Chapter Three

Quantum Electrodynamics

A comprehensive treatment of quantum electrodynamics (QED) cannot be given here in a single chapter. In any case, that would be beyond the scope of this book, and, as with the classical theory, the general subject has already been covered in a number of very fine textbooks. The full theory is, in fact, quite complicated and requires a lengthy formulation if it is to be laid out in a systematic way. Actually, the final result of this theory is a fairly simple set of prescriptions (Feynman rules) for the calculation of rates of electromagnetic processes. Instead of trying to "derive" the rules in a rigorous and pedestrian manner, we shall try to take an intuitive approach—like Feynman did—to shortcut the path leading to the rules themselves. In an attempt to make the reasoning seem more natural, we start first with a restricted limiting form of the theory. This limiting form is the non-relativistic theory restricted to the case where the photons are radiation-field particles. That is, the photons are considered to exist in given observable states described by their wave vector (k) and polarization unit vector (ε). The photons are considered to couple to non-relativistic charges and magnetic moments, and the rates (cross sections) for the various processes are computed by elementary perturbation theory. To calculate these rates we need the perturbation Hamiltonian terms (H') associated with these couplings, and we can obtain these by elementary, but rigorous, methods. Having formulated this restricted corner of the subject, we then try to make the jump to the relativistic theory. In later chapters, when the theory is applied to treat various processes, the same approach will be taken. That is, the non-relativistic limit will be treated first, to be followed by the more general case where the particles can have relativistic energies.

The non-relativistic theory is, to be sure, an important subject in itself, with many useful applications. Some processes are sufficiently complicated that analytic results cannot be derived in general covariant formulations, although exact formulas can be derived in the non-relativistic theory with its restricted energy domain. This chapter treats only the developments of the theory for its use in later applications, and it differs from the way the subject is presented in the standard textbooks. Modern books generally formulate the covariant theory from the very beginning, a result of the great success of this general and preferred approach. The whole subject has an interesting history, and we begin with a short outline of it.

3.1 BRIEF HISTORICAL SKETCH

A lot of people have the impression that QED started in 1948 with the work of Feynman, Schwinger, and others. The real beginning came in 1905 with the introduction of the concept of the photon. This concept was not readily accepted by twentieth-century physicists; only by about, say, 1917 did the particle nature of electromagnetic radiation begin to be considered seriously.[1] In 1922, the particle character of the photon was seen clearly in the observation of the Compton effect which demonstrated the existence of momentum and energy of individual photons. With the subsequent developments in quantum mechanics in 1925 and the description of processes in terms of a general time-dependent Schrödinger equation

$$H\psi = i\hbar\,\partial\psi/\partial t, \tag{3.1}$$

the formulation of a quantum theory of electromagnetism followed, principally in the work of Dirac, Heisenberg, Pauli, and Fermi.

The essence of QED—a quantum theory of the electromagnetic field—is the introduction of the photon and the description of the fields in terms of a collection of these particles in various one-particle states. That is, the field is specified in terms of the values of the occupation numbers of the particle (photon) states. The classical Maxwell theory serves as a guide for QED, since in the limit of large occupation numbers the theory must approach that of the classical continuum. In a quantum-mechanical formulation describing many charged particles and photons, the wave function ψ in Equation (3.1) would contain information on the charges and photons, and the Hamiltonian would consist of parts associated with each plus an interaction part:

$$H = H_{\text{ch}} + H_{\text{ph}} + H_{\text{int}}, \tag{3.2}$$

just as in classical theory. Dirac gave the first such treatment of charged particles and photons in 1927 for the special case of radiation-field photons. As in classical electrodynamics, the quantum mechanics of the radiation field is simpler than that for general electromagnetic fields.

General QED, formulated to describe electromagnetic effects not confined to radiation fields, presents more formidable problems, especially in connection with the requirements of gauge invariance. The first formulations of the general theory was by Heisenberg and Pauli in 1929 and Fermi in 1930. The theory was then successfully applied in the early 1930s to a number of problems involving the interaction of charged particles and photons. At the same time, however, it was found that there were difficulties with the theory when applied to certain problems. In particular, it was found that the theory seemed incapable of calculating effects to higher order in the electromagnetic coupling constant α. That is, quantities like cross sections, calculated by perturbation theory in a power series in α, came out in satisfactory form in lowest order, but yielded infinities in the higher-order corrections. The divergence problems, which we discuss briefly later in this chapter, show up in the evaluation of these "radiative corrections" and, in particular, in the

---

[1]Even as late as 1917, only Einstein was actively advocating acceptance of the idea of the photon (or "quantum") [see A. Pais, *Rev. Mod. Phys.* **51**, 861 (1979)].

attempt to calculate the electromagnetic self-energy of an electron. As first shown by Weisskopf, the divergence in the self-energy problem is not as serious as in classical electrodynamics, when treated by relativistic QED; the divergence, instead of being linear, is now of a logarithmic type.

Quantum electrodynamics remained in this state until the late 1940s when there began a new phase in theoretical and experimental developments on the subject. The experiments were partly a result of advances in microwave techniques from research carried out during World War II. The first important measurement was the (1947) determination of the *Lamb shift*, a displacement of the $2s_{1/2}$ level with respect to the $2p_{1/2}$ level in atomic hydrogen. The small level shift, suspected from earlier spectroscopic measurements, provided a challenge for theory, and an approximate calculation of the effect was carried out by Bethe soon after the experimental discovery. This was an important discovery, and the method used in the theoretical calculation of the effect was also of great significance, for the level shift corresponds to a "radiative correction," and the measurement showed that the effect is real and finite. Bethe's was the first calculation making use of "renormalization" techniques, thereby eliminating the necessity of dealing with the ("ultraviolet") divergences that occur in these higher-order calculations. The idea, first introduced into QED by Kramers[2] in 1937, removes the divergences by noting that the observed mass of the electron includes the electromagnetic self-energy. Then, when the higher-order radiative effects are expressed in terms of the observed mass, the divergent integrals subtract out.

In the years 1948–1950 there was a great deal of work on QED, especially on the theoretical side. Independently, Feynman and Schwinger reformulated the theory, providing a description that was relativistically covariant at every stage. Actually, it turned out that work on covariant QED had already begun in Japan in 1943 by Tomonaga; however, Tomonaga's papers, first published in Japanese, were unknown to almost everyone. Feynman's approach first seemed hard to comprehend, and Schwinger's papers were not easy to read. The work of Feynman was highly intuitive and, although no one else really knew why, with his formulation he was apparently able to perform calculations of a number of processes with great ease. The relationship between the work of Tomonaga, Feynman, and Schwinger with field theory was shown by Dyson. It is Dyson's formulation of covariant QED and covariant perturbation theory that is found in all subsequent textbooks on the subject.

The advantage of a fully covariant formulation of QED became evident, and the theory, with the help of renormalization procedures, was shown to be capable of performing calculations to very high accuracy. In particular, radiative corrections to atomic energy levels (Lamb shift) and to the magnetic moment of the electron were evaluated and found to be in excellent agreement with new very precise measurements. In its ability to calculate quantities that can be measured—to an accuracy of one part in 10^8 in atomic energy levels—the theory is highly successful. Perhaps it should be emphasized, nevertheless, that modern QED is fundamentally no different

---

[2]See, for example, the note on p. 453 in H. A. Kramers, *Quantum Mechanics*, New York: Dover Publ. Inc., 1964.

from the Dirac-Heisenberg-Pauli-Fermi theory described in 1930. The computational capabilities are much more advanced in the fully covariant theory, however, and this again demonstrates the advantages of a covariant formulation. For example, using the modern methods, the derivation of a formula for, say, some cross section might take a couple of hours, while months of work would be required using the older non-covariant procedures. For some calculations, even the methods of modern QED involve very lengthy mathematical manipulations requiring the evaluation of a great many integrals, etc. These higher-order calculations, which employ the techniques of renormalization, are beyond the capability of the older formulations of QED. Still, the theory is not in a totally satisfactory state in that the divergences have not been eliminated. For example, we cannot compute the electromagnetic self-energy of the electron; indeed, the theory still gives infinity for this quantity. If we can forget about difficulties like that, the theory can be applied with confidence to the calculation of virtually any observable electromagnetic phenomenon.

3.2 RELATIONSHIP WITH CLASSICAL ELECTRODYNAMICS

Purely classical electrodynamics (CED) is a continuum theory where the concept of a field particle (photon) is foreign. It is also a limiting case of the more general theory (QED), so that it is contained within the latter. This relationship can be exploited, as it is in the following section (3.3), to employ CED as a guide in formulating QED. There is, of course, a limit to what can be done along these lines, since the more general theory is greater in content and describes phenomena that are beyond the capability of the classical theory. Nevertheless, like all fundamental theories, QED is simple in its foundations, and it is not difficult to formulate it with the help of our knowledge of its limiting form.

The classical domain of a radiation field corresponds to the condition

$$\bar{n} \gg 1, \tag{3.3}$$

where $\bar{n}$ is the photon occupation number; this is essentially a specification of the applicability of a continuum-description theory of the photon field. For a particular process involving the interaction of these photons, the classical limit requires that the effects having to do with individual photons be negligible. That is, the particle (photon) characteristics must not play a role in this limit. It would seem that the condition that the photons are "soft" would be sufficient to satisfy this requirement. The soft-photon limit would confine the photon energy and momentum (that is, its kinematic properties) to be small compared with the energies and momenta (and their changes) of the charged particles involved in the particular process. For most processes this criterion would be sufficient for the validity of a classical treatment. However, for a particular process[3] it may be that the photon kinematic properties play a key role even in the soft-photon limit. In such cases, a semi-classical calculation is not completely valid if its only quantum-mechanical aspect is the introduction

---

[3]An example, treated later, would be electron-electron bremsstrahlung in the Born-approximation limit. This process must be given a quantum-mechanical treatment even for the case where the photons are soft.

of the photon concept. Some processes, such as photon-photon scattering, simply do not exist in a purely classical theory. This process, which can be thought of as involving virtual electron-positron pairs in an intermediate state, requires the fully relativistic QED for its treatment. However, immediately upon the introduction of quantum-mechanical ideas, it becomes fairly clear that such processes must be possible. In fact, even without a detailed quantum-mechanical calculation, we can have some idea of the order of magnitude of the associated cross section for this higher-order process. Moreover, for some processes, the relativistic effects are such that they can be included in a calculation in an approximate way by modifying the (large) argument of some logarithmic factor.

For a process that does have an essentially classical description, it may be that only certain kinematic properties of the photon need be limited for the domain of applicability. For example, it may be that only the photon momentum is important in kinematic considerations while the photon energy is not limited. Or it may be that a classical description is valid even though the photon momentum is comparable to that of the charged particles involved. Bremsstrahlung and Compton scattering in the non-relativistic limit are examples of, respectively, each of these cases.

Sometimes it is an accident that a certain classical formula is identical to a particular quantum-mechanical one. If this happens to be so, it is important to understand the reason for the coincidence. There is certainly one aspect of QED that is highly significant, and this has to do with the fundamental coupling constant associated with electromagnetic processes. The results (2.143) and (2.149), derived classically and obtained later in this chapter in a quantum-mechanical formulation, indicate how α ($= e^2/\hbar c \cong 1/137$) determines the probability of photon emission. The electromagnetic coupling is weak (but not *very* weak), so that the calculation of processes can be carried out through the use of perturbation theory. Also, in the comparison of classical and perturbation-theory formulas, because α is small, we can understand why identical results are obtained. In the calculation of, say, photon-emission probability, the classical formula really gives the probability of *one or more* photons being produced, while the perturbation-theory result is for the probability of production of *one* photon. Because multiple photon production is improbable,[4] the "one or more" of the classical formula essentially means one.

The degree to which quantum mechanics must be included in a problem involving the electromagnetic interaction depends on the details of the process. For example, in bremsstrahlung it may be that "photon emission" can be described classically while the rest of the process (scattering, say) is treated quantum mechanically. The experimental or observational conditions can also dictate how some phenomenon is handled theoretically. If individual photons are not being detected and quantum mechanics does not play a role in the process, then classical electrodynamics is an appropriate description. Quantum electrodynamics can also be employed in such a problem, but it simply makes no sense to take the more general approach. In fact, the exact quantum-mechanical treatment might be so complicated that the problem cannot be handled this way. The relationship between CED and QED will become clearer as we treat the various electromagnetic processes in later chapters.

---

[4]However, this is not so at ultrahigh energies [see Equation (3.191)].

3.3 NON-RELATIVISTIC FORMULATION

3.3.1 Introductory Remarks

By confining our treatment to describe only processes involving radiation-field photons, we can simplify the subject. Thus, our photons are always real[5] and their parameters (polarization and wave vector) are numbers that are fixed in the initial and/or final states. This relaxes, to a certain degree, the extent to which quantum mechanics must be employed in the description of the photon field and its interaction. However, it is inappropriate to refer to our treatment as "semi-classical," since all the necessary quantum mechanics is included, and the formulation—with, however, its limited applicability—is completely rigorous. We do include the particle aspects of the photon field, that is, the effects of the individual-particle momentum ($\hbar\boldsymbol{k}$) and energy ($\hbar ck = \hbar c|\boldsymbol{k}|$). Also, we give a complete quantum mechanical description of the charged particles that interact with the photons. The only restriction placed on the charge motion is that $\beta = v/c \ll 1$.

A wide variety of problems can be worked out with the help of the formulation outlined here. It can even be applied to problems involving pair production and annihilation, as long as the charged particles involved have non-relativistic kinetic energies. For some problems or applications, the non-covariant development is actually more convenient, even for the case where the charges are in relativistic motion. However, nowadays everyone learns QED in the modern fully covariant form. There are, as a result, few treatments of non-relativistic, non-covariant QED in modern textbooks. Perhaps, therefore, the simplified approach taken in this chapter will be helpful to people not inclined to delve into the details of the full theory.

