamma_\mu\gamma_\upsilon F_{\mu\upsilon}$ from the ordinary potential $\not{a} = \gamma_\mu A_\mu$, appearing in the Dirac equation.

Since this is precisely what the first term of Eq. (28-10) does, one can say that this part of the self-action correction looks like a correction to the electron's magnetic moment, so that

$$\mu_{elec} = (e/2m)[1 + (e^2/2\pi)]$$

Note that this result [and (28-9) and (28-10)] does not depend on the cutoff λ, and hence λ can now be taken to be infinity

Part 29

It has been shown that when a particle is scattered by a potential, the primary effect is that of $\not{a}$, and that for diagram I (Fig. 28-2) a correction term arises which

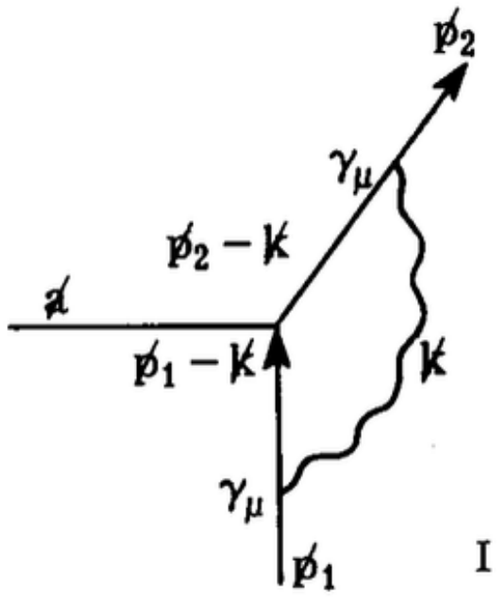


FIG. 28-2

$$\frac{e^2}{2\pi}\left[2\left(\ln\frac{m}{\lambda_{min}}-1\right)\left(1-\frac{2\theta}{\tan 2\theta}\right)+\theta\tan\theta+\frac{4}{\tan 2\theta}\right.$$

$$\left.\times\int_0^\theta \alpha\tan\alpha\right]\not{a}+\frac{e^2}{8\pi m}(\not{q}\not{a}-\not{a}\not{q})\frac{2\theta}{\sin 2\theta}+\frac{e^2}{2\pi}r\not{a}$$

It remains to show that the combined effect of diagrams II and III (Fig. 28-2),

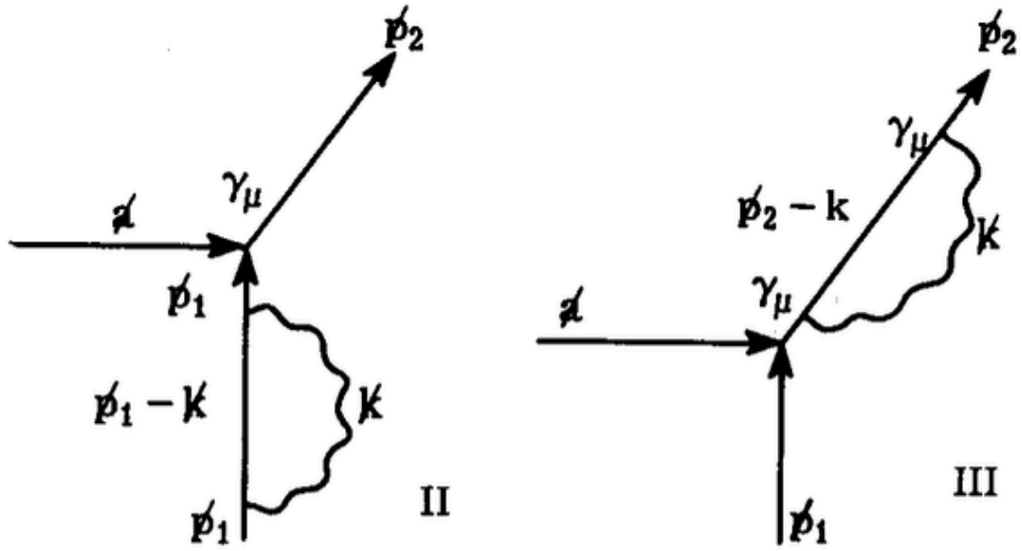


FIG. 28-2

when considered along with the effect of the mass correction, is another correction term,

$$-(e^2/2\pi)r\not{a}$$

just canceling the last term in the preceding expression.

It is recalled that the necessity for considering the effect of the mass correction together with the self-action represented in diagrams I, II, and III is that the theory being developed must contain the experimental mass rather than the “theoretical" mass.

Suppose that in the Dirac equation

$$(i\not{\nabla} - m_{th})\Psi = e\not{A}\Psi$$

m_{th}, the theoretical mass, is replaced by m - Δm, where m is the experimental mass; then

$$(i\not{\nabla} - m)\Psi = e(\not{A} + \Delta m)\Psi$$

The mass correction Δm is just a number, so that in momentum representation it is a δ function of momentum.

Hence from the form of the foregoing equation, it is seen to behave like a potential with zero momentum and involves no matrices.

Diagrammatically its effect may be represented as in Fig. 29-1.