3.3.2 Classical Interaction Hamiltonian

Consider first the classical motion of a charge in an electromagnetic field described in terms of a vector potential $\boldsymbol{A}(\boldsymbol{r}, t)$ and scalar potential $\Phi(\boldsymbol{r}, t)$. If there is present, in addition, a velocity-independent potential $V(\boldsymbol{r}, t)$, the total Lagrangian function would be (see Chapter 1)

$$L = \tfrac{1}{2}mv^2 + (q/c)\boldsymbol{A} \cdot \boldsymbol{v} - q\Phi - V. \tag{3.4}$$

As is well known, this Lagrangian yields the correct equation of motion including the Lorentz force. The corresponding Hamiltonian is

$$H = \frac{1}{2m}\left(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right)^2 + q\Phi + V. \tag{3.5}$$

It should be noted that

$$\boldsymbol{p} = \partial L/\partial \boldsymbol{v} = m\boldsymbol{v} + (q/c)\boldsymbol{A} \tag{3.6}$$

---

[5]That is, they are not "virtual." A virtual photon is one that is emitted and reabsorbed and exists only in intermediate states. If the photons in intermediate states are described by plane wave states, the amplitude and coupling derived here for radiation photons can still be used to describe these states, however.

is the canonical momentum and *not* the particle momentum; it is this quantity that is replaced by the operator $-i\hbar\nabla$ in going to a Schrödinger equation. The particular form of the Hamiltonian should also be noted, in particular, the first term in Equation (3.5); it is the same as that for a free particle with the replacement

$$\boldsymbol{p} \to \boldsymbol{p} - (q/c)\boldsymbol{A}. \tag{3.7}$$

The prescription (3.7) for the formation of a Hamiltonian function including electromagnetic interactions is sometimes referred to as a "Principle of Minimal Electromagnetic Coupling." The coupling is "minimal" in the same sense that it yields a particular combination of terms involving the potentials Φ and $\boldsymbol{A}$ such that the gauge invariance of the theory is satisfied.[6] In covariant form, in terms of the gradient operator, the relation (3.7) would become, for quantum-mechanical wave equations,

$$\partial_\mu \to \partial_\mu - (iq/\hbar c)A_\mu. \tag{3.8}$$

The Hamiltonian (3.5), when written as

$$H = p^2/2m + V + H_{\text{int}}, \tag{3.9}$$

yields an interaction part

$$H_{\text{int}} = -(q/2mc)(\boldsymbol{p} \cdot \boldsymbol{A} + \boldsymbol{A} \cdot \boldsymbol{p}) + (q^2/2mc^2)\boldsymbol{A} \cdot \boldsymbol{A} + q\Phi \tag{3.10}$$

associated with the coupling to the electromagnetic field. In this expression, since we want to make a transition to a quantum-mechanical formulation, the terms $\boldsymbol{p} \cdot \boldsymbol{A}$ and $\boldsymbol{A} \cdot \boldsymbol{p}$ are not equated. When $\boldsymbol{p}$ is replaced by the gradient operator and the term $\nabla \cdot \boldsymbol{A}$ operates on a function ψ to the right, the identity

$$\nabla \cdot \boldsymbol{A}\psi = (\nabla \cdot \boldsymbol{A})\psi + \boldsymbol{A} \cdot \nabla\psi \tag{3.11}$$

is employed. As we have seen in Chapter 2, for radiation fields a gauge can be chosen such that $\nabla \cdot \boldsymbol{A} = 0$ *and* $\Phi = 0$. In this case, for a system of charges interacting with an electromagnetic radiation field, the part of the Hamiltonian associated with this coupling is

$$H_{\text{int}} = \frac{i\hbar}{c} \sum_\alpha \frac{q_\alpha}{m_\alpha} \boldsymbol{A}(\boldsymbol{r}_\alpha, t) \cdot \nabla_\alpha + \frac{1}{2c^2} \sum_\alpha \frac{q_\alpha^2}{m_\alpha} \boldsymbol{A}(\boldsymbol{r}_\alpha, t) \cdot \boldsymbol{A}(\boldsymbol{r}_\alpha, t). \tag{3.12}$$

This expression is a generalization of Equation (3.10), employing the identity (3.11) and summing over particles (α), replacing the momenta by the operators $-i\hbar\nabla_\alpha$.

The formulation given above is adequate to calculate the rate at which the system of charges undergoes transitions as a result of the action of the perturbation (3.12). The state of the charge system is described in terms of a wave function $\psi(\boldsymbol{r}_\alpha, t)$ with the index α running over 1 to N (= number of charges). The charges can be in free-particle states or bound in atoms, molecules, or nuclei. However, the electromagnetic field, expressed in terms of the vector potential $\boldsymbol{A}$ in Equation (3.12), is treated as an "external" potential, that is, as a specified function of $\boldsymbol{r}$ and t. This corresponds to a classical description of the photon field, and it is important to understand the limitations of the treatment and why it is adequate for a

---

[6]See, for example, F. Rohrlich, *Classical Charged Particles*, Reading, MA: Addison-Wesley Publ. Co., 1965; W. Pauli, *General Principles of Quantum Mechanics*, Berlin: Springer-Verlag, 1980.

certain class of problems. The formulation works only for radiation-field photons, that is, for photons in free-particle (plane-wave) states. These photons are thus in fixed states, either generated prior to incidence on the charge system or as detected after interaction with the system. The simplified treatment cannot handle problems that require consideration of electromagentic perturbations corresponding to purely intermediate-state photons[7] ("photons" that are emitted and reabsorbed during the process). Although the photons must be either incoming or outgoing on the charge system, they can be created (or annihilated) as a result of the perturbation (3.12). The elementary theory given here is capable of calculating these processes without introducing the formalism of quantum field theory and annihilation and creation operators, etc.

Photons existing only in intermediate states do not have observable characteristics and, in fact, have a spectrum of kinematic properties. A full quantum-mechanical treatment would be required to evaluate processes involving such states, and this apparatus is provided by conventional quantum field theory. However, the formulation outlined here is adequate to treat problems in which the charged particles exist in (unobserved) intermediate states "in between" the action of perturbations. The elementary theory is also capable of treating problems involving pair production and annihilation as long as the charged-particle kinematic energies are non-relativistic. However, the non-relativistic theory cannot treat problems in which there are "virtual" charged pairs in intermediate states, since the characteristic energies of these pairs is such as to force a relativistic treatment. An example of a process of this type is photon-photon scattering—a phenomenon that does not exist in purely classical theory. As mentioned earlier, photons can scatter off one another because of the possibility of virtual electron-positron pairs in an intermediate state. The incoming and outgoing photons couple to these "particles" and this allows the process to take place. For center-of-mass photon energies such that $\varepsilon_1 = \varepsilon_2 \ll mc^2$, one might think that a non-relativistic theory would be capable of computing the cross section for the process. However, if pairs are produced in an intermediate state, they will have characteristic kinetic energies $\sim mc^2$ even for the scattering of low-energy photons.

Finally, on the subject of the interaction Hamiltonian, we should introduce an additional term associated with coupling of the electromagnetic field to a particle's intrinsic magnetic moment. Charge coupling to the "orbital" motion is contained in the Hamiltonian (3.12), but (permanent) magnetic moments in a magnetic field have an energy $-\boldsymbol{\mu}\cdot\boldsymbol{B}$, and a corresponding term would have to be added to the Hamiltonian. In terms of the vector potential of the radiation field, for a collection of moments, the expression

$$H_{\mathrm{int}(\mu)} = -\sum_{\alpha} \boldsymbol{\mu}_{\alpha} \cdot \mathrm{curl}\, \boldsymbol{A}(\boldsymbol{r}_{\alpha}, t) \tag{3.13}$$

should be added to H_{int} in Equation (3.12). This term, like that for the $\boldsymbol{A}\cdot\boldsymbol{p}$ coupling, is linear in the electromagnetic field amplitude.

---

[7]This is a rather loose terminology; the designation "photon" should perhaps be reserved for free-particle states. On the other hand, all photons are eventually absorbed and in that sense could even be regarded as virtual.

3.3.3 Quantum-Mechanical Interaction Hamiltonian

For the calculation of the rates for various processes involving coupling to photons, we need the forms for the Hamiltonian terms associated with these couplings. Basically, the couplings are the forms H_{int} for which we must now substitute in the appropriate expression for the vector potential corresponding to a single propagating photon. In terms of the photon wave vector $\boldsymbol{k}$ and unit polarization vector $\boldsymbol{\varepsilon}$, the vector potential is[8]

$$\boldsymbol{a}(\boldsymbol{r}, t) = a_0 \boldsymbol{\varepsilon} \cos(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t). \tag{3.14}$$

In the oscillatory term with the phase $\phi = \boldsymbol{k} \cdot \boldsymbol{r} - \omega t$, if we write

$$\cos \phi = \tfrac{1}{2}(e^{i\phi} + e^{-i\phi}), \tag{3.15}$$

in evaluating the rate for the process by perturbation theory, only one of the terms $(e^{\pm i\phi})$ would be picked out, depending on whether the photon is incoming or outgoing. The task, then, is to determine the amplitude factor a_0; this would then establish the precise forms for each of the three kinds of terms in H_{int}. That is, we would have the $\boldsymbol{a} \cdot \boldsymbol{p}$, $\boldsymbol{a} \cdot \boldsymbol{a}$, and $\boldsymbol{\mu} \cdot \text{curl } \boldsymbol{a}$ couplings that, for various processes, determine the associated rate or cross section.

It is easy to fix the parameter a_0, and there are at least two simple procedures for doing this.[9] The most direct method fixes a_0 by relating the photon energy flux computed with the vector potential (3.14) to what it should be for a simple propagating photon of energy $\hbar\omega = \hbar ck$. If for photon states we take a unit normalization volume, the number density of photons in the state described by $\boldsymbol{\varepsilon}$ and $\boldsymbol{k}$ would be $n_\gamma = 1$. In terms of the photon energy density u_γ, the magnitude of the photon energy flux (Poynting vector) in the direction of $\boldsymbol{k}$ would be

$$S = c\hbar\omega = cu_\gamma. \tag{3.16}$$

If E and B are the magnitude of the electric and magnetic fields carried by the photon, then

$$u_\gamma = (\langle E^2 \rangle + \langle B^2 \rangle)/8\pi = \langle E^2 \rangle /4\pi, \tag{3.17}$$

where the brackets denote a time average. With the electric field $\boldsymbol{E}$ derived from $-(1/c)\partial \boldsymbol{a}/\partial t$, and the time average $\langle \cos^2 \phi \rangle = \langle \sin^2 \phi \rangle = \tfrac{1}{2}$, we combine the above two equations to give

$$a_0 = (2\pi \hbar c^2/\omega)^{1/2}, \tag{3.18}$$

which should be then substituted into Equation (3.14) and employed to fix the forms for H_{int}.

In the other method (see Footnote 9) used to evaluate a_0, some process is computed classically and quantum mechanically and the results are then compared. The simplest process to consider is soft-photon production when a charged particle is accelerated suddenly to a velocity v, say. The probability can be computed for the acceleration to be accompanied by the production of a photon of frequency within $d\omega$.

---

[8]We adopt notation here in which the lower case letter ($\boldsymbol{a}$) is used for the vector potential associated with the state of a single photon.

[9]See R. J. Gould, *Astrophys. J.* **362**, 284 (1990).

This can be done by employing the soft-photon formulas of classical electrodynamics that were derived in Chapter 2, and the comparison with a quantum-mechanical calculation then yields precisely the result (3.18).

The treatment referred to in Footnote 9 uses the notation R, T, and S for the $\boldsymbol{a} \cdot \boldsymbol{p}$, $\boldsymbol{a} \cdot \boldsymbol{a}$, and $\boldsymbol{\mu} \cdot \text{curl } \boldsymbol{a}$ perturbations, standing for "radiation," "two-photon," and "spin." The $\boldsymbol{a} \cdot \boldsymbol{a}$ coupling is quadratic in the electromagnetic field; that is, it involves two fields which can represent two photons. Then, if we write $\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2$ for the total field,

$$\boldsymbol{a} \cdot \boldsymbol{a} = a_1^2 + a_2^2 + 2\boldsymbol{a}_1 \cdot \boldsymbol{a}_2. \tag{3.19}$$

If, for example, there is one incoming and one outgoing photon, the factor 2 multiplying $\boldsymbol{a}_1 \cdot \boldsymbol{a}_2$ would account for the possibility that labels 1 and 2 (fields 1 and 2) could describe either photon. If, on the other hand, we had two outgoing photons, the factor 2 would be regarded as accounting for a direct and exchange amplitude. We represent the coupling T as a "two-photon vertex" for which the photon states are described by $\boldsymbol{\varepsilon}$, $\boldsymbol{k}$ and $\boldsymbol{\varepsilon}'$, $\boldsymbol{k}'$, and to remind the reader of the origin of the factor of 2, we enclose it in parentheses in the formula for T. In the expression for S (coupling to the "spin" moment), the curl operator on the plane-wave photon state can be written as a cross product involving $\boldsymbol{k}$ and $\boldsymbol{\varepsilon}$. For coupling to a charge $q = ze$ and magnetic moment $\boldsymbol{\mu}$, the formulas for the three types of H_{int} perturbations can be summarized as

$$R = \frac{ze}{m} \left(\frac{2\pi \hbar}{\omega} \right)^{1/2} \boldsymbol{\varepsilon} \cdot \boldsymbol{p}, \tag{3.20}$$

$$T = \frac{(2)\pi z^2 e^2 \hbar}{m} \frac{\boldsymbol{\varepsilon} \cdot \boldsymbol{\varepsilon}'}{(\omega\omega')^{1/2}}, \tag{3.21}$$

$$S = -ic \left(\frac{2\pi \hbar}{\omega} \right)^{1/2} \boldsymbol{\mu} \cdot (\boldsymbol{k} \times \boldsymbol{\varepsilon}). \tag{3.22}$$

3.3.4 Perturbation Theory

The standard time-dependent perturbation theory employed for the calculation of rates of processes in quantum-mechanical systems was first developed by Dirac. The procedure, often referred to as the "method of variation of constants," is very general in that it is formulated for an arbitrary system described by a Schrödinger equation

$$(H_0 + H')\psi = i\hbar \, \partial \psi / \partial t. \tag{3.23}$$

Here H_0 is the unperturbed Hamiltonian, which satisfies

$$H_0 \psi^{(0)} = i\hbar \, \partial \psi^{(0)} / \partial t, \tag{3.24}$$

for which there is a spectrum of stationary-state solutions

$$\psi_m^{(0)} = u_m^{(0)} \exp(-i E_m^{(0)} t / \hbar), \tag{3.25}$$

the $u_m^{(0)}$ being independent of time but functions of the coordinates of the particles. The general state of the unperturbed system is

$$\psi^{(0)} = \sum_m a_m^{(0)} \psi_m^{(0)}, \tag{3.26}$$

a superposition of possible states.

The solution to the "perturbed" Schrödinger equation (3.23) is written in the form

$$\psi = \sum_m a_m(t) \psi_m^{(0)}, \tag{3.27}$$

that is, as a linear combination of the unperturbed wave functions. This is a convenient procedure since often we have a problem in which, before the perturbation acts, the system is in some unperturbed state $\psi^{(0)} = \psi_0^{(0)}$ at, say, $t = -\infty$. Thus,

$$\begin{aligned} a_0^{(0)} &= 1, \\ a_m^{(0)} &= 0 \quad (m \neq 0), \end{aligned} \tag{3.28}$$

the subscript 0 referring to the initial state. As a result of the perturbation H', the system makes transitions to different states. After the action of the perturbation, at $t = +\infty$, the system will be in the final state

$$\psi_f(t = +\infty) = \sum_m a_m(t = +\infty) \psi_m^{(0)}, \tag{3.29}$$

and the probability of a particular state k will be given by

$$W_k = |a_k(t = +\infty)|^2. \tag{3.30}$$

Perturbation theory works when H' is small and $W_k \ll 1$, except for W_0. Since we are interested in transitions of various types, the case $k \neq 0$ is important. In general, if the perturbation H' is sufficiently weak,

$$a_m(t) = \delta_{0m} + a_m'(t), \tag{3.31}$$

with $\left|a_m'(t)\right|^2 \ll 1$. If the general solution (3.27) is substituted into the perturbed equation (3.23), an equation for the a's is obtained:

$$(H_0 + H') \sum_m a_m \psi_m^{(0)} = i\hbar \sum_m (a_m \partial \psi_m^{(0)}/\partial t + \dot{a}_m \psi_m^{(0)}). \tag{3.32}$$

Multiplying from the left by $\psi_k^{(0)}$ and integrating over the spatial volume dV (which may be multidimensional), employing the forms (3.25), we have an equation for a_k:

$$\sum_m a_m e^{i\omega_{km}t} H_{km}' = i\hbar \dot{a}_k. \tag{3.33}$$

The orthogonality of the $\psi_m^{(0)}$ has been employed to get a single term on the right, and

$$\omega_{km} = \left(E_k^{(0)} - E_m^{(0)}\right)/\hbar, \tag{3.34}$$

$$H_{km}' = \int \left(u_k^{(0)}\right)^* H' u_m^{(0)} \, dV. \tag{3.35}$$

Note that in the definition (3.35) the matrix element is computed from time-independent wave functions $u_k^{(0)}$ and $u_m^{(0)}$. The perturbation Hamiltonian H' may still be a function of time, so that H'_{km} may be as well. We consider the important special cases where H' is time independent and later where H' is an harmonic function of t.

For the condition (3.28) in which the system started in the initial state $\psi_0^{(0)}$, the coefficients are given by the expression (3.31). When this is substituted into the (exact) equation (3.33) the main contribution on the left side is from $m = 0$ and we have, in a first iteration,

$$a'_k(t) = -(i/\hbar) \int_{-\infty}^{t} e^{i\omega_{k0}t} H'_{k0} \, dt. \tag{3.36}$$

With H'_{k0} independent of time[10] the expressions (3.36) can be integrated; writing the lower limit as $-T$, we have

$$a'_k(t) = \frac{H'_{k0}}{E_0 - E_k} (e^{i\omega_{k0}t} - e^{-i\omega_{ko}T}). \tag{3.37}$$

This result can then be employed in Equation (3.31) and then substituted back into the exact equation (3.33) to give a second-iteration result:

$$i\hbar \dot{a}'_k = e^{i\omega_{k0}t} H'_{k0} + \sum_m \frac{H'_{m0}}{E_0 - E_m} e^{i\omega_{m0}t} e^{i\omega_{km}} H'_{km}. \tag{3.38}$$

In this step, the second term in parentheses in Equation (3.37) has been ignored; this term, for $T \to \infty$, does not contribute, being highly oscillatory and having the essential value zero. Since $\omega_{km} + \omega_{m0} = \omega_{k0}$, the result (3.38) is the same as that for the first iteration in Equation (3.36) with

$$H'_{k0} \to H'_{k0} + \sum_m H'_{km} \frac{1}{E_0 - E_m} H'_{m0}. \tag{3.39}$$

For reasons to be discussed later, the factors in the above sum have been written in this special way.

To simplify the equations, we express them with only the first term in the perturbation series (3.39). By means of the Equations (3.30) and (3.36), the probability for the transition $0 \to k$ occurring due to the action of the perturbation H' can be expressed explicitly in terms of an integral[11] involving the matrix element H'_{k0}:

$$W_k = \frac{1}{\hbar^2} \left| \int_{-\infty}^{\infty} e^{i\omega_{k0}t} H'_{k0}(t) \, dt \right|^2. \tag{3.40}$$

This is an important formula that exhibits a number of significant features. In particular, when the perturbation is a slowly varying function of time, the oscillatory exponential factor makes the value of the integral small. That is, when the perturbation is "adiabatic" (slow) the probability of a transition is small and, instead, the

---

[10] We shall see shortly how to modify our formulas for the case where $H'_{k0} \propto e^{i\omega t}$, as in electromagentic perturbations.

[11] In fact, this integral is 2π times the Fourier transform of $H'_{k0}(t)$.

system adjusts itself gradually to the instantaneous effect of the perturbation. This effect is known as the *Adiabatic Theorem*.

Another very important result can be derived from Equations (3.40) or (3.30) and (3.36). For transitions to continuum states, the transition probability per unit time is of interest. For a particular one (k) of these states, this is given by

$$\frac{\Delta W_k}{\Delta t} = \lim_{\tau \to \infty} \frac{1}{2\tau} \frac{1}{\hbar^2} \left| \int_{-\tau}^{\tau} e^{i\omega_{k0}t} H'_{k0} \, dt \right|^2. \qquad (3.41)$$

When H'_{k0} is independent[12] of time, it can be taken out of the integral, which is then equal to $2i(H'_{k0}/\omega_{k0}) \sin \omega_{k0}\tau$. But since

$$\lim_{\tau \to \infty} \frac{\sin^2 \omega_{k0}\tau}{\omega_{k0}^2 \tau} = \pi \delta(\omega_{k0}) = \pi \hbar \delta(E_k - E_0), \qquad (3.42)$$

we have

$$\frac{\Delta W_k}{\Delta t} = \frac{2\pi}{\hbar} \left| H'_{k0} \right|^2 \delta(E_k - E_0). \qquad (3.43)$$

Here the δ-function manifests energy conservation in the overall process, and the result (3.43) is often written with a "density of final states" $\rho(E_k) = dN_k/dE_k$ instead of the δ-function. Since a summation over final states is always performed in the applications of the result for $\Delta W_k/\Delta t$, the δ-function is employed then, and the effect is the same as with the inclusion of $\rho(E_k)$. The formula (3.43) has many applications and was called "Golden Rule Number Two" by Fermi. It is well to emphasize its general applicability in that it is based essentially on the perturbed general Schrödinger equation (3.23), which does not specify the form of the Hamiltonian. The Golden Rule has assumed that H' is not an explicit function of time; however, we can also employ the result when H' is a harmonic function of time for our special applications.

Another aspect of the Golden Rule formula ought to be emphasized. Although it was derived in a perturbation theory formulation, the formula itself is more general in that it holds *even if the perturbation is not weak*. To understand this, we should recognize that the formulation was a description of "flow of probability," and as long as the expired time is not long, the final-state coefficients $a'_k(t)$ will be small even if the perturbation is not weak. Actually, if H' is strong, the effective matrix element H'_{k0} can be regarded as the result of the multiple action of some coupling evaluated to higher order, assuming that the perturbation series converges.

In one special application of the Golden Rule formula this more general validity is inherently assumed. The application is to the determination of the ratio of cross sections for processes in the forward and reverse directions. The cross-section ratio yields the corresponding ratio of the associated phase-space factors for the forward and reverse processes. The squared matrix elements are the same because the effective coupling Hamiltonian is Hermitian. Then, for example, if the processes

---

[12]Here we mean that H' is not an *explicit* function of time. The results can still be applied, for example, in the calculation of scattering cross sections when the scattering potential (which would act as H') is time independent. In such scattering, the particle experiences a variable perturbation, but as a result of the dependence on position coordinate and not as a result of explicit time dependence.

themselves involve complicated multiple actions of a strong coupling, the ratio of the forward and reverse cross sections can be obtained from the Golden Rule formula.

3.3.5 Processes, Vertices, and Diagrams

The possible electromagnetic processes are determined by the interaction Hamiltonians (3.20), (3.21), and (3.22). For an overall process there may be other perturbations involved and, correct to second order in the perturbations, the rate would be determined by the effective perturbation matrix element (3.35). We are restricting our treatment to processes involving only radiation-field photons, since this allows simplification. At the same time, it limits the number of processes that our elementary theory can treat. However, the techniques employed in the restricted theory are very similar to those in general QED, thereby providing a useful introduction to the subject.

Of special interest are processes that occur as a result of the action of two perturbations, one or both of which are electromagnetic (that is, involving production or destruction of photons). If U and V are the interaction Hamiltonians associated with the perturbations, the effective perturbation matrix element for the combined process would be, generalizing[13] the result (3.39),

$$H'_{f0} = \sum_I U_{fI} \frac{1}{E_0 - E_I} V_{I0} + \sum_{I'} V_{fI'} \frac{1}{E_0 - E_{I'}} U_{I'0}. \tag{3.44}$$

Here I and I' are "intermediate" states and these states are summed over. The probability of the process between the initial state 0 and the final state f would be proportional to $\left|H_{f0}\right|^2$ and would determine a cross section or transition probability. The states 0 and f are observed or specified in the sense that the system is considered to be measured or detected in these states. The intermediate states are, of course, not observed, and it is essential in the evaluation of the overall $0 \to f$ process that all of these accessible intermediate states be included in the total amplitude for the process.

The probability of a process resulting from two perturbations can be written in the form

$$W_{f0} = \left| \sum_I A_{fI} A_I A_{I0} \right|^2. \tag{3.45}$$

An individual amplitude is the product of these factors, reading from right to left: (i) an amplitude (A_{I0}) associated with the action of a perturbation acting as to cause the system to make a transition $0 \to I$, (ii) an amplitude (A_I) that is only a function of the intermediate state, and (iii) an amplitude (A_{fI}) from a perturbation acting to cause a transition $I \to f$. If a process can take place through the action of a single perturbation, there is no intermediate state and there would be only the amplitude A_{f0} in lowest order. On the other hand, for a given process, even in lowest order in a

---

[13]This can be done by simply replacing H' with $U + V$, yielding the four types of second-order terms (UU, VV, UV, VU), each with an energy denominator.

coupling constant,[14] the total amplitude may involve both a simple direct amplitude A_{f0} and a combination of amplitudes involving intermediate states as in the form (3.45). That is, in general, the total amplitude for a process can involve a number of perturbations and intermediate states.

The "intermediate-state amplitude" A_I is just the factor $(E_0 - E_I)^{-1}$ involving the energy denominators. It corresponds to, in modern covariant perturbation theory, the Feynman *propagator* or propagation factor. Although the idea has limited physical meaning, being simply a factor in a perturbation-theory development, it is a convenient notion to introduce. As is indicated in the notation in Equation (3.45), we can think of the total amplitude for the process involving two perturbations as if A_I were an amplitude for propagation in the intermediate state between the two interactions. Basically, this is the reason why the factor with the energy denominator in Equation (3.39) was written in between the two matrix elements. It is also convenient (but not necessary) to introduce a pictorial representation of the process whereby a particle undergoes the transition $0 \rightarrow f$ by means of the two perturbations U and V and the intermediate states I and I' [see Equation (3.44)]. The diagrams represent terms in a perturbation series and are nothing more than a bookkeeping device. Note that time runs vertically in these pictures, with the initial state indicated at the bottom and the final state at the top. The horizontal scale can represent position in a rough sense, so that the picture is like that of a "world line," taking some terminology from relativity. An actual path or line for the particle is, of course, not implied, since this would be a classical notion foreign to our quantum-mechanical formulation. These diagrams are not to be taken "literally"; rather, they serve as a guide or reminder in writing down terms in a perturbation series.

Two diagrams are indicated in Figure 3.1, corresponding to the second-order perturbation for the case where the perturbations U and V are different and able to act in either order to cause the transition $0 \rightarrow I$ or $I' \rightarrow f$. An important example of a problem of this type would be that of bremsstrahlung, which takes place through the combined process of scattering (U, say) and photon production (V). Here, by "photon production" is meant the interaction (3.20) between the charge and the field of the outgoing photon. Charged particles are always subjected to the photon-emission Hamiltonian [Equations (3.20) as well as (3.21) and (3.22)]. However, energy conservation does not permit an isolated charge to produce a photon. An additional perturbation is required, such as a scattering potential, and through the combined action of both perturbations the photon can be produced. This has a corresponding feature in classical radiation theory in which emission takes place because of the charge's acceleration (as a result of the scattering potential).

The interaction associated with photon emission can be represented in a picture by a *vertex* with a wavy line designating the photon. Both the Hamiltonian (3.20) corresponding to the interaction of the charge with the radiation field and the expression (3.22) from the interaction of the intrinsic magnetic moment can be represented by a vertex of the type in Figure 3.2. Again with time running in the vertical direction, this picture is a designation of a matrix element H'_{ba} in which a and b refer to the

---

[14]As noted already, the perturbation (3.20) is first order in the charge while the perturbation (3.21) is second order.
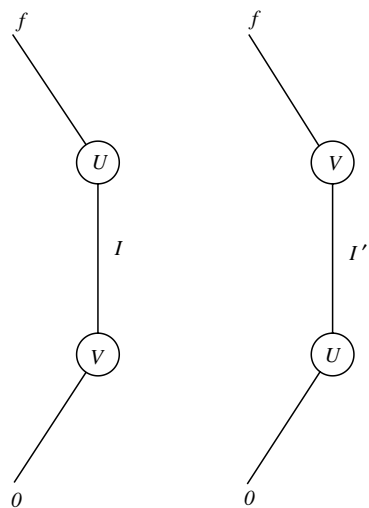
Figure 3.1 Pictorial representation of the action of two perturbations. The vertices represent
the matrix elements of the U and V interactions and the line between the vertices
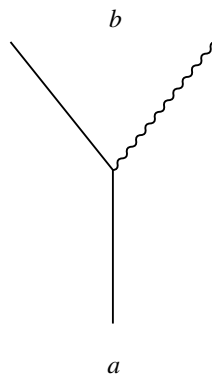represents the energy denominator factor.