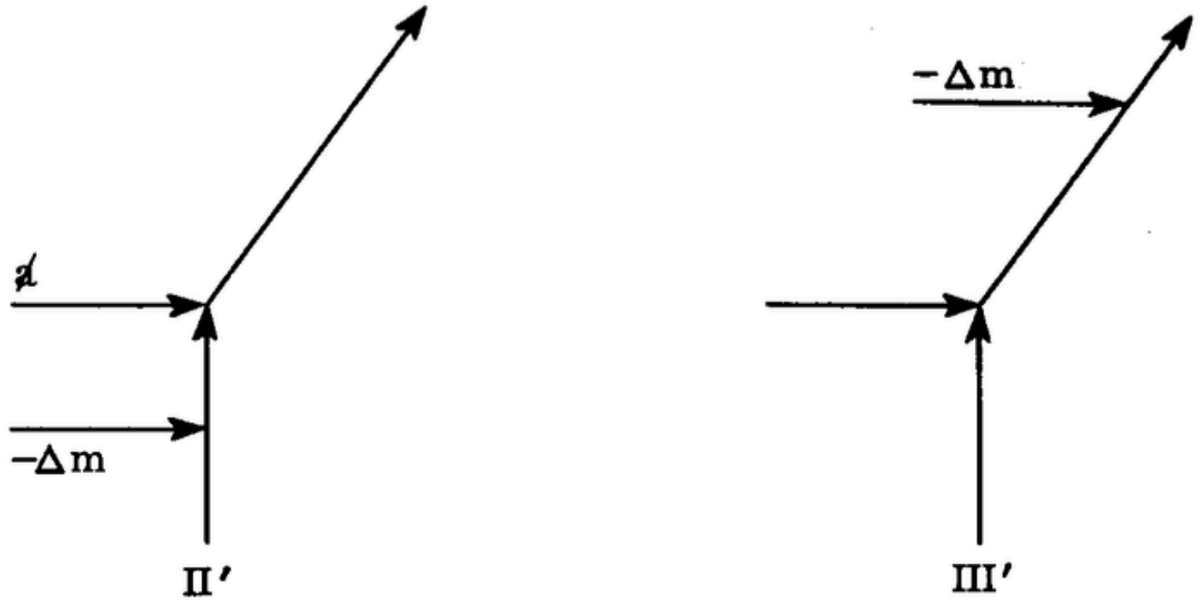


FIG. 29-1

The minus sign is used because the effect of the mass correction Δm is to be subtracted from the results obtained from diagrams I, II, an III (Fig. 28-2) alone. For diagram II the amplitude would appear to b

For diagram II the amplitude would appear to be

$$\tilde{u}_2 \not{p} \frac{1}{\not{p}_1 - m}\left(\frac{4\pi e^2}{i} \int \gamma_\mu \frac{1}{\not{p}_1 - \not{k} - m} \gamma_\mu \frac{1}{k^2 - \lambda_{min}} \frac{d^4k}{(2\pi)^4} \frac{-\lambda}{k^2 - \lambda^2}\right)$$

and for diagram II' (Fig. 29-1),

$$-\tilde{u}_2 \not{p} [1/(\not{p}_1 - m)] (\Delta m) u_1$$

But the part of the amplitude for diagram II (Fig 28-2) contained in the parentheses is just Δm u_1 so that II and II' seem to cancel.

A similar result applied for diagrams III and III'.

This is an error, however, arising from the fact that both of these amplitudes are infinite, owing to the factor p-m in the denominator.

Hence their difference is indeterminate.

But by subtracting them properly it will be found that their difference does not vanish.

The method proposed to accomplish this subtraction will, in fact, give the combined effect of the self-action and mass correction of both diagrams II and III and Il' and III'.

It is based on the fact that an electron is never actually free.

An electron's history will have always involved a series of scatterings, as will its future.

These scatterings will be considered as occurring at long but finite time intervals.

It will be sufficient to calculate the effect of self-action and the mass correction between any two of these scatterings, since the result will evidently be the same between each pair of them.

Then, the effect will be accounted for simply by regarding a correction, equal to that calculated for one of the intervals between scatterings, as being associated with the potential at each scattering (number of intervals equals number of scatterings).

Then, considering a single scattering event as here, this correction to the potential represents all the effects of diagrams II, III, II', and Ill'.

For an electron which is not quite free, $p^2 \neq m^2$ exactly, but instead

$$p^2 = m^2(1 + \epsilon)^2$$

where

$$m\epsilon = \hbar/T$$

by the uncertainty principle, and T is the interval between scatterings.

Since T is large, ϵ is a small quantity.

Let $p = (1 + \epsilon)\not{p}_0$, where $\not{p}_0$ is the momentum of a free electron.

If $\not{a}$ and $\not{b}$ are the momentum representatives of the scattering potentials at a and b (any two scatterings), then the matrix of the amplitude to go from the initial state at a to the final state at b without any perturbations is

$$\not{b}\,\frac{1}{\not{p} - m}\,\not{a} = \not{b}\,\frac{\not{p} + m}{p^2 - m^2}\,\not{a} = \frac{\not{b}(\not{p} + m)\not{a}}{2m^2\epsilon}$$

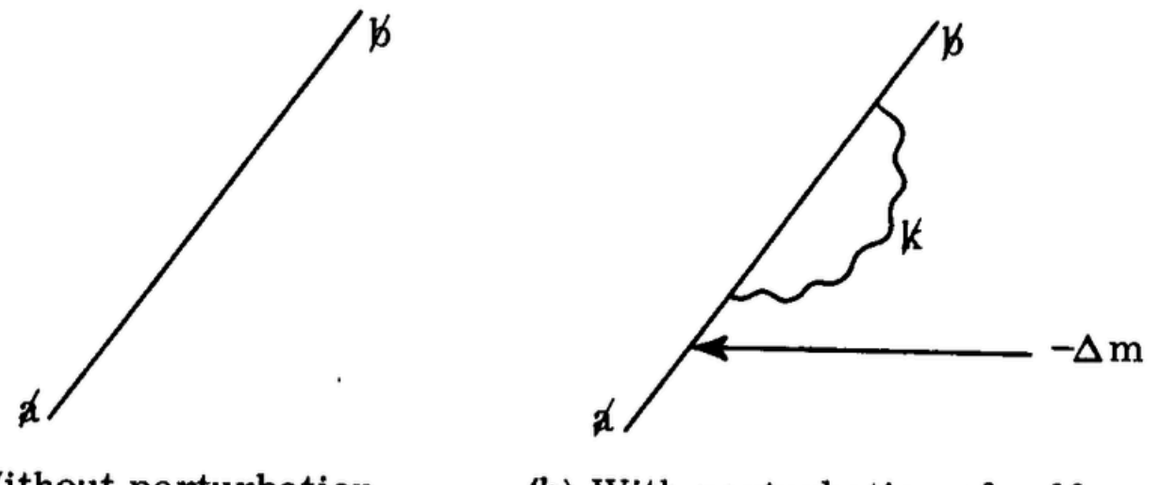
up to terms of order ϵ.

With the perturbations of self-action and mass correction, this matrix is

$$i4\pi e^2 \int \not b \frac{1}{\not p - m} \gamma_\mu \frac{1}{\not p - \not k - m} \gamma_\mu \frac{1}{\not p - m} \not a \frac{d^4k}{k^2 - \lambda^2_{min}} \frac{-\lambda^2}{k^2 - \lambda^2}$$

$$- \not b \frac{1}{\not p - m} \Delta m \frac{1}{\not p - m} \not a$$

It is the value of this matrix compared to that of the unperturbed matrix which gives the desired correction term (see Fig. 29-2).