Figure 3.2 Photon-coupling vertex corresponding to photon "production" (outgoing photon).

initial and final states, respectively, with the state b containing a photon. Since an
isolated charge or magnetic moment cannot emit a photon, the vertex of Figure 3.2
must be just part of a diagram if a photon process is being represented. As noted
above, it could be part of the bremsstrahlung diagrams, and we see that the two
diagrams correspond to emission "before and after" the scattering. Again, it should
be emphasized that this terminology, and the diagrams, should not be taken literally.
"Emission" is associated with the probability for the process, and the probability is
the squared total amplitude. The total amplitude itself is not observed nor are its
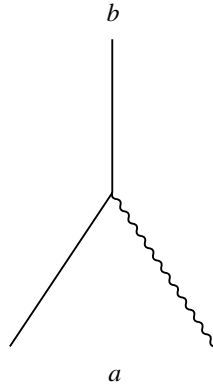individual components associated with individual diagrams.

Figure 3.3 Photon-coupling vertex associated with photon "absorption".

Photons can be absorbed as well, and the vertex or matrix element of the Hamiltonian (3.20) associated with interaction of an incoming photon by this coupling would be represented as in Figure 3.3. Basically, the vertices of Figures 3.2 and 3.3 are the same, being the time reverse of one another. However, they do represent distinctly different interactions or perturbations, since one involves an incoming photon and the other an outgoing photon. This means, for example, that in the evaluation of the process of photon scattering by a free charge (Compton scattering), two diagrams occur that involve the vertices of Figures 3.2 and 3.3. That is, if these vertices correspond to the interactions U and V, respectively, of Figure 3.1, the "absorption" and "emission" in the two-step scattering process can occur in either order.[15] Although the basic nature of the interaction is the same (for U and V) in this case, the two photons involved are different and so, therefore, are the associated perturbations. In a later chapter, we give a more complete discussion of Compton scattering, and we shall see that the two diagrams referred to above actually do not contribute to the cross section. The whole contribution comes from a third diagram that is associated with the interaction Hamiltonian (3.21). This diagram, a vertex associated with particle coupling to two photons or two photon fields, would be, for one incoming and one outgoing photon, as shown in Figure 3.4. That is, the coupling involves the point interaction of the fields of four particles: the incoming and outgoing charges and the two photons. The interaction is of higher order than that associated with the coupling (3.20) (and the diagrams in Figure 3.2 or 3.3), being second order in the charge.

The intrinsic magnetic moment coupling (3.22) to the photon field has a vertex looking like that for the coupling (3.20), that is, like Figures 3.2 or 3.3. However, for processes involving non-relativistic charges, this coupling is weaker than the

---

[15]Again, we do not mean this to be taken literally. Without performing an experiment to do so—and thereby disturbing the process—we cannot know which perturbation acted first. The separate diagrams refer only to mathematical terms in a perturbation-theory formulation.
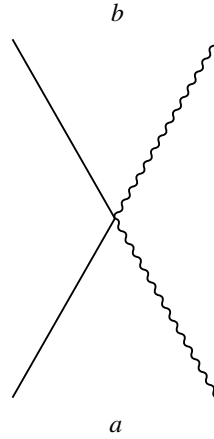
Figure 3.4 Diagram representing the coupling (vertex) of two photon fields, in this case an incoming and an outgoing photon.

lowest-order term R. Since $\mu_\alpha \sim q_\alpha \hbar / m_\alpha c$, we see that

$$R/S \sim k/k_\alpha \sim \hbar\omega/(m_\alpha c^2 E_\alpha)^{1/2}, \qquad (3.46)$$

where E_α is the particle kinetic energy. The ratio (3.46) is small, being $\sim (E_\alpha/m_\alpha c^2)^{1/2}$ for $\hbar\omega \sim E_\alpha$. If, on the other hand, a particle had no charge but possessed a magnetic moment, the coupling (3.22) would be the only interaction with the photon field. It is perhaps appropriate to remark at this point that covariant perturbation theory for spin-$\frac{1}{2}$ particles has only one basic interaction or vertex. This vertex is of the type in Figure 3.2 (or 3.3), associated with a linear coupling to the electromagnetic field. The non-relativistic theory, on the other hand, has three basic couplings (3.20)–(3.22), and so the transition to the relativistic theory is somewhat complex. The non-relativistic, non-covariant theory is really quite different from the covariant formulation, and there is not a one-to-one correspondence between the vertices and diagrams in our treatment here and those associated with the relativistic theory.

The "elementary" particles have an important fundamental property—their particle-antiparticle symmetry. That is, in addition to the photon, proton, neutron, electron, pion, muon, and neutrino (electron and muon types), there are the corresponding antiparticles. The antiphoton is absolutely identical to the photon ($\bar{\gamma} = \gamma$). The neutron is also neutral, but it has a magnetic moment and for $\bar{n}$ there is a sign difference in the relation between magnetic moment and spin. The antineutrino has a helicity opposite in sign to that of the neutrino. The antielectron (positron) is identical to the electron except for being oppositely charged (which also has a corresponding effect on its magnetic moment). This is also true of the other charged "elementary" particles[16] (proton, muon, and charged pion). The particle-

---

[16]On the other hand, the "strange" particles do not possess this symmetry between counterparts of opposite charge.
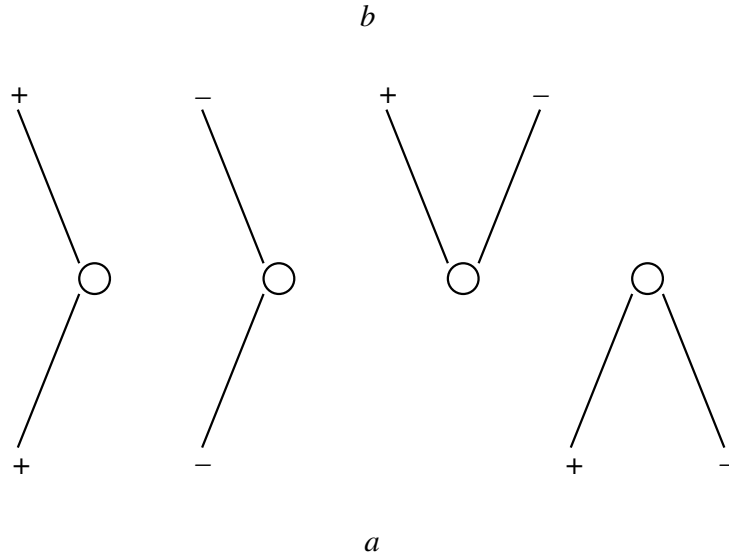
$$b$$

Figure 3.5 Particle and antiparticle diagrams.

antiparticle symmetry means, for example, that in quantum electrodynamics a single unified theory describes electron and positron processes. In particular, it means that, consistent with charge conservation, for any perturbation matrix element

$$V_{ba} = \langle b|V|a \rangle \,, \tag{3.47}$$

the only requirement on the states a and b is that they be consistent with charge conservation. This can be represented pictorially as in Figure 3.5, which gives the four possibilities for interactions involving electrons and/or positrons. In these vertices or diagrams, the circle can represent any kind of interaction. It can correspond to some external potential or one of the three electromagnetic perturbations (3.20)–(3.22). For the latter perturbations, the photons involved can be either incoming or outgoing in all possibilities, so that there are many subcategories[17] for the vertices in Figure 3.5. Some of these specific vertices may have a matrix element that is zero; this is, in fact, the case, as will be seen later.

The two vertices on the right of Figure 3.5 would be part of processes involving pair production and annihilation, respectively. Because of the basic symmetry of the theory, a unified treatment of these processes can be given in terms of a basic interaction that can cause a variety of processes. For example, bremsstrahlung is closely related to pair production (and annihilation), the diagrams for the processes being essentially the same when one set is placed on its side. These processes will be treated in detail in later chapters.

---

[17]For the interactions (3.20) and (3.22), each of the vertices in Figure 3.5 can involve either an incoming or an outgoing photon. For the interaction (3.21), each vertex could involve an incoming and an outgoing photon, two incoming photons, or two outgoing photons.

3.4 RELATIVISTIC THEORY

Although the treatment of non-relativistic QED in the last section is fairly self-contained, only a very superficial formulation of the relativistic theory can be given here.[18] A substantive exposition of modern QED would require much more space, but perhaps we can provide some understanding of the essence of the theory in this brief outline.

3.4.1 Modifications of the Non-Covariant Formulation

The early ($\sim$1930) formulations of relativistic QED were not manifestly covariant and were somewhat complicated in their derivations of the foundations and in the associated computational techniques in the applications of the theory. The newer ($\sim$1950) methods are preferable and benefit from the more "natural" formulation in terms of covariant equations. Although no new physics is introduced in the covariant formulation, it is quite different in its mathematical layout and in its subsequent computational methods. In the interest of also retaining some simplicity, let us now consider how the non-relativistic theory might be modified.

Some aspects of the theory can be considered to be very general, and we can expect to carry over some of the principles introduced in the non-relativistic development. Again, it is convenient to formulate a perturbation theory, and certain features of the non-covariant formulation in the last section are retained. The non-covariant perturbation theory is still a valid procedure in applications involving even relativistic particles, since it is based on the general Schrödinger equation (3.23) and the superposition principle. Transitions are viewed as the result of the action of perturbations, and more than one perturbation may be required to produce a non-vanishing total amplitude for a certain process. In the mathematical formulation, the perturbations result in matrix elements associated with transitions between two states, and we have found it convenient to introduce the notion of a "vertex" in the pictorial representation of the transition. One or both of these states may be "intermediate" so that the "transition" involving a particular vertex may not be directly related to an observable event; in that case, the associated matrix element is just one of the factors in the expression for the total amplitude for the process. To be consistent with the general literature on the subject, we use the notation, say, M_{lk} for one of these matrix elements (vertices) corresponding to the $k \rightarrow l$ transition.

In an amplitude for a process in which intermediate states are involved, it can be expected that there will be a factor corresponding to the energy denominator $(E_0 - E_I)^{-1}$ in the non-covariant formulation. This factor is a function only of the characteristics of the intermediate state, and in covariant perturbation theory it is called a (Feynman) *propagator*. For now we designate it as P_I for an intermediate state I, and we try to give simple arguments for inferring its expected form. For a process corresponding to some (total) transition $0 \rightarrow f$ in which an arbitrary number of intermediate states are involved, the probability or rate would be given

---

[18]There are a large number of excellent textbooks on relativistic QED. A selection of these, chosen for their variety of approach and penetration of the subject, are listed at the end of this chapter.

by an expression of the form

$$W_{f0} = \left| \sum_{I_k \cdots I_n} M_{f I_n} P_{I_n} M_{I_n I_m} \cdots M_{I_l I_k} P_{I_k} M_{I_k 0} \right|^2 . \tag{3.48}$$

The total amplitude in this expression is meant to be read from right to left. Loosely stated, the matrix element $M_{I_k 0}$ causes the transition from the initial state 0 to the intermediate state I_k, there is then a "propagation" (P_{I_k}) to the next interaction or vertex $M_{I_l I_k}$, and so forth, to the last interaction $M_{f I_n}$. The M's are just the perturbation Hamiltonians, and, if there is more than one such perturbation, any one of them can be among the string of factors in Equation (3.48). That is, there can be various combinations of interactions resulting in a variety of processes, and for a given process there may be several combinations yielding the total amplitude for the process.

In covariant perturbation theory, the matrix elements M_{lk} are evaluated from a spacetime integration over the invariant four-dimensional volume $d^4 x$ or, in the momentum representation, over $d^4 p$. This results in a fundamental difference from the non-covariant theory in which matrix elements are evaluated from an integration over the three-dimensional spatial volume $d^3 \boldsymbol{r}$ (or $d^3 \boldsymbol{p}$). The time (as well as space) integration in the matrix elements yields a simplification in the perturbation theory, which is described conveniently in terms of a diagrammatic representation. As noted by Stückelberg in 1942 and fully exploited by Feynman in 1948, positrons—holes in the sea of electrons in negative energy states in Dirac's relativistic electron theory—can be described as electrons moving backward in time.[19] In a process that involves, for example, two interactions, that is, two matrix elements, non-covariant perturbation theory must include the two types of diagrams indicated in Figure 3.6. One involves pair production with the particle of opposite sign to the incoming particle annihilating with the latter; this diagram corresponds to an intermediate state with three particles. On the other hand, in the covariant theory, in which there is an integration over all t for each interaction matrix element, both diagrams on the left in Figure 3.6 are effectively included in a single one. That is, in following the particle's "world line" in covariant theory, we do not distinguish the two diagrams on the left, which are regarded as essentially the same. This can also be seen through consideration of a Lorentz transformation or rotation in the space-time plane; in another Lorentz frame, the pair-production diagram is that on the far left in the figure. The non-covariant theory is, as we have stated earlier, valid—even for relativistic problems. However, it is not as convenient as the covariant theory, which is fundamentally the more natural approach. In non-relativistic problems, the non-covariant theory is more convenient, and the pair-producing diagram is negligible. This is because the corresponding intermediate-state energy is very large, and the factor $(E_0 - E_I)^{-1}$ is very small in the term in the effective perturbation Hamiltonian [see Equation (3.44)].