(a) Without perturbation

(b) With perturbation of self-action and mass correction

FIG. 29-2

Now, for two noncommuting (or commuting) operators A and B, the following expansion is true:

$$\frac{1}{A+B} = \frac{1}{A} - \frac{1}{A} B \frac{1}{A} + \frac{1}{A} B \frac{1}{A} B \frac{1}{A} + \cdots$$

Then one can write

$$\frac{1}{\not p - \not k - m} = \frac{1}{\not p_0 + \epsilon \not p_0 - \not k - m} \approx \frac{1}{\not p_0 - \not k - m} - \frac{1}{\not p_0 - \not k - m}$$

$$\times \epsilon \not p_0 \frac{1}{\not p_0 - \not k - m} + \cdots$$

and the foregoing matrix becomes

$$i4\pi e^2 \int \not{b} \frac{\not{p}+m}{2m^2\epsilon} \gamma_\mu \frac{1}{\not{p}_0 - \not{k} - m} \gamma_\mu \frac{\not{p}+m}{2m^2\epsilon} \not{a} \frac{d^4k}{k^2-\lambda^2_{min}} \frac{-\lambda^2}{k^2-\lambda^2}$$

$$-i4\pi e^2 \int \not{b} \frac{\not{p}+m}{2m^2\epsilon} \gamma_\mu \frac{1}{\not{p}_0 - \not{k} - m} p_0 \frac{1}{\not{p}_0 - \not{k} - m} \gamma_\mu \frac{d^4k}{k^2-\lambda^2_{min}}$$

$$\times \frac{\not{p}+m}{2m^2} \not{a} \left(\frac{-\lambda^2}{k^2-\lambda^2}\right) - \not{b} \frac{\not{p}+m}{2m^2\epsilon} \Delta m \frac{\not{p}+m}{2m^2\epsilon} \not{a}$$

The first and last terms are identical, up to terms of order ϵ, hence may be canceled.

The integral in the second term has already been done essentially in computing diagram I (Fig. 28-2), except here $\not{p}_0$ replaces $\not{a}$, $\not{p}_1$, and $\not{p}_2$, that $\not{a} = \not{p}_2 - \not{p}_1 = 0$ in this case and gives the result

$$-\frac{e^2}{2\pi} r \not{b} \frac{\not{p}+m}{2m^2\epsilon} \not{p}_0 \frac{\not{p}+m}{2m^2} \not{a}$$

To this order in ϵ the $\not{p}$'s in the numerator may be replaced by $\not{p}_0$'s.

It is also noted that since $\not{p}_0 u = mu$,

$$(\not{p}_0 + m)\not{p}_0(\not{p}+m) \equiv 2m^2(\not{p}+m)$$

so that the foregoing result may be written

$$-\frac{e^2}{2\pi} r \not{b} \frac{\not{p}+m}{2m^2\epsilon} \not{a}$$

This is just - $(e^2/2\pi)\mathbf{r}$ times the matrix for no perturbation.

Hence the *correction term* due to diagrams II, III, II', and III' is obtained simply by replacing the scattering potential $\not{A}$ by $-(e^2/2\pi)\mathbf{r}\not{A}$, as was stated earlier.

It should be noted that the difficulty in obtaining the proper subtraction of the self-action and mass corrections just clarified does not represent a "divergence" problem of quantum electrodynamics.

It is a typical problem which could as well arise in nonrelativistic quantum mechanics if, for example, one chose some nonzero value as a reference of potential, that is, regarded a free electron as moving in a uniform nonzero potential.

It may be easily verified that this would give rise to an "energy correction" for the free electron analogous to the mass correction involved here.

Then in computing the amplitude for a scattering process where one used a "theoretical energy" and subtracted the effect of the "energy correction, " the difference of infinite terms would appear if one used free-electron wave functions. In this simple case the infinite term would, indeed, cancel upon proper subtraction but in principle the problem is the same as the present one.

Finally, the complete correction term arising from self-action and mass correction is

$$\frac{e^2}{2\pi}\left[2\left(\ln\frac{m}{\lambda_{min}}-1\right)\left(1-\frac{2\theta}{\tan 2\theta}\right)+\theta\tan\theta+\frac{4}{\tan 2\theta}\right.$$

$$\left.\times\int_0^\theta \alpha\tan\alpha\,d\alpha\right]\not{A}+\frac{e^2}{8\pi m}(\not{q}\not{A}-\not{A}\not{q})\frac{2\theta}{\sin 2\theta}$$

RESOLUTION OF THE FICTITIOUS "INFRARED CATASTROPHE"

From the correction term just determined, it is seen that, to order e^2 the cross section for scattering of an electron with the emission of *no photons* is

$$\sigma = \sigma_0 \left\{ 1 - \frac{e^2}{2\pi} \left[2\left(\ln \frac{m}{\lambda_{min}} - 1 \right) \left(1 - \frac{2\theta}{\tan 2\theta} \right) + \begin{pmatrix} \text{terms not} \\ \text{dependent} \\ \text{on } \lambda_{min} \end{pmatrix} \right] \right\}$$

where σ_0 is the cross section for the potential $\mathcal{A}$ only.

This cross section diverges logarithmically as $\lambda_{min} \rightarrow 0$, and it is this divergence which was formerly referred to as the "infrared catastrophe."

This result, however, arises from the physical fact that it is impossible to scatter an electron with the emission of no photons.

When the electron is scattered, the electromagnetic field must change from that of a charge moving with momentum p_1 to that for momentum p_2.

This change of the field is necessarily accompanied by radiation.

In the theory of bremsstrahlung, it was shown that the cross section for emission of one low-energy photon is

$$\sigma = \sigma_0 \frac{e^2}{\pi} \frac{d\Omega_\omega}{4\pi} \left(\frac{\omega p_1 \cdot e}{p_1 \cdot q} - \frac{\omega p_2 \cdot e}{p_2 \cdot q} \right)^2 \frac{d\omega}{\omega}$$

One can show that the integral over all directions and the sum over polarizations of the foregoing cross section is

$$\sigma = \sigma_0 (2e^2/\pi)[1 - (2\theta/\tan 2\theta)]\ d\omega/\omega$$

where $\sin^2 \theta = -(\not{p}_2 - \not{p}_1)^2/4m^2$.

Thus the probability of *emitting any* photon between k = 0 and k = K_m is

$$\sigma_0 \frac{2e^2}{\pi}\left(1 - \frac{2\theta}{\tan 2\theta}\right)\int_0^{K_m} \frac{d\omega}{\omega} = \sigma_0 \frac{2e^2}{\pi}\left(1 - \frac{2\theta}{\tan 2\theta}\right) \ln \frac{K_m}{\lambda_{min}}$$

which diverges logarithmically.

Therefore, the dilemma of the diverging scattering cross section actually arises from asking an improper question: What is the chance of scattering with the emission of no photons?

Instead, one should ask: What is the chance of scattering with the emission of no photon of energy greater than K_m?

For there will always be some very soft photons emitted.

the

Then, effectively, what is sought in answer to the last question is the chance of scattering and emitting no photon, the chance of emitting one photon of energy below K_m, and the chance of two and more photons below K_m (but these terms are of order e^4 and higher and hence are neglected).

Each of these terms is infinite, actually, but is kept finite temporarily by the artifice of the λ_{min}.

Their sum, however, does not diverge, as may be seen by gathering the previous results and by writing

$$\begin{aligned}
&\text{Chance of scattering and emitting no photon of energy} > K_m \\
&= \sigma_0 \left\{ 1 - \frac{e^2}{\pi} \left[2 \left(\ln \frac{m}{\lambda_{min}} - 1 \right) \left(1 - \frac{2\theta}{\tan 2\theta} \right) + (\text{terms inde-} \right. \right. \\
&\qquad \left. \left. \text{pendent of } \lambda_{min}) \right] \right\} + \sigma_0 \frac{2e^2}{\pi} \left(1 - \frac{2\theta}{\tan 2\theta} \right) \ln \frac{K_m}{\lambda_{min}} + (\text{terms} \\
&\qquad \text{of order } e^4) \\
&= \sigma_0 \left[\left(1 - \frac{e^2}{\pi} 2 \ln \frac{m}{K_m} \right) \left(1 - \frac{2\theta}{\tan 2\theta} \right) \right] + \begin{pmatrix} \text{terms independent} \\ \text{of } \lambda_{min} \text{ and of} \\ \text{order } e^4 \end{pmatrix}
\end{aligned}$$

This does not depend on λ_{min} and hence resolves the "infrared catastrophe."

It has been shown by Bloch and Nordsieck that the same idea applies to all others.

It is interesting that the largest term in the quantum-electrodynamic corrections to the scattering cross section, namely,

$$-(2e^2/\pi)\,[1 - (2\theta/\tan 2\theta)]\,\ln\,(m/K_m)$$

may be obtained from classical electrodynamics, since such long wavelengths are involved.

The other terms have small effects.

To date, the scattering experiments have been accurate enough to verify the existence of the large term but not accurate enough to verify the exact contributions of the smaller terms.

Hence they do not provide a nontrivial test of quantum electrodynamics.

These same considerations apply in any process involving the deflection of free electrons.

The best way to handle the problem is to calculate everything in terms of the λ_{min} and then to ask only questions which can have sensible answer as verified by the eventual elimination of the λ_{min}.

Part 30 ANOTHER APPROACH TO THE INFRARED DIFFICULTY

Instead of introducing an artificial mass, assume no weak photons contribute.

Thus we must subtract from the previous results the contributions of all photons with momentum magnitude less than some number $k_0 >> a$.

The previous result is

$$\not{a}\{1 + (e^2/2\pi)[2 \ln (m/\lambda_{min} - 1)(1 - 2\theta/\tan 2\theta)] + \theta \tan \theta$$
$$+ (4/\tan 2\theta) \int_0^\theta y \tan y \, dy]\} \tag{30-1}$$

The term to be subtracted is

$$(e^2/2\pi) \int_0^{k_0} \gamma_\mu (p_2 - \not{k} + m)(k^2 - 2p_2 \cdot k_2)^{-1} \not{a}(\not{p}_1 - \not{k} + m)$$
$$\times (k^2 - 2p_1 \cdot k_1)^{-1} \gamma_\mu \, d^4 k/(k^2 - \lambda^2_{min}) \tag{30-2}$$

We assume $k_0 << p_1$ or p_2, and neglect both $\not{k}$ and the first two k^2 in this integral.

Then using $\not{p}_1 \gamma_\mu = 2p_\mu - \gamma_\mu \not{p}_1$, the integral is approximately

$$x = -\frac{e^2}{2\pi} \frac{\not{a}}{2} \int \left[\frac{p_{2\mu}}{p_2 \cdot k} - \frac{p_{1\mu}}{p_1 \cdot k}\right]^2 \frac{d^4 k}{k^2 - \lambda^2_{min}} \tag{30-3}$$

Then

$$x = e^2/2\pi\{[1 - (2\theta/\tan 2\,\theta)]\,[2 \ln (2k_0/\lambda_{min} - 1)] + [4\theta/\tan 2\theta]$$
$$\times [(1/2\theta) \int_0^{2\theta} (y/\tan y)\, dy - 1]\} \tag{30-4}$$

This is the term to be subtracted from expression (30-1)

Using $\sin^2\theta = q^2/4m^2$, for small q, Eq. (30-4) become

$$x = (e^2/2\pi)\,(2q^2/3m^2)[\ln\,(2k_0/\lambda_{min}) - (5/6)]$$

Subtracting this from Eq. (30-1), also with q small, gives

$$\not{a}\,\{1 + (e^2/4\pi)(4q^2/3m^2)[\ln\,M/\lambda_{min}) - (3/8) - \ln\,(2k_0/\lambda_{min}) + (5/6)]\} \qquad (30\text{-}5)$$

The last term is $[\ln\,(M/2k_0) + (11/24)]$.

EFFECT ON AN ATOMIC ELECTRON

Consider the hydrogen atom with a potential $V = e^2/r$ and a wave function $\phi_0(R)\exp(-iE_0t) = \phi_0(x_\mu)$.

Take the wave function to be normalized in the conventional manner.

The effect of the self-energy of the electron is to shift the energy level by an amount

$$\Delta E = e^2 \int \tilde{\phi}_0(x_2,t_2)\,\gamma_\mu\,K_+{}^V(2,1)\,\gamma_\mu\,\delta_+(s_{1,2}{}^2)\,\phi_0(x_1,t_1)\,d^3x_1\;d^3x_2\;dt_2 - \Delta m \int \tilde{\phi}(x,t)\,\phi(x,t)\,d^3x \qquad (30\text{-}6)$$

The first integral is written down from Fig. 30-1.

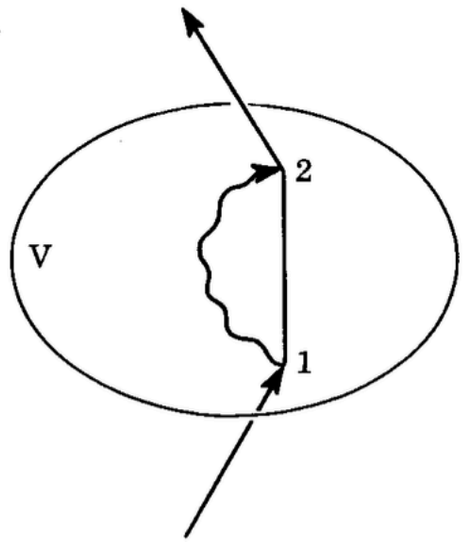


FIG. 30-1

The second is the free-particle effect as noted in previous parts.