There is another feature of covariant perturbation theory that may be considered a consequence of the four-dimensional integrations in matrix elements. Matrix

---

[19]This can be seen in the phase factor $k_\mu x_\mu = \boldsymbol{k} \cdot \boldsymbol{r} - \omega t$ in the wave function of a propagating particle. Replacing $\omega \, (= E/\hbar)$ by $-\omega$ corresponds to a description in which the direction of motion is changed unless $t \to -t$.
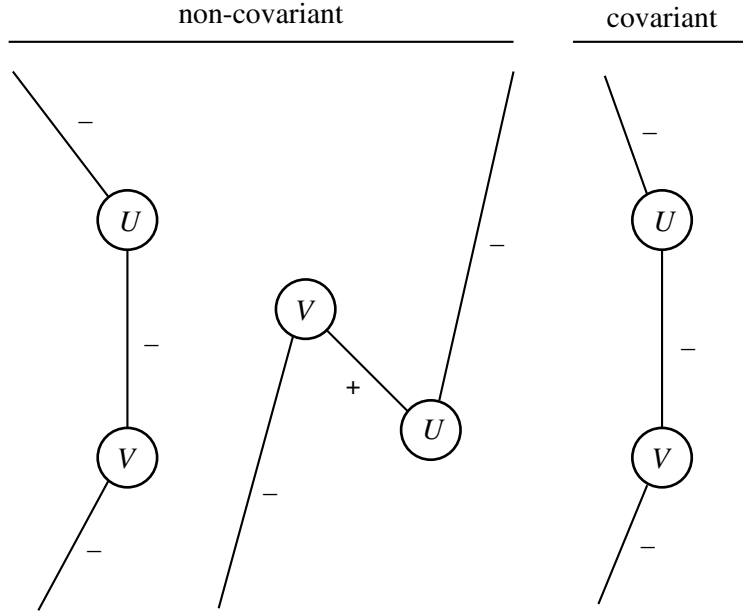
Figure 3.6 Diagrams representing the action of two perturbations in the non-covariant formulation and in the covariant formulation.

elements in the non-covariant theory yield momentum-conservation δ-functions, since the position-space integrals are always of the form

$$\int e^{i\boldsymbol{\kappa}\cdot\boldsymbol{r}} d^3\boldsymbol{r} = (2\pi)^3 \delta^{(3)}(\boldsymbol{\kappa}), \tag{3.49}$$

[see Equation (2.68)], where $\delta^{(3)}(\boldsymbol{\kappa}) = \delta(\kappa_x)\delta(\kappa_y)\delta(\kappa_z)$, and

$$\boldsymbol{\kappa} = \sum_{\text{in}} \boldsymbol{k} - \sum_{\text{out}} \boldsymbol{k} \tag{3.50}$$

is just the difference between the sum of the wave vectors (or momenta) for the incoming and outgoing particles involved in a vertex. Momentum conservation results because the matrix elements are evaluated from

$$M_{lk} = \int u_l^* M u_k \, d^3\boldsymbol{r} \quad \text{(non-cov.)}, \tag{3.51}$$

with the two spatial wave functions being products of (unperturbed) factors $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$. In the covariant theory, the matrix elements involve the space- and time-dependent free-particle wave functions (3.25)

$$M_{lk} = \int \psi_l^* M \psi_k \, d^4x \quad \text{(cov.)}, \tag{3.52}$$

with the ψ's of the form $\exp(ik_\mu x_\mu)$. The integrals then yield the result

$$\int e^{i\kappa_\mu x_\mu} d^4x = (2\pi)^4 \delta^{(4)}(\kappa_\mu) = (2\pi)^4 \delta(\kappa_0)\delta^{(3)}(\boldsymbol{\kappa}), \tag{3.53}$$

corresponding to conservation of *both* momentum and energy at each vertex. This then represents a fundamental difference from non-covariant perturbation theory where there was not energy conservation at vertices and in intermediate states. Consequently, diagrams and factors in amplitudes for processes do not have the same precise meaning in the non-covariant and covariant formulations.

The starting point for all relativistic theories is an Action Principle:

$$\delta \int \mathcal{L} \, d^4 x = 0, \tag{3.54}$$

where $\mathcal{L}$ is a scalar Lagrangian density. The Lagrangian is a function of the various field components $\phi^\alpha(x_\mu)$ and their derivatives $\partial \phi^\alpha / \partial x_\mu$, where the index α refers to the various types of particles (photons, electrons, etc.). For the integral (3.54) to have a stationary value for independent variations of the field components, the set of field equations must hold:

$$\frac{\partial \mathcal{L}}{\partial \phi^\alpha} - \frac{\partial}{\partial x_\mu} \left(\frac{\partial \mathcal{L}}{\partial(\partial \phi^\alpha / \partial x_\mu)} \right) = 0. \tag{3.55}$$

The Lagrangian is constructed to yield the correct equations of motion, including the effects of interactions, when substituted into the set of relations (3.55). If only electrons, positrons, and photons are involved, $\mathcal{L}$ would be of the form

$$\mathcal{L} = \mathcal{L}_{\mathrm{D}} + \mathcal{L}_{\mathrm{M}} + \mathcal{L}_{\mathrm{I}}, \tag{3.56}$$

consisting of a Dirac (electrons and positrons), Maxwell (photons), and interaction part. Since the Dirac equation is quantum mechanical, the Lagrangian (3.56) is as well, although further steps must be taken to develop the full quantum field theory. Since the theory must allow for production and destruction of photons and particles (in pairs), it is inherently "many-particle" in nature. That is, the ψ in the Dirac Lagrangian is not a one-particle wave function but an operator for the Dirac *field*.

The older formulations of relativistic QED, developed around 1930, were from a "Hamiltonian" approach with commutation relations introduced for the field operators. We now know that this method of field quantization is not very good. It is cumbersome and not fully covariant in its formulation. Feynman, Schwinger, and Dyson showed around 1950 how to reformulate the subject in a much better way. Dyson, in particular, gave a systematic treatment of QED by making the appropriate modifications in the non-covariant field theory to make the formulation covariant at every stage. In addition to demonstrating the connection with the work of Schwinger, Dyson was able to provide a basis for the so-called Feynman rules (and Feynman diagrams) as a logical consequence of covariant perturbation theory. Rather than try to reproduce this formulation, which would have to be lengthy, we attempt to give a shortcut superficial account that may at least provide some understanding of the theory and its applications.

3.4.2 Photon Interactions with Charges without Spin

The simplest, relativistic, one-particle wave equation corresponding to the non-relativistic Schrödinger equation was first written down by Schrödinger himself. However, to avoid confusion with *the* Schrödinger equation, it is referred to as the

Klein-Gordon equation, named after others who first considered relativistic wave equations. If ψ is a scalar function of only r and t (that is, not describing any spin states), a free-particle wave equation could be written as $i\hbar\partial\psi/\partial t = H\psi$ with $H = (p^2c^2 + m^2c^4)^{1/2}$ as the Hamiltonian, replacing p^2 by the operator $-\hbar^2\nabla^2$ as in the non-relativistic theory. This wave equation with the square root is unwieldy, and there is also the problem of the appropriate sign outside the squre root. No one has made progress with the equation and, instead, an equation was formed from the relativistic relation $p_\mu p_\mu = -m^2c^2$, replacing p_μ with $-i\hbar\partial_\mu$ and allowing the terms to operate on $\psi(r, t)$. This is what is now known as the Klein-Gordon (KG) equation[20]:

$$(\partial_\mu\partial_\mu - \kappa^2)\psi = (\Box^2 - \kappa^2)\psi = 0, \qquad (3.57)$$

where $\kappa = mc/\hbar$. For particles satisfying such one-particle equations, a field theory can be formulated to describe many-particle states. If there are several types of such particles, characterized by their mass (κ) parameter, the Lagrangian density for the scalar field of each type would be of the form[21]

$$\mathcal{L} = \mathcal{L}_{KG} = -\tfrac{1}{2}(\partial_\mu\phi\partial_\mu\phi + \kappa^2\phi^2); \qquad (3.58)$$

by the Lagrange equation (3.55), the field then satisfies a Klein-Gordon equation of the form (3.57):

$$(\Box^2 - \kappa^2)\phi = 0. \qquad (3.59)$$

The sign and factor in front of the Lagrangian function (3.58) are arbitrary for yielding Equation (3.59). However, with this choice the canonical momentum is

$$\pi = \partial\mathcal{L}/\partial\dot{\phi} = \dot{\phi}/c^2, \qquad (3.60)$$

and the Hamiltonian

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L} = \tfrac{1}{2}\pi^2 + \tfrac{1}{2}(\nabla\phi)\cdot(\nabla\phi) + \tfrac{1}{2}\kappa^2\phi^2 \qquad (3.61)$$

is positive definite.

One problem with the Klein-Gordon equation is that if the probability current density is given, as in the non-relativistic Schrödinger theory, by

$$j = (\hbar/2mi)(\psi^*\nabla\psi - \psi\nabla\psi^*), \qquad (3.62)$$

for the conservation equation

$$\partial\rho/\partial t + \nabla\cdot j = \partial_\mu j_\mu = 0 \qquad (3.63)$$

to hold, the probability density must be given by

$$\rho = \frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \frac{\partial\psi^*}{\partial t}\psi\right). \qquad (3.64)$$

Although this expression reduces to the correct non-relativistic limit, it is not positive definite in general. The theory remained in this stage until 1934 when Pauli and

---

[20]In the various textbooks, the reader will notice sign differences in terms in equations associated with the relativistic theory. This has to do with the somewhat arbitrary choice for the "metric." For example, in many texts $\Box^2$ is meant to designate $(1/c^2)\partial^2/\partial t^2 - \nabla^2$, which is the same as our $\partial_\mu\partial_\mu$ but with a minus sign.

[21]In this function the ∂_μ is meant to operate only on the ϕ to its immediate right.

Weisskopf reinterpreted the density ρ of the Klein-Gordon theory in terms of a charge density. The formulation of a theory of scalar particles of positive and negative charge is conveniently done by introducing a complex scalar field

$$\phi = 2^{-1/2}(\phi^{(1)} + i\phi^{(2)}), \tag{3.65}$$

where $\phi^{(1)}$ and $\phi^{(2)}$ are two real scalar fields. They can be written in terms of ϕ and ϕ^* as

$$\phi^{(1)} = 2^{-1/2}(\phi + \phi^*),$$
$$\phi^{(2)} = (1/2^{1/2}i)(\phi - \phi^*). \tag{3.66}$$

The total Lagrangian $\mathcal{L} = \mathcal{L}^{(1)} + \mathcal{L}^{(2)}$, with $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$ of the form (3.58), is then, very simply,

$$\mathcal{L} = -\partial_\mu \phi^* \partial_\mu \phi - \kappa^2 \phi^* \phi. \tag{3.67}$$

Here ϕ^* and ϕ are to be considered as independent fields (instead of $\phi^{(1)}$ and $\phi^{(2)}$). Both ϕ and ϕ^* satisfy Equation (3.59), as we see on applying the Lagrange equation (3.55) separately for ϕ and ϕ^* using the Lagrangian (3.67). Further, $\pi = \partial\mathcal{L}/\partial\dot\phi = \dot\phi^*/c^2$ and $\pi^* = \partial\mathcal{L}/\partial\dot\phi^* = \dot\phi/c^2$; the Hamiltonian is

$$\mathcal{H} = \pi\dot\phi + \pi^*\dot\phi^* - \mathcal{L}$$
$$= \pi^*\pi + (\boldsymbol{\nabla}\phi^*) \cdot (\boldsymbol{\nabla}\phi) + \kappa^2\phi^*\phi. \tag{3.68}$$

We can construct a four-current density of the form

$$j_\mu = \text{const}(\phi^{(1)}\partial_\mu\phi^{(2)} - \phi^{(2)}\partial_\mu\phi^{(1)}), \tag{3.69}$$

and since $\square^2\phi^{(1)} = \kappa^2\phi^{(1)}$ and $\square^2\phi^{(2)} = \kappa^2\phi^{(2)}$,

$$\partial_\mu j_\mu = 0. \tag{3.70}$$

The current can be written conveniently in terms of ϕ and ϕ^*:

$$j_\mu = \text{const}\left[\phi\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} - \phi^*\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\right], \tag{3.71}$$

with, for example, $\rho = \text{const}(\pi\phi - \pi^*\phi^*)$. This gives some indication of how the theory is applied to describe scalar particles of equal and opposite sign.

This theory, with the Lagrangian (3.67), can be used as a starting point for the description of the interaction of charged pions (π^+, π^-) with photons. From the free-particle Lagrangian (3.67), the interaction Lagrangian can be obtained by replacing ∂_μ with $\partial_\mu - (iq/\hbar c)A_\mu$:

$$\mathcal{L}_{\text{KG}} + \mathcal{L}_{\text{I}} = -\left(\partial_\mu\phi - \frac{iq}{\hbar c}A_\mu\phi\right)^*\left(\partial_\mu\phi - \frac{iq}{\hbar c}A_\mu\phi\right) - \kappa^2\phi^*\phi. \tag{3.72}$$

The interaction Lagrangian is then (see, also, Footnote 21)

$$\mathcal{L}_{\text{I}} = \frac{iq}{\hbar c}(\phi\partial_\mu\phi^* - \phi^*\partial_\mu\phi)A_\mu + \left(\frac{q}{\hbar c}\right)^2\phi^*\phi A_\mu A_\mu. \tag{3.73}$$

To the Lagrangian (3.72) should be added that for a pure photon field. This is usually referred to as the *Maxwell Lagrangian* ($\mathcal{L}_M$) and a convenient form has been introduced by Fermi[22]:

$$
\begin{aligned}
\mathcal{L}_M &= -\frac{1}{16\pi} F_{\mu\nu} F_{\mu\nu} - \frac{1}{8\pi}(\partial_\mu A_\mu)^2 \\
&= -\frac{1}{16\pi}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) - \frac{1}{8\pi}(\partial_\mu A_\mu)^2.
\end{aligned}
\tag{3.74}
$$

The total Lagrangian for charged spinless bosons interacting with photons is then

$$
\mathcal{L} = \mathcal{L}_{KG} + \mathcal{L}_M + \mathcal{L}_I.
\tag{3.75}
$$

The coupling terms in the interaction Lagrangian (3.73) are similar in form to the forms (3.20) and (3.21) that arise in the non-relativistic theory. And, of course, the diagrams that are introduced as a guide in a perturbation-theory description look the same. Because the non-relativistic theory has fundamental differences from the covariant formulation, we have not previously referred to them as "Feynman diagrams," believing that the designation should be reserved for the covariant theory. Although we cannot give a complete description of this theory, we can infer something about it in a simple way, and the form (3.73) can be employed for this purpose.