The kernel K_+^V is not well enough determined to make exact calculation of this integral possible.

An approximate calculation can be made with the form

$$K_+^V(2,1) = \sum_{+n} \exp[-iE_0(t_2 - t_1)]\ \tilde{\phi}_n(x_2)\phi_n(x_1) \qquad t_2 > t_1$$

$$-\ \text{similar sum over negative energies for}\ \ t_2 < t_1$$

The photon propagation kernel can be expanded as

$$\delta_+(s_{1,2}^2) = 4\pi \int \exp[-ik(t_2 - t_1) + ik(x_2 - x_1)]\ d^3k/2k(2\pi)^{-3}$$

$$t_2 > t_1$$

$$= 4\pi \int \exp[+ik(t_2 - t_1) + ik(x_2 - x_1)]\ d^3k/2k(2\pi)^{-3}$$

$$t_2 < t_1$$

Using these expressions, Eq. (30-6) becomes

$$\Delta E = \sum_{+n} \int [\alpha_\mu \exp(-i\mathbf{K}\cdot\mathbf{R})]_{0n} (E_n + K - E_0)^{-1} [\alpha_\mu \exp(i\mathbf{K}\cdot\mathbf{R})]_{n0} \times d^3k/4\pi k - \sum_{-n} \int [\alpha_\mu \exp(-i\mathbf{K}\cdot\mathbf{R})]_{0n} (|E_n| + \omega + E_0)^{-1} \times [\alpha_\mu \exp(i\mathbf{K}\cdot\mathbf{R})]_{n0} d^3k/4\pi k - (\Delta m \text{ term}) \quad (30\text{-}7)$$

This form implies the use of ϕ^* instead of $\tilde{\phi}$ and $\alpha_4 = 1$, $\alpha_{1,2,3} = \alpha$.

Another approach to the motion of an electron in a hydrogen atom is the following.

Consider the electron as a free particle intermittently scattered by the Coulomb potential.

The scatterings cause a phase shift in the wave function of the order of (Rydberg/$\hbar$).

Thus the period between scatterings is of the order T = $\hbar$/Rydberg.

Take the lower limit k_0 of the momentum of the "self-action" photons as very large compared to the Rydberg.

Then it is very probable that an emitted photon will be reabsorbed before two interactions between the electron and the potential have taken place; it is very improbable for two or more scatterings to take place between emission and absorption (see Fig. 30-2).

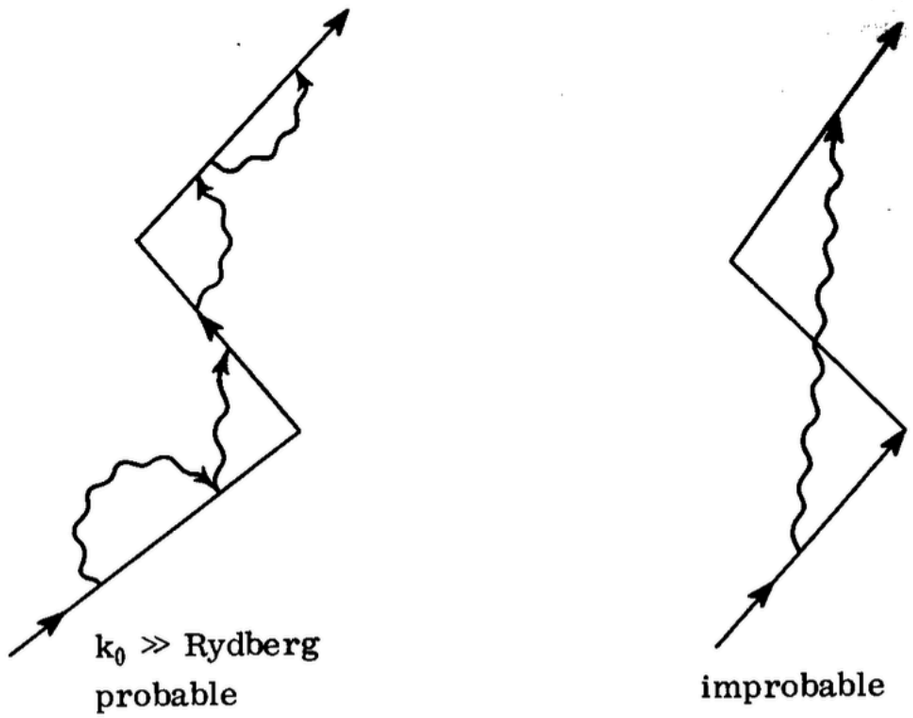


FIG. 30-2

Then the correction to the potential is that computed in Eg. (30-5) for small q (plus anomalous moment correction).

This is

$$(e^2/4\pi)(4q^2/3m^2)(\ln m/2k_0 + 11/24)\, \not{V}$$

in momentum space.

To transform to ordinary space, use

$$q^2 \not{V} = (q_4^2 - Q^2)\, \not{V} \rightarrow (\partial^2/\partial t^2 - \nabla^2)\, V$$

Thus the correction is

$$-(e^2/3\pi m^2)(\log M/2k_0 + 11/24)\, \nabla^2 V \qquad (30\text{-}7')$$

This correction is of greatest importance for the s state, since with a Coulomb potential $\nabla^2 V = 4\pi Ze^2\delta(\mathbf{R})$, and only in the s states is $\phi(\mathbf{R})$ different from 0 at $\mathbf{R} = 0$.

The choice of k_0 is determined by the inequalities $m \gg k_0 \gg$ Rydberg.

A satisfactory value is $k_0 = 137$ Ryd.

With such a k_0, the effect of photons of $k < k_0$ must be included.

This will be done by separating the effect into the sum of three contributing effects.

It will be seen that two of these effects are independent of the potential V and thus are canceled by similar terms in the Δm correction for a free particle.

Thus for only one situation must the effect be computed.

In all cases, since k is small, the nonrelativistic approximation to expression (30-7) may be used.

(1) The contribution of negative energy states: Neglecting k with respect to m gives

$$(|E_n| + k + E_0) \approx 2m$$

The matrix element for α_4 is very small, and only the elements for $\boldsymbol{\alpha}$ need be considered. Then the sum over negative states is

$$\sum_{-n} \int [(\boldsymbol{\alpha}_{0n}) \cdot (\boldsymbol{\alpha}_{n0})/2m] k^2 \, dk/k$$

If this sum is continued for +n, a negligible term of order v^2/c^2 is added. Thus the sum is approximately

$$- \sum_{\text{all states}} \int [(\alpha_{0n}) \cdot (\alpha_{n0})/2m] \; k^2 \, dk/k = (\alpha \cdot \alpha)_{00} \, k^2 \, dk/2mk$$

$$= 3k_0^{\,2}/4m$$

This is independent of V, and thus is canceled by a similar quantity in the Δm term.

(2) Longitudinal positive energy states ($\alpha_\mu \rightarrow \boldsymbol{\alpha} \cdot \mathbf{k}/k$): As an exercise the reader may show

$$\boldsymbol{\alpha} \cdot \mathbf{k} \exp(i\mathbf{k} \cdot \mathbf{R}) = H \exp(i\mathbf{k} \cdot \mathbf{R}) - \exp(i\mathbf{k} \cdot \mathbf{R}) H$$

Then

$$[(\boldsymbol{\alpha} \cdot \mathbf{k}/k) \exp(i\mathbf{k} \cdot \mathbf{R})]_{n0} = (E_n - E_0)/k [\exp(i\mathbf{k} \cdot \mathbf{R})]_{n0}$$

and the contribution of these terms summed over positive energy states gives

$$\int [1 - (E_n - E_0)^2/k^2] \exp(i\mathbf{k}\cdot\mathbf{R})_{0n} \exp(-i\mathbf{k}\cdot\mathbf{R})_{n0} (E_n + k - E_0)^{-1} \, d^3k/4\pi k$$

$$= \int (E_n - E_0 + k) \exp(i\mathbf{k}\cdot\mathbf{R})_{0n} \exp(-i\mathbf{k}\cdot\mathbf{R})_{n0} \, d^3k/4\pi k^3$$

$$= \int [H \exp(i\mathbf{k}\cdot\mathbf{R}) - \exp(i\mathbf{k}\cdot\mathbf{R})H]_{0n} [\exp -(i\mathbf{k}\cdot\mathbf{R})]_{n0} \, d^3k/4\pi k^3$$

Writing $H = p^2/2m$ (V commutes with the exponent), this becomes

$$\int [(p + k)^2/2m - p^2/2m + k] \, d^3k/k^3$$

This term is independent of V, and thus is also canceled by the Δm correction.

(3) Transverse positive energy states: Since k_0 is large compared to the size of the atom, the dipole approximation can be used.† The general term in the sum of Eq. (30-7) becomes

$$\int (\alpha_{tr})_{0n} (\alpha_{tr})_{n0} (E_n + k - E_0)^{-1} \, d^3k/k \qquad (30\text{-}8)$$

Writing

$$(E_n + k - E_0)^{-1} = 1/k - (E_n - E_0)/(E_n + k - E_0)k$$

the term in $1/k$ can be split off from the rest of the integral as a quantity independent of V and thus canceled by the Δm correction. Further, by averaging over directions,

$$(\alpha_{tr})_{0n} (\alpha_{tr})_{n0} = 2/3(\boldsymbol{\alpha})_{0n} \cdot (\boldsymbol{\alpha})_{n0} = (2/3m^2)(\mathbf{p})_{0n} \cdot (\mathbf{p})_{n0}$$

in the nonrelativistic approximation. Thus the integral of Eq. (30-8) is

$$(2/3m^2)(\mathbf{p})_{0n} \cdot (\mathbf{p})_{n0} (E_n - E_0) \log (k_0 + E_n - E_0)/(E_n - E_0)$$

Using the relation

$$\mathbf{p}_{n0}(E_n - E_0) = (\mathbf{p}H - H\mathbf{p})_{n0} = (\nabla V)_{n0}$$

and the fact that $k_0 \gg E_n - E_0$, one part of the sum over transverse positive energy states is

$$\ln\ k_0 \sum_n \mathbf{p}_{0n} \cdot (\nabla V)_{n0} = 1/2\ \ln\ k_0 (\nabla^2 V)_{00}$$

This cancels with the $\ln\ k_0$ of Eq. (30-7′), leaving the final correction as

$$(2e^2/3\pi m^2) \sum_{+n} \mathbf{p}_{n0} \cdot \mathbf{p}_{0n} (E_n - E_0) \{ \log [M/2(E_n - E_0)] + (11/24) \}$$

+ anomalous moment correction

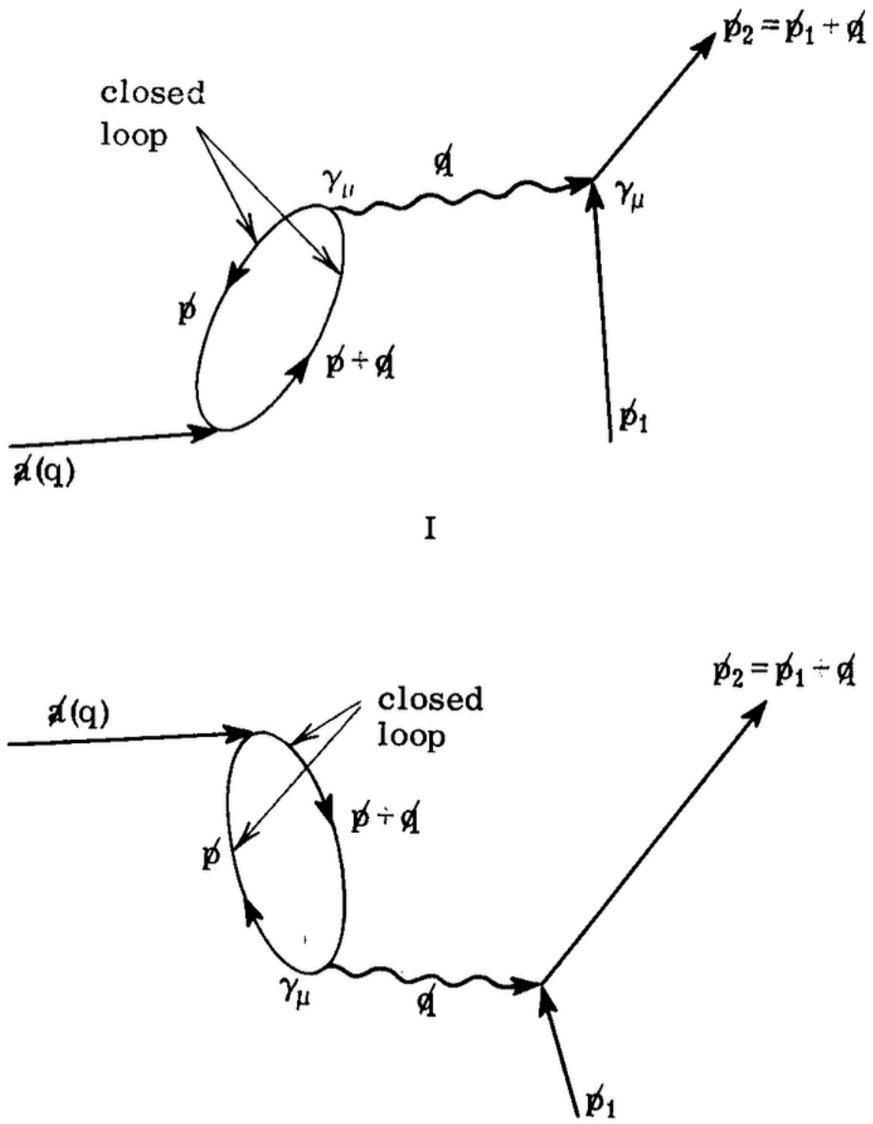
This sum has been carried out numerically to be compared with the observed Lamb shift.