The first term on the right in the Lagrangian (3.73) has a coupling strength linear in the charge and is associated with, say, the interaction of two charged meson fields with a photon field. It corresponds to a vertex with two meson lines and one photon line like that in Figure 3.2 or 3.3. All lines can be incoming or outgoing. If one meson (field) is incoming and the other is outgoing, the matrix element associated with the interaction or vertex is of the form

$$
M_{ba}^{(q)} \propto q\varepsilon_\mu(k_{a\mu} + k_{b\mu}),
\tag{3.76}
$$

where ε_μ is the photon polarization four-vector arising from the photon field $A_\mu \propto \varepsilon_\mu \exp(\pm i k_\mu x_\mu)$, where the $+ (-)$ would be associated with an incoming (outgoing) photon, and $k_{a\mu}$ and $k_{b\mu}$ are the wave vectors for the incoming and outgoing mesons. Since the meson fields are of the form $\exp(ik_{a\mu}x_\mu)$ and $\exp(ik_{b\mu}x_\mu)$, the matrix element is a result of the integration

$$
\begin{aligned}
M_{ba}^{(q)} &\propto q\varepsilon_\mu(k_{a\mu} + k_{b\mu}) \int d^4x \, e^{i(k_{a\mu} - k_{b\mu} \pm k_\mu)x_\mu} \\
&= (2\pi)^4 q\varepsilon_\mu(k_{a\mu} + k_{b\mu})\delta^{(4)}(k_{a\mu} - k_{b\mu} \pm k_\mu).
\end{aligned}
\tag{3.77}
$$

The δ-function is simply a manifestation of energy and momentum conservation at the vertex.

---

[22]We shall not employ the specific form for $\mathcal{L}_M$ given here, but quote the result for completeness (see references at end of chapter). It is expressed here in c.g.s. or Gaussian units, but most books on QED use the Heaviside-Lorentz units, which have the advantage of eliminating factors of $(4\pi)^{-1}$ in equations for electromagnetic phenomena. In these (HL) units, however, the electronic charge has a different value; for example, the fine-structure constant is given by $(e^2/4\pi\hbar c)_{HL} = \alpha \simeq 1/137$. The reader should be aware of the use of these different unit systems.

The second term in the interaction Lagrangian (3.73) is second order in the charge and involves two meson fields and two photon fields (see Figure 3.4). It yields a matrix element of the form

$$M_{ba}^{(q^2)} \propto q^2 \varepsilon_{1\mu} \varepsilon_{2\mu} \int d^4x \, e^{i(k_{a\mu} - k_{b\mu} \pm k_{1\mu} \pm k_{2\mu})x_\mu}$$
$$= (2\pi)^4 q^2 \varepsilon_{1\mu} \varepsilon_{2\mu} \delta^{(4)}(k_{a\mu} - k_{b\mu} \pm k_{1\mu} \pm k_{2\mu}),$$

(3.78)

where the $+$ $(-)$ sign on the photon four-momenta is associated with incoming (outgoing) states.

The form of the couplings can be obtained in another way by making the replacement $\partial_\mu \rightarrow \partial_\mu - (iq/\hbar c)A_\mu$ directly in the Klein-Gordon equation $(\partial_\mu \partial_\mu - \kappa^2)\phi = 0$. The "perturbed" equation is then

$$(\Box^2 - \kappa^2)\phi = (iq/\hbar c)[A_\mu \partial_\mu \phi + \partial_\mu(A_\mu \phi)] + (q/\hbar c)^2 A_\mu A_\mu \phi.$$

(3.79)

The right-hand side of this equation represents source terms $S^{(q)}$ and $S^{(q^2)}$ and is linear in the charged boson field ϕ. Transition matrix elements would be evaluated from integrals of the form $\int \phi_b^* S_a d^4x$. The q-term of the source (3.79), with ∂_μ operating on ϕ_a and A_μ, yields

$$M_{ba}^{(q)} \propto q\varepsilon_\mu(k_{a\mu} + k_{b\mu} \pm k_\mu) \int d^4x \, e^{i(-k_{b\mu} + k_{a\mu} \pm k_\mu)}$$
$$= (2\pi)^4 q\varepsilon_\mu(k_{a\mu} + k_{b\mu}) \, \delta^{(4)}(k_{a\mu} - k_{b\mu} \pm k_\mu),$$

(3.80)

identical to the result (3.77). In a similar manner, the q^2-term (3.78) is again obtained.

The other fundamental type of quantity in the amplitude for a process is the Feynman propagator P_I. Let us see[23] if the form for this factor can be inferred from simple arguments. We consider again the effects of the action of the two perturbations V and U in a process involving a scalar charged particle (see Figure 3.6 and Section 4.1). As we have already seen, in a covariant formulation, the two diagrams of the non-covariant theory are replaced by one. The essential form of the Feynman propagator for scalar particles can be obtained if we start from the non-covariant amplitude in terms of energy denominators and try to inject some simple ideas of the modern theory. One is a more suitable covariant definition of interaction matrix elements in terms of "entry" and "exit" states: $\langle \text{exit}|M|\text{entry}\rangle$. With this description, we follow the "world line" of a particle even it goes backward in time. For a process involving, say, a π^-, the right-hand, non-covariant diagram in Figure 3.6 would correspond to the initial production (by U) of a π^+, π^- pair with the π^+ then annihilating (V) with the incident π^-. The diagram on its left would be associated with an amplitude

$$A_1 = U_{fI} P_I V_{I0},$$

(3.81)

[23]Here we are relying heavily on R. P. Feynman, *Theory of Fundamental Processes*, New York: W. A. Benjamin, Inc., 1962.

with $P_I = (E_0 - E_I)^{-1}$. For the diagram with pair production (or with the π^- going backward in time), taking the exit-entry definition of matrix elements, the amplitude is

$$A_2 = V_{I0} P_I' U_{fI}, \tag{3.82}$$

where $P_I' = [E_0 - (E_I + E_0 + E_f)]^{-1} = -(E_0 + E_I)^{-1}$, since $E_0 = E_f$. Then

$$A_1 + A_2 = U_{fI}(P_I + P_I')V_{I0}, \tag{3.83}$$

where

$$P_I + P_I' = 2E_I/(E_0^2 - E_I^2). \tag{3.84}$$

The result (3.84) can be cast in another form that exhibits its covariant character. In attempting to obtain the form for the Feynman propagator, we should remember that the ultimate expression should be characteristic of the scalar nature of the particle (that is, not of its interactions) and be determined by, for example, its free-particle wave equation. Although we are considering electromagnetic phenomena, we can consider some special process convenient for the determination of the propagator for the intermediate state. We can, for example, consider the interaction V to be some (external) scattering potential (say, produced by some very heavy particle) such that the incident π^- suffers a change in the direction of its momentum but not in its magnitude. The interaction U can be a photon-emission vertex, and let us assume that the energies of the photons are infinitesimally small; then, even for the non-covariant diagram on the right in Figure 3.6, the photon momentum and energy in the intermediate (and final) state are negligible. Dropping, for simplicity, the subscript I for the intermediate state, we have $E_I^2/c^2 = -p^2 + \boldsymbol{p}^2$, where $p^2 \equiv p_\mu p_\mu$. Further, $E_0^2/c^2 = m^2c^2 + \boldsymbol{p}_0^2 = m^2c^2 + \boldsymbol{p}^2$. Thus we obtain[24]

$$(E_0^2 - E_I^2)/c^2 = p^2 + m^2c^2 \quad (p^2 \equiv p_\mu p_\mu). \tag{3.85}$$

We see that the resulting expression is an invariant. Note also that, although $p_\mu = (iE/c, \boldsymbol{p})$ for the intermediate state, $p^2 = p_\mu p_\mu \neq -m^2c^2$. The other factor $(2E_I)$ in the propagator (3.84) is of less significance but has some meaning in terms of the transition to the covariant formulation. In the covariant theory, a different normalization convention is more appropriate. The conservation of probability equation (3.70) requires that j_μ be a four-vector function of x_μ, and this is accomplished by adopting a normalization convention[25] $\int |u|^2 d^3\boldsymbol{r} = 2E$ (rather than unity) for free-particle states. This makes the probability density ρ [see also Equation (3.64)] the 0-component of a four-vector. Since u_I always appears twice in the total amplitude, this explains the occurrence of the factor $2E_I$. The essential factor in the covariant Feynman propagator for scalar particles is then

$$P_F(\text{scalar}) = (p^2 + m^2c^2)^{-1}. \tag{3.86}$$

---

[24]Often in the literature and in texts the propagator for scalar particles is given as $(p^2 - m^2c^2)^{-1}$. The sign difference has to do simply with the convention (metric) chosen for the four-dimensional scalar product. Again, the reader will have to be aware of these various sign conventions that are employed. Also, units with c (and $\hbar$) equal to unity are almost always used in this subject.

[25]Taking $2E$ rather than E just happens to be more convenient.

An alternative simplified approach[26] to obtaining P_F is illuminating in a different way. The coupled field equation (3.79) can be expressed in the form

$$(\Box^2 - \kappa^2)\phi = -S, \tag{3.87}$$

where the right-hand side can be considered a source term. It is written with a minus sign for convenience [compare Equation (2.71)]. Both ϕ and S are functions of x_μ and their Fourier transforms can be introduced. With f standing for either ϕ or S,

$$f(x_\mu) = \int f(k_\mu) e^{ik_\nu x_\nu} d^4k, \tag{3.88}$$

$$f(k_\mu) = (2\pi)^{-4} \int f(x_\mu) e^{-ik_\nu x_\nu} d^4x. \tag{3.89}$$

Substituting into the equation (3.87), we obtain the inhomogeneous solution ($k^2 = k_\mu k_\mu$)

$$\phi(x_\mu) = \int (k^2 + \kappa^2)^{-1} S(k_\mu) e^{ik_\nu x_\nu} d^4k', \tag{3.90}$$

which yields the factor $(k^2 + \kappa^2)^{-1}$ as in the result (3.86). The spacetime form of the propagator can be obtained by writing $S(k_\mu)$ in the terms of $S(x_\mu)$ by means of the relation (3.89). We then have

$$\phi(x_\mu) = \int D(x_\mu - x'_\mu) S(x'_\mu) d^4x', \tag{3.91}$$

where

$$D(x_\mu - x'_\mu) = \int \frac{e^{ik_\mu(x_\mu - x'_\mu)}}{k^2 + \kappa^2} \frac{d^4k}{(2\pi)^4} \tag{3.92}$$

is the propagator in the position representation. The form (3.91) exhibits how ϕ depends on S at other spacetime points.

3.4.3 Spin-$\frac{1}{2}$ Interactions

The general theory of the interactions of relativistic spin-$\frac{1}{2}$ particles with the electromagnetic field (photons) is what is commonly known as quantum electrodynamics. The spin-$\frac{1}{2}$ case is much more important than that for spin-0 because nature's smallest-mass charged particles, electrons and positrons, are of this type. Unfortunately, the theory is, on the whole, more complicated than that for spin-0 which we have just outlined. Thus, in imposing brevity, we shall have to be quite superficial and cannot really formulate the subject to provide the foundation necessary to perform extensive calculations. Nevertheless, certain basic characteristics of the theory can be seen without getting deeply into a more substantive exposition.

Basically, the content of the theory is determined by the form of the one-particle relativistic wave equation for the electron and positron, that is, the Dirac equation. In seeking this equation, Dirac was guided very much by a feeling that it should

---

[26]We are, again, relying heavily on R. P. Feynman, *Theory of Fundamental Processes*, New York: W. A. Benjamin, Inc., 1962.

have a form similar to the general time-dependent Schrödinger equation (3.1). Like the non-relativistic equation—and not like the Klein-Gordon equation—it should be first-order in the time derivative $\partial_t = \partial/\partial t$. But, since covariant equations have spatial derivatives ($\partial_j = \partial/\partial x_j$; $j = 1, 2, 3$) appearing in the same manner as time derivatives, it would seem that only the first derivatives ∂_j might be expected. The simplest such equation of this form is

$$(1/c)\partial_t\psi + \alpha_j\partial_j\psi + i\kappa\beta\psi = 0; \tag{3.93}$$

here α_j and β are numerical coefficients and κ ($= mc/\hbar$) is the basic inverse-length parameter characteristic of the particle whose state is described by $\psi(\mathbf{r}, t)$. The factor $1/c$ in the first term is natural in order to give the same dimensions as the second and third terms, if β and the three α_j are to be dimensionless. The factor i in the third term is arbitrary (but convenient) and could, alternatively, be incorporated into β.

Equation (3.93) is the free-particle Dirac equation, and it remains to determine the nature of β and the α_j. The wave function ψ is, moreover, allowed to have more than one component in order, for example, to describe particle spin; that is,

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix}. \tag{3.94}$$

In fact, the number of components has to be four if (for spin-$\frac{1}{2}$) two spin substates ($s_z = \pm 1/2$) are to be allowed as well as the two types of charge ($e^{\pm}$) or energy or antiparticle (hole) states in the theory. the four components are also called for by the resulting requirements on the coefficients β and α_j; these have to be *matrices* and at least 4×4. The form (3.93) chosen for the relativistic wave equation inherently implies an application to spin-$\frac{1}{2}$ particles with finite mass. The matrices β and α_j are determined by requiring that ψ also satisfy the Klein-Gordon equation (3.57), which is essentially a consequence of the relativistic relation $p_\mu p_\mu = -m^2c^2$ required also for free particles of spin-$\frac{1}{2}$. An equation of this type can be obtained if we operate with $(1/c)\partial_t - \alpha_l\partial_l - i\kappa\beta$ from the left on the Dirac equation (3.93). The coefficients and signs in this operator are fixed by the requirement that there be no terms in ∂_t and $\partial_t\partial_j$. Comparison with Equation (3.57) then yields the following requirements[27] on β and the α_j:

$$\tfrac{1}{2}(\alpha_j\alpha_l + \alpha_l\alpha_j) = \delta_{jl},$$

$$\beta\alpha_j + \alpha_j\beta = 0, \tag{3.95}$$

$$\beta^2 = I \quad \text{(unit matrix } I_{jl} = \delta_{jl}\text{)}.$$

It is possible to obtain solutions to these equations for β and the α_j in terms of the 2×2 Pauli spin matrices and the 2×2 identity matrix. However, instead of employing such explicit expressions, it is better not to indicate a particular representation, and

---

[27]The first of the relations (3.95), obtained through consideration of the term in $\partial_j\partial_l$, results in writing the coefficient of $\partial_j\partial_l$ ($= \partial_l\partial_j$) in the most general form to include all terms in the dummy indices j and l.

instead to regard the commutation relations (3.95) as fundamental and make use of identities derived directly from them. The β and α_j must be, in general, non-commuting operators or matrices as we see, in particular, from the second of the identities (3.95). Rather than employing these identities, it is, in fact, preferable to cast them—and the Dirac equation (3.93)—into four-dimensional covariant form; this is easily done.

Because of the last of the identities (3.95), if we operate from the left on the Dirac equation with β, we eliminate the matrix as a factor in the third term. The equation is then

$$\beta(1/c)\partial_t\psi + \beta\alpha_j\partial_j\psi + i\kappa\psi = 0, \tag{3.96}$$

and, in terms of $x_\mu = (ict, x_1, x_2, x_3)$ and ∂_μ, we see that the equation simplifies if we define

$$\begin{aligned}
\gamma_0 &= \beta, \\
\gamma_j &= i\alpha_j\beta = -i\beta\alpha_j,
\end{aligned} \tag{3.97}$$

and $\gamma_\mu = (\gamma_0, \gamma_1, \gamma_2, \gamma_3)$. The equation (3.96) then becomes

$$(\gamma_\mu\partial_\mu + \kappa)\psi = 0, \tag{3.98}$$

or, in terms of $p_\mu = -i\hbar\partial_\mu$,

$$(\not{p} - imc)\psi = 0, \tag{3.99}$$

where

$$\not{p} = \gamma_\mu p_\mu. \tag{3.100}$$

The covariant form (3.99) of the Dirac equation[28] is extremely simple. Moreover, the new Dirac matrices (3.97) satisfy the very compact identity

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\delta_{\mu\nu}, \tag{3.101}$$

which replaces the three identities (3.95). The identity (3.101) is easily obtained by operating on the first of the relations (3.95) from the right and left with β.

The notation $\not{p}$, read "p dagger" or "p slash," was introduced by Feynman. For any four-vector, we define

$$\not{B} = \gamma_\mu B_\mu. \tag{3.102}$$

With A_μ and B_μ being four-vectors, the identity (3.101) yields the useful relation

$$\not{A}\not{B} + \not{B}\not{A} = 2A_\mu B_\mu = 2A \cdot B. \tag{3.103}$$

There are many other identities involving Dirac matrices, and often the relation (3.103) is employed in proving the more complicated ones.

---

[28]The equation appears in this form in some textbooks, but more often it is exhibited without the factor i multiplying mc. The difference has to do with the convention chosen for the metric (see, also, Footnote 24), which we avoid throughout by taking imaginary time components of four-vectors, etc. Our subsequent expressions, such as the Feynman propagator for spin-$\frac{1}{2}$ particles, will, as a result of our notation, also contain the factor i.

To include the effects of electromagnetic interactions, the replacement $\partial_\mu \rightarrow \partial_\mu - (iq/\hbar c)A_\mu$ or

$$\not{p} \rightarrow \not{p} - (q/c)\not{A} \tag{3.104}$$

can be made in the free-particle Dirac equation (3.99). The equation is then

$$(\not{p} - imc)\psi = (q/c)\not{A}\psi = (q/c)\gamma_\mu A_\mu \psi. \tag{3.105}$$

Immediately we see two important characteristics of the covariant theory of electron-photon interactions. The source or coupling is of the γ_μ type—one of the fundamental kinds of coupling in covariant perturbation theory. In covariant QED for spin-$\frac{1}{2}$ particles, there is just this one type of coupling. This is to be compared with the non-relativistic formulation, which must include *three* types of coupling [see Equations (3.12) and (3.13) and Figures 3.2 and 3.4] associated with the $A \cdot p$, $A \cdot A$, and $\mu \cdot \text{curl } A$ terms in the interaction Hamiltonian. The single coupling term in the covariant theory accounts for all of these interaction terms in the non-relativistic theory (even the magnetic moment coupling!). As a result, the number of diagrams representing perturbations in a covariant formulation is significantly reduced for the spin-$\frac{1}{2}$ case. The only vertex is one involving two (charged) particle lines and one photon line—as in Figure 3.2. As we have seen, basically the occurrence of only this type of coupling is a result of the inherent simplicity of the Dirac equation, being linear in the momentum (or gradient) operator.

The other important factor in covariant perturbation theory that can be inferred from Equation (3.105) is the Feynman propagator for spin-$\frac{1}{2}$ particles. Comparison of Equation (3.105) with the corresponding field-source equation (3.87) for scalar particles suggests a propagator of the form

$$P_F(\text{spin-}\tfrac{1}{2}) = (\not{p} - imc)^{-1}. \tag{3.106}$$

Making use of the identity (3.103), this can also be written

$$\frac{1}{\not{p} - imc} \cdot \frac{\not{p} + imc}{\not{p} + imc} = \frac{\not{p} + imc}{\not{p}^2 + m^2c^2} = \frac{\not{p} + imc}{p^2 + m^2c^2}, \tag{3.107}$$

where, as in the scalar formulas, $p^2 = p_\mu p_\mu$. Again, the denominator is not zero because $p^2 \neq -m^2c^2$ in intermediate states.

For both the treatment of scalar (spin-0) QED in Section 4.2 and of spinor (spin-$\frac{1}{2}$) QED in Section 4.3, we have given only a sketchy outline. Not only are the derivations not rigorous, for example, those to obtain P_F (spin-0) and P_F (spin-$\frac{1}{2}$), but we have not derived the multiplying constants. This is not important; there are other ways of obtaining the factors. For example, we could compute some formula for which there is also a classical or non-relativistic derivation and compare the results to determine the factor. Of course, there is a detailed systematic method for obtaining the results and prescriptions in covariant perturbation theory. This takes much more time. However, we have given a fairly complete and self-contained treatment of the non-relativistic theory. Perhaps a comparison of this formulation with our more sketchy outline of the relativistic theory will help in understanding the latter. We shall be making some applications of the results derived here and that should also help in the understanding.

3.4.4 Invariant Transition Rate

Although a very general result, the Fermi Golden Rule formula (3.43) is inappropriate in that form for applications involving covariant perturbation theory. We can rewrite it for this purpose, however, with the resulting expression containing factors that are manifestly Lorentz invariant. The most important feature of the relativistic form of the transition rate formula is the invariant nature of the amplitude for the process. To distinguish it from the non-covariant amplitude (M_{f0}) we use the notation $\mathcal{M}_{f0}$ for the quantity. All of the physics of a process is contained in the corresponding $\mathcal{M}_{f0}$; the rest of the factors in the transition rate formula are kinematic in nature.

One characteristic of a relativistic formulation that is different concerns the normalization convention for free-particle wave functions. It is no longer appropriate[29] to take $|u|^2 = 1$; instead we adopt

$$|u|^2 = 2E \tag{3.108}$$

(some textbooks take $2E/m$ or E/m). This type of normalization results naturally from a relativistic formulation, as can be seen from formula (3.64), for example, and has a simple explanantion or interpretation. To provide an invariant probability (volume integral of $|u|^2$), Lorentz-Fitzgerald contraction along the direction of motion then compensates the factor E. Thus, with our modified normalization we make the replacement

$$\left|M_{f0}\right|^2 = \frac{\left|\mathcal{M}_{f0}\right|^2}{\left(\prod_{(f)} 2E\right)\left(\prod_{(0)} 2E\right)}; \tag{3.109}$$

here the products are over the $2E$ factors for all the incoming (0) and outgoing (f) particles. There would also be normalization factors associated with intermediate states, but these can be assumed to be contained within $\mathcal{M}_{f0}$. These factors could be incorporated within the Feynman propagators, for example, since for every pair of intermediate states, there is a P_F.

There are two other modifications in the non-covariant formula (3.43) that are to be made. The transition matrix element M_{f0} always yields a (total) momentum-conservation δ-function even when multiple perturbations are involved, arising from an integral of the type $\int e^{i\mathbf{k}\cdot\mathbf{r}} d^3\mathbf{r} = (2\pi)^3\delta^{(3)}(\mathbf{k})$, where $\mathbf{k}$ is the total momentum change. We extract this ubiquitous factor from M_{f0} and combine it with the energy-conservation δ-function:

$$(2\pi)^3\delta^{(3)}\left(\sum_{(f)}\mathbf{p} - \sum_{(0)}\mathbf{p}\right)\delta\left(\sum_{(f)}E - \sum_{(0)}E\right) = (2\pi)^3\delta^{(4)}\left(\sum_{(f)}p_\mu - \sum_{(0)}p_\mu\right), \tag{3.110}$$

with the four-dimensional covariant δ-function now expressing energy and momentum conservation.

---

[29]In this subsection, we simplify the algebra by setting $c = 1$ and $\hbar = 1$; this provides an easier comparison with treatments in modern textbooks. We also take a unit normalization volume.

The second additional modification of the non-covariant transition rate formula is, like the factor (3.110), essentially kinematic. It concerns the implied factor, eventually integrated, for the differential number of states[30] for the (N) outgoing particles:

$$dN_f = \prod_{(f)}^{N} d^3\boldsymbol{p}/(2\pi)^3. \tag{3.111}$$

Instead of counting the states of N particles, momentum conservation could be applied and $N-1$ factors of $d^3\boldsymbol{p}/(2\pi)^3$ could be taken, and the momentum-space δ-function [times $(2\pi)^3$] could be omitted. However, it is better to express results in this more symmetric way. It allows the employment of the covariant δ-function (3.110) and also allows the N factors $2E$ in the denominator of Equation (3.109) to be combined with the $d^3\boldsymbol{p}$ in product (3.111). This ratio $d^3\boldsymbol{p}/E$ is an invariant [see Equation (1.47)]; it can also be written in another form. With $p^2 \equiv p_\mu p_\mu = \boldsymbol{p}^2 - E^2$ (= inv.), we can write $p^2 + m^2 = (E_p + E)(E_p - E)$, where E_p is the positive square root of $\boldsymbol{p}^2 + m^2$. Then, if it is understood that only the positive values of E will be included in integrations, a factor $\delta(E - E_p)dE$ can be inserted to multiply each $d^3\boldsymbol{p}$. But, with the implied inclusion of only positive energies, $\delta(E - E_p) = 2E\delta(p^2 + m^2)$. Then

$$d^3\boldsymbol{p}/2E = \delta(p^2 + m^2)\, d^4p, \tag{3.112}$$

which is manifestly invariant.

The transition rate can now be expressed in the desired covariant form:

$$\frac{\Delta W}{\Delta t} = (2\pi)^4 \frac{\left|\mathcal{M}_{f0}\right|^2}{\prod_{(0)}(2E)}\, \rho, \tag{3.113}$$

where

$$\rho = \left(\prod_{(f)} 2\pi\delta(p^2 + m^2)\frac{d^4p}{(2\pi)^4}\right)\delta^{(4)}\left(\sum_{(f)} p_\mu - \sum_{(0)} p_\mu\right) \tag{3.114}$$

is the invariant phase space density. The factors of $2E$ for the incident particles in the denominator of Equation (3.113) are kinematic in nature. An important feature of this result is, as already emphasized, the invariant nature of the matrix element $\mathcal{M}_{f0}$. In fact, in some applications, it can even be possible to make an educated guess as to its form from considerations of invariance. There are many applications of the transition rate formula, with the physics of particular processes being contained in $\mathcal{M}_{f0}$. In particular, it is employed to compute lifetimes and cross sections for processes.

---

[30]If the particles have spin and we are not interested in their polarization states, there will also be a sum over these spin substates. Also, if, as is usually the case in the application considered, the incident particles are unpolarized, there would be an average over these polarizations. For simplicity in the notation, we are omitting spin summations and averages from the transition rate formula.

3.5 SOFT-PHOTON EMISSION

When the photon energy and momentum are small, that is, when the photon is "soft," some very general and useful results can be derived. In this section, we give quantum-mechanical derivations of certain formulas applicable in the non-relativistic limit and in the general case. All of the expressions are valid only in the Born approximation, and they will be compared with corresponding results derived through classical electrodynamics. The formulas are identical to the classical ones, but it turns out that the non-relativistic expressions derived quantum mechanically have a more general validity, being applicable even away from the soft-photon limit.

3.5.1 Non-Relativistic Limit

Consider the motion of a charge in the presence of some perturbation V that does not involve the spin of the particle.[31] The charge also feels a photon-emission perturbation, the lowest-order type being the interaction Hamiltonian (3.20) designated H_q'. The charge always experiences the perturbation H_q'; that is, it is always trying to produce photons. However, as mentioned earlier in this chapter, an isolated charge without internal structure cannot do this because of energy conservation. The combined action of the two perturbations V and H_q' does allow the phenomenon of photon emission to take place. The total perturbation Hamiltonian is then

$$H' = V + H_q', \tag{3.115}$$

and this perturbation can account for a variety of processes. The non-electromagnetic part V could represent various mechanisms. For example, it could be some scattering potential, or it could even be some interaction that causes the creation of a charge as in β-decay. In the latter process, a charge of opposite sign must be created (as in neutron decay: $n \rightarrow p + e^- + \overline{v}_e$), or the charge on a proton must be transferred to a (positive) electron as in positron-producing β-decay. It is always the electromagnetic interactions with the e^+ or e^- that are important for photon production because of their small mass. Further, regarding the application in β-decay, the e^- or e^+ involved must be non-relativistic for the applicability of the formulation in this subsection. There are such β-decays, an example being the decay of tritium (H^3), for which the maximum e^- energy is about 18.7 keV.

We are interested in this section in evaluating the probability that a soft photon of energy within $\hbar d\omega$ accompanies some radiationless process, the photon emission being just a small perturbation on the rest of the overall process. Both interactions V and H_q' are treated as perturbations; that is, in lowest order, they are considered to "act once." This is always a valid assumption for the electromagnetic perturbation

---

[31]We ignore spin to simplify the formulation. This assumption is really not necessary, and, if spin interactions are involved, there is no difficulty in modifying the equations by adding the spin coordinate to a total description of the particle state: $\psi_{\text{tot}} = \psi(r, t)\psi_{\text{spin}}$. If there are no spin interactions, the orthogonality of the spin eigenfunctions would simply tell us that the spin coordinate remains unchanged in the overall process and can be ignored. In the following subsection (5.2), we consider a problem where spin plays a major role.
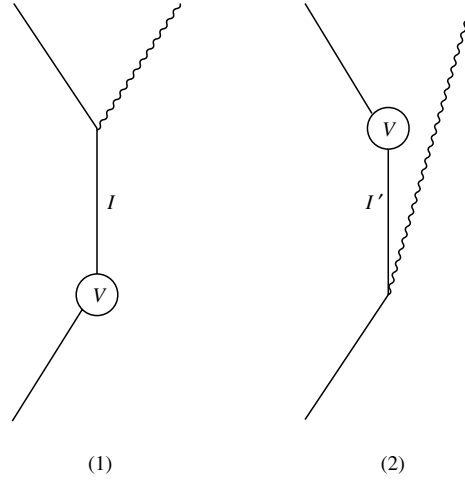
(1)   (2)

Figure 3.7 Diagrams representing the combined action of two perturbations, one being some
coupling or perturbation V and the second representing an electromagnetic cou-
pling, in this case associated with the production of an outgoing photon.