Part 31 CLOSED-LOOP PROCESSES, VACUUM POLARIZATION

Another process which is still of first order in e^2 has not been considered in the scattering by a potential.

Instead of the potential scattering the particle directly, it can do so by first creating a pair which subsequently annihilates, creating a photon which does the scattering.

Diagram I (Fig 31-1) applies to this process; diagram II applies to a similar process, with the order in time changed slightly.



I

II

FIG. 31-1

The amplitude for these processes is

$$i4\pi e^2 \sum_{\substack{\text{spin states} \\ \text{of u}}} (\tilde{u}_2 \gamma_\mu u_1) \frac{1}{q^2} \int \left(\tilde{u} \frac{1}{\not p - m} \gamma_\mu \frac{1}{\not p + \not q - m} \not a u \right) (2\pi)^{-4} \, d^4 p \tag{31-1}$$

where u is the spinor part of the closed-loop wave function.

The first parenthesis is the amplitude for the electron to be scattered by the photon; $1/q^2$ is the photon propagation factor; and the second parenthesis is the amplitude for the closed-loop process which produces the photon.

The expression is integrated over p because the amplitude for a positron of any momenta is desired.

In the sum over four spin states of u, two states take care of the processes of diagram I and two states take care of the processes of diagram II.

No projection operators are required, so the method of spurs may be used directly to give

$$i4\pi e^2 \,(\tilde{u}_2 \gamma_\mu u_1)\, \frac{1}{q^2} \int \mathrm{Sp}\left[\frac{1}{\not p - m}\,\gamma_\mu\,\frac{1}{\not p + \not q - m}\,\not a\right] (2\pi)^{-4}\, d^4 p \tag{31-2}$$

a form which contains both I and II (so as usual it is not necessary to make separate diagrams for processes whose only difference is the order in time).

This integral also diverges, but a photon convergence factor, as used in the previous parts, is of no value because now the integral is over p, the momentum of the positron in the intermediate step.

The method which has been used to circumvent the divergence difficulty is to subtract from this integral, a similar integral with m replaced by M.

M is taken to be much larger than m, and this results in a type of cutoff in the integral over p.

When this is done, the amplitude is found to be

$$(\tilde{u}_2 \gamma_\mu u_1) a_\mu (e^2/\pi) [-(1/3) \ln (M/m)^2 - (1 - \theta/\tan \theta) \times (4m^2 + 2q^2)/3q^2 + 1/9] \quad (31\text{-}3)$$

where $q^2 = 4m^2 \sin^2 \theta$, which, for small q, becomes

$$(\tilde{u}_2 \gamma_\mu u_1) a_\mu (e^2/\pi) [-(1/3) \ln (M/m)^2 + 2q^2/15] \quad (31\text{-}4)$$

Notice that $(\tilde{u}_2 \gamma_\mu u_1) = (\tilde{u}_2 \not{a} u_1)$, so that, considering only the divergent part of the correction, the effective potential is

$$\not{a} \{1 + (e^2/\pi)[-(1/3) \ln (M/m)^2]\} \quad (31\text{-}5)$$

The 1 comes from the theory without radiative corrections, while the e^2 term is the correction due to processes of the type just described.

Thus the correction can be interpreted as a small reduction in the effect of all potentials, and one can introduce an experimental charge e_{exp} and a theoretical charge e_{th} related by

$$e_{exp} = e_{th} + \Delta e \quad (31\text{-}6)$$

where $\Delta(e^2) = -(e^2/3) \ln (M/m)^2$, in a manner analogous to the mass correction described in Part 28.

This is referred to as "charge renormalzation."

The other term

$$(2/15)(e^2/\pi)q^2\not{A}$$

is more interesting, since it represents a perturbation $2e^2/15\pi(\nabla^2 V)$.

This correction is responsible for 27 Mc in the Lamb shift and the $\{\ln [m/2(E_n - E_0)]+ (11/24)\}$ term in (30-7') is replaced by $\{\ln [m/2(E_n - E_0)] + (11/24) - (1/5)\}$.

The 1/5 term is due to the "polarization of the vacuum."

SCATTERING OF LIGHT BY A POTENTIAL

One possible process for the scattering of light, and an indistinguishable alternative, is indicated by the diagrams in Fig. 31-2.

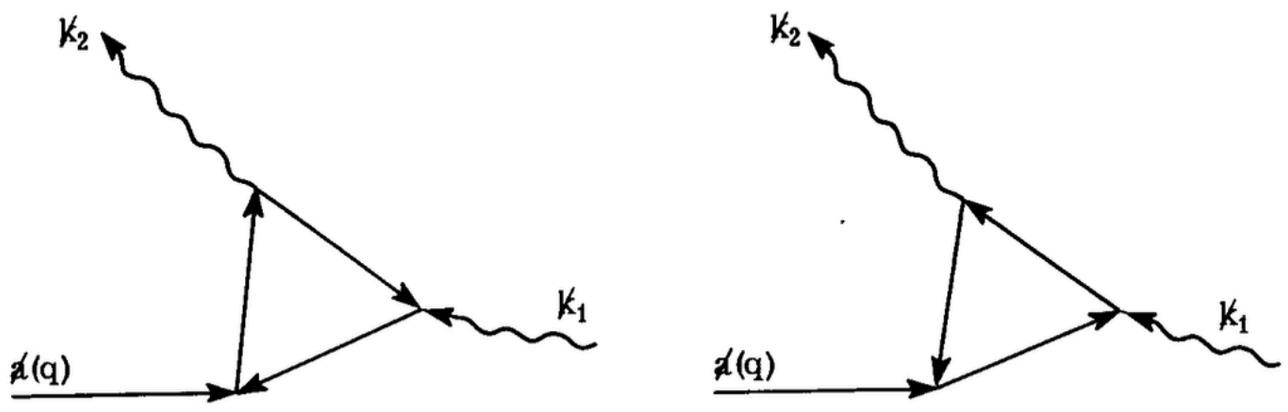


FIG. 31-2

The second diagram differs from the first only in the direction of the arrows of the electron lines.

Reversing such a direction is equivalent to changing an electron to a positron.

Thus the coupling with each potential would change sign.

Since there are three such couplings, the amplitude for the second process is the negative of that for the first. Since the amplitudes add, the net amplitude is zero.

In general, any closed-loop process of this type involving an odd number couplings to a potential (including photon), has zero net amplitude

Note, if you set up the integrals for each of the two diagrams in Fig. 31-2, you can show that they are equal and opposite in sign.

However, the higher-order processes shown in Fig. 31-3 can take place.

The amplitude for the process is

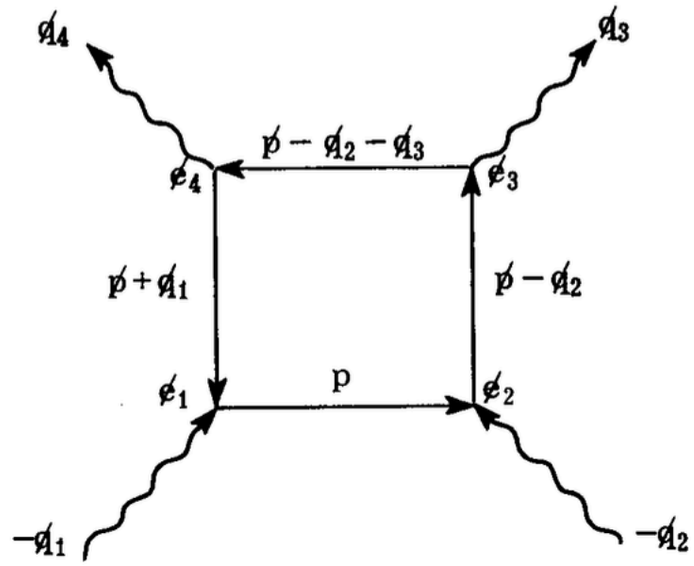
$$-(4\pi e^2)^2 \int \mathrm{Sp}\,[\not\epsilon_1(\not p - m)^{-1}\, \not\epsilon_2(\not p_1 - \not q_2 - m)^{-1}\, \not\epsilon_3(\not p - \not q_2 - \not q_3 - m)^{-1}$$

$$\times \not\epsilon_4(\not p + \not q_1 - m)^{-1}\,]\,(2\pi)^{-4}\, d^4k$$

plus five similar terms resulting from permuting the order of photons as in the figure.

This integral appears to diverge logarithmically. But when all six alternatives are taken into account, the sum leaves no divergent term.

More complicated closed-loop processes are convergent.



alternatives

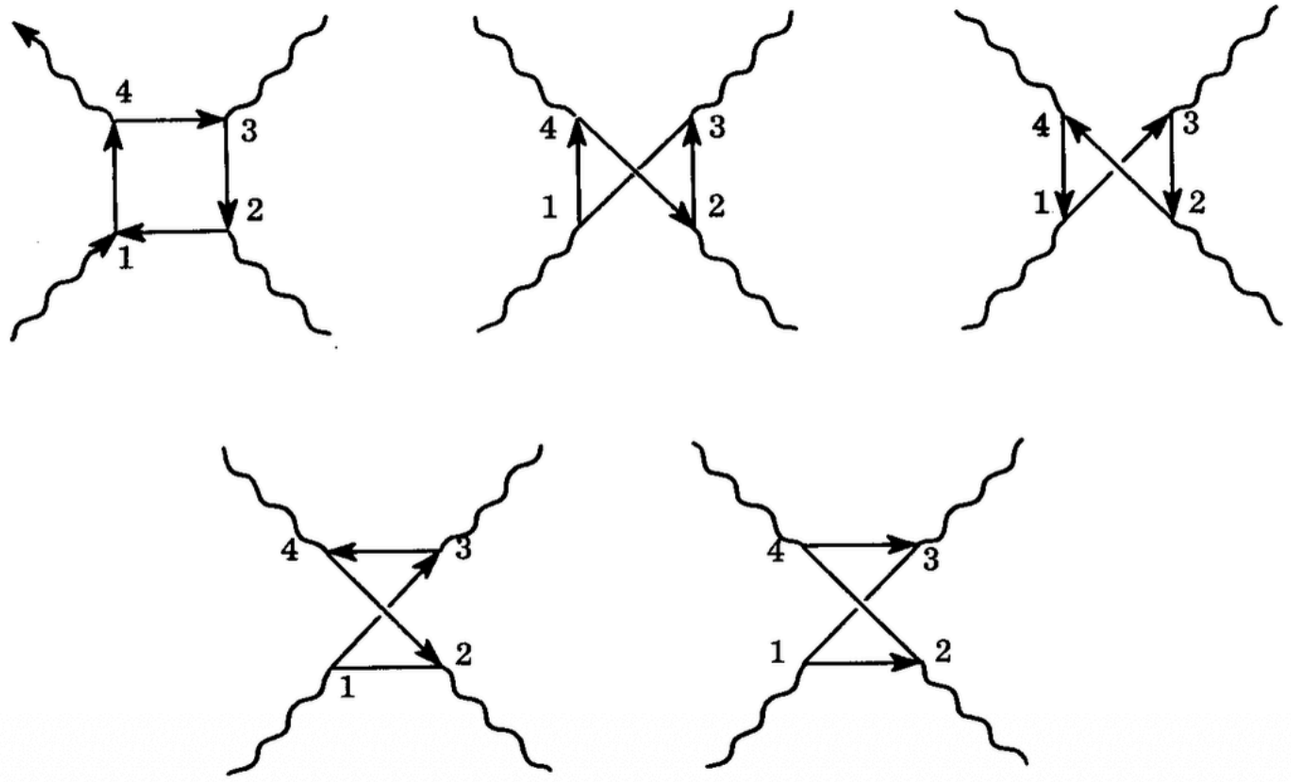


FIG. 31-3

That should have satisfied anyone’s desire to see detailed calculations.