H'_q, since the interactions is fundamentally weak. The smallness of the fine structure
constant guarantees this: $\alpha \approx 1/137$. Treating V as a perturbation that acts once is a
more restricting assumption that generally confines the application to the domain of
the Born approximation. The Born approximation would require the energy of the
charge to be sufficiently large if, for example, V is a Coulomb scattering potential.
Specifically, in Coulomb scattering where the charges are ze and $z'e$, the criterion
for Born approximation would be simply[32] $zz'e^2/\lambda \ll mv^2$, where $\lambda = \hbar/mv$ is
the de Broglie wavelength. The requirement is then $\beta = v/c \gg zz'e^2/\hbar c = zz'\alpha$
or that the particle energy is $E \gg (zz')^2 \mathrm{Ry}$.

   With both V and H'_q acting once, the process can be represented in terms of the
diagrams in Figure 3.7. In our non-covariant formulation, there would also be pair
production-annihilation diagrams (see Figure 3.6), but these give a negligible contri-
bution in the non-relativistic limit (see previous section). The effective perturbation
Hamiltonian matrix element for the process is [see Equation (3.44)]

$$H'_{f0} = \sum_I \frac{\langle f|H'_q|I\rangle \langle I|V|0\rangle}{E_0 - E_I} + \sum_{I'} \frac{\langle f|V|I'\rangle \langle I'|H'_q|0\rangle}{E_0 - E_{I'}}. \qquad (3.116)$$

In the evaluation of matrix elements, let us, for convenience, set the normalization
volume L^3 equal to unity in all subsequent formulae; factors involving L always
cancel in final expressions for probabilities, cross sections, etc. For outgoing pho-

---

[32]This is a requirement that the characteristic scattering potential is small compared with the particle
kinetic energy. The probability current associated with the amplitude of the scattered wave would, as a
result, be small compared with that of the wave incident on the scattering center. The Born approximation
is essentially that of a perturbation or iteration procedure.

tons (wave vector $\boldsymbol{k}$ and polarization $\boldsymbol{\varepsilon}$), the matrix element associated with the photon-emission vertex (between plane-wave states a and b) will be given by [see Equation (3.20)]

$$\langle b|H_q'|a\rangle = (2\pi\hbar^3/\omega)^{1/2}(q/m)(\boldsymbol{\varepsilon}\cdot\boldsymbol{k}_a)\int e^{i(\boldsymbol{k}_a-\boldsymbol{k}_b-\boldsymbol{k})\cdot\boldsymbol{r}}d^3\boldsymbol{r}, \tag{3.117}$$

where $\boldsymbol{k}_a$ and $\boldsymbol{k}_b$ are the wave vectors of the charged particle. The integral is just a three-dimensional, momentum-conservation, Dirac δ-function, which can also be expressed as a Kronecker δ:

$$\int e^{i(\boldsymbol{k}_a-\boldsymbol{k}_b-\boldsymbol{k})\cdot\boldsymbol{r}}d^3r = (2\pi)^3\delta^{(3)}(\boldsymbol{k}_a-\boldsymbol{k}_b-\boldsymbol{k}) = \delta_{\boldsymbol{k}_a,\boldsymbol{k}_b+\boldsymbol{k}}. \tag{3.118}$$

The Dirac or Kronecker δ-functions are eventually used in the intermediate-state summations in Equation (3.116). Leaving out the trivial δ-function factor, the matrix element (3.117) can be written very simply:

$$\langle b|H_q'|a\rangle = (2\pi\hbar^3/\omega)^{1/2}q(\boldsymbol{\varepsilon}\cdot\boldsymbol{v}_a), \tag{3.119}$$

where $\boldsymbol{v}_a$ is the velocity of the charge in state a. The matrix element of V does not involve the photon and is only a function of the charge wave vectors:

$$\langle d|V|c\rangle \equiv V_{dc} = V_{dc}(\boldsymbol{k}_d,\boldsymbol{k}_c). \tag{3.120}$$

Further, we allow the source of the potential V to have "internal structure" such that it can possess an excitation energy χ. For example, it could be an atom or it could be a nucleus when the potential V causing a β-decay.

The energies of the various states are given by

$$\begin{aligned}
E_0 &= (\hbar k_0)^2/2m + \chi_0, \\
E_f &= (\hbar k_f)^2/2m + \hbar ck + \chi_f, \\
E_I &= (\hbar k_I)^2/2m + \chi_f, \\
E_{I'} &= (\hbar k_{I'})^2/2m + \hbar ck.
\end{aligned} \tag{3.121}$$

Since $E_0 = E_f$, there is also the relation

$$(\hbar k_0)^2/2m + \chi_0 = (\hbar k_f)^2/2m + \hbar ck + \chi_f. \tag{3.122}$$

Moreover, because of momentum conservation at the photon-emission vertices,

$$\boldsymbol{k}_I = \boldsymbol{k}_f + \boldsymbol{k}, \quad \boldsymbol{k}_{I'} = \boldsymbol{k}_0 - \boldsymbol{k}. \tag{3.123}$$

Making use of the identities (3.122) and (3.123), we obtain the energy denominators

$$\begin{aligned}
E_0 - E_I &= \hbar ck - \hbar^2\boldsymbol{k}\cdot\boldsymbol{k}_f/m - \hbar^2k^2/2m, \\
E_0 - E_{I'} &= -\hbar ck + \hbar^2\boldsymbol{k}\cdot\boldsymbol{k}_0/m - \hbar^2k^2/2m.
\end{aligned} \tag{3.124}$$

We now make the assumption that the photon energy ($\hbar ck$) is of the order of or less than that of the charged particle [say, $E_\alpha = (\hbar k_\alpha)^2/2m$]; then, [see also Equation (3.46)]

$$k/k_\alpha \lesssim \hbar k_\alpha/mc = v_\alpha/c = \beta_\alpha \ll 1. \tag{3.125}$$

In other words, for non-relativistic particles, when the photon and particle energies are comparable, the photon momentum is small. Under these conditions, the photons are always soft—at least as regards their momenta. Also, as we see easily, the three terms on the right-hand side of each of the Equations (3.124) are successively smaller in the ratio β_α.

There are two especially simple and important applications of the above results. The first is to the problem of particle production such as in β-decay. We ask for the probability that a photon (energy within $\hbar d\omega$) accompanies the production of a particle of given velocity $\beta_f c$. Except for summing over photon states $dN = d^3k/(2\pi)^3$, this differential probability is given by, in Born approximation,

$$dw = \frac{\left|H'_{f0}\right|^2}{\left|V_{f0}\right|^2}\, dN. \tag{3.126}$$

In this process, there is only diagram (1) of Figure 3.7, and the energy denominator is given by $\hbar ck$ in good approximation. Further, since $\boldsymbol{\varepsilon} \cdot \boldsymbol{k} = 0$, by the first equation (3.124), $\boldsymbol{\varepsilon} \cdot \boldsymbol{k}_I = \boldsymbol{\varepsilon} \cdot \boldsymbol{k}_f$, and the photon-emission matrix element is given by

$$\langle f|H'_q|I\rangle = (2\pi\hbar/\omega)^{1/2}q(\boldsymbol{\varepsilon} \cdot \boldsymbol{v}_f). \tag{3.127}$$

Because of the inequality (3.125) and the equation (3.120), the matrix element involving V is given by

$$\langle I|V|0\rangle = V_{I0} \approx V_{f0}. \tag{3.128}$$

Thus,

$$dw = \frac{\left|\langle f|H'_q|I\rangle\right|^2}{(\hbar ck)^2}dN, \tag{3.129}$$

with

$$dN = (2\pi)^{-3}k^2dk\, d\Omega, \tag{3.130}$$

in terms of the solid angle element Ω for the outgoing photon. In terms of $\omega\,(= kc)$ and $z = q/e$, we have, very simply,

$$dw = z^2(\alpha/4\pi^2)(d\omega/\omega)(\boldsymbol{\varepsilon} \cdot \boldsymbol{\beta}_f)^2 d\Omega, \tag{3.131}$$

where α is the fine-structure constant.

The expression (3.131) is the result of a quantum-mechanical derivation of a formula identical to the semi-classical expression (2.142). However, now we see that the expression is not restricted to the soft-photon limit when the Born approximation holds. We can also sum over polarizations for the outgoing photon. The two possible (linear) polarization states are perpendicular to $\boldsymbol{k}$. With $\boldsymbol{k}$ and $\boldsymbol{\beta}$ defining a plane, one of these polarization reference directions ($\boldsymbol{\varepsilon}_1$) can be taken perpendicular to this plane. The other polarization unit vector ($\boldsymbol{\varepsilon}_2$) is in the plane and perpendicular to $\boldsymbol{k}$ and gives the whole contribution (the reader can construct a simple diagram to indicate this). Thus,

$$\sum_{\text{pol}}(\boldsymbol{\varepsilon} \cdot \boldsymbol{\beta}_f)^2 = (\boldsymbol{\varepsilon}_2 \cdot \boldsymbol{\beta}_f)^2 = \beta_f^2 \sin^2\theta, \tag{3.132}$$

where θ is the angle between $\boldsymbol{k}$ and $\boldsymbol{\beta}_f$. Summed over polarizations, the angular distributions of the photon emission is then

$$dw = z^2(\alpha/4\pi^2)(d\omega/\omega)\beta_f^2 \sin^2\theta \, d\Omega. \tag{3.133}$$

Finally, integrating over angles for the outgoing photon,

$$dw = z^2(2\alpha/3\pi)\beta_f^2 d\omega/\omega. \tag{3.134}$$

The three formulas (3.131), (3.133), and (3.134) exhibit the characteristic "infrared divergence" factor $d\omega/\omega$, indicating an infinite probability of emitting infinitesimally soft photons. However, we emphasize again that, in the Born approximation, these non-relativistic expressions are not restricted to the soft-photon limit. That is, $\hbar\omega$ does not have to be small compared with the charged-particle kinetic energy. The general nature of these formulas has not been emphasized in the literature.

Let us turn to another important application of the above general formulation. This is to the problem of photon emission (bremsstrahlung) in the scattering of a charged particle. We need not specify the nature of the scattering potential, although the important application would be for Coulomb scattering (the photon emission would then be Coulomb bremsstrahlung). The Born approximation is assumed, however, since the scattering potential is treated as a perturbation that acts only once. Both diagrams in Figure 3.7 give contributions to the amplitude for the process, so that

$$H'_{f0} = H'_1 + H'_2, \tag{3.135}$$

and the energy denominators are given by

$$\begin{aligned} E_0 - E_I &\approx \hbar ck, \\ E_0 - E_{I'} &\approx -\hbar ck. \end{aligned} \tag{3.136}$$

The matrix elements of the electromagnetic perturbation (photon emission) have the form (3.127); that is, $\langle f|H'_q|I\rangle \propto \boldsymbol{\varepsilon} \cdot \boldsymbol{v}_f$ and $\langle I'|H'_q|0\rangle \propto \boldsymbol{\varepsilon} \cdot \boldsymbol{v}_0$. Both matrix elements of the scattering potential can be approximated by the (radiationless) expressions (3.128), because of the inequality (3.125). Because the two energy denominators (3.136) are equal but opposite in sign, we have the important result

$$H'_{f0} \propto \boldsymbol{\varepsilon} \cdot (\boldsymbol{v}_f - \boldsymbol{v}_0) = \boldsymbol{\varepsilon} \cdot \Delta\boldsymbol{v}. \tag{3.137}$$

The formulas for the photon-emission probabilities are then the same as the results (3.131), (3.133), and (3.134) with $\boldsymbol{\beta}_f$ replaced by $\Delta\boldsymbol{\beta}$. They are again identical to the semi-classical expressions (2.142) and (2.143). However, we now see that, basically because of the inequality (3.125), they are not restricted to the soft-photon limit—at least when the Born approximation holds. We return to make use of these formulas when we consider the general bremsstrahlung process in a later chapter.

3.5.2 Emission from Spin Transitions

Phenomena having to do with a particle's intrinsic magnetic moment associated with its spin have only a limited classical analog. This intrinsic moment is without spatial extent in a quantum-mechanical formulation, and the special case of spin-$\frac{1}{2}$ is particularly interesting and unique. The effects associated with photon coupling

to this moment are small, especially in the soft-photon limit. It is well to demonstrate this, if only as an illustration of the general theory developed in this chapter. Interest in the problem should be more than academic, however; it is important to demonstrate the magnitude of the effects in comparison with those involving photon interactions with particle "orbital" motion (interaction with the charge). There is, in fact, a particle in nature that has a magnetic moment but no charge. The neutron is electrically neutral but has a magnetic moment of magnitude[33] $\mu_n = -1.913\mu_M$, where $\mu_M = e\hbar/2M_pc$ is the nuclear Bohr magneton. The neutron moment—as well as most of the proton's moment—can be thought to be a result of the cloud of virtual pions that is associated with the nucleon. To regard the neutron's moment as point-like and "rigid" is thus an approximation, but to do so using the experimental magentic moment should take into account the virtual pion field. The approximation can be expected to be valid for photon energies much less than the pion rest energy (~ 100 MeV); this is in the non-relativistic domain for neutron motion. The other relevant application for spin-$\frac{1}{2}$ particles is to the electron, but photon interactions with that particle's charge are much greater.

For particles of spin-$\frac{1}{2}$, the spin angular momentum is given by $\boldsymbol{J}_{\text{spin}} = \hbar\boldsymbol{s}$ with $\boldsymbol{s} = \frac{1}{2}\boldsymbol{\sigma}$, $\boldsymbol{\sigma}$ being the set of 2×2 Pauli spin matrices: $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. The spin wave function has two components representing the amplitudes for the two possible values for the spin component in some specified direction. In terms of its spin magnitude μ_0, the spin magnetic moment is given by

$$\boldsymbol{\mu}_s = \mu_0\boldsymbol{\sigma}, \tag{3.138}$$

and in the (σ^2, σ_z) representation where the z-direction is chosen to describe one component of $\boldsymbol{\sigma}$, the three Pauli matrices are (see any book on quantum mechanics)

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.139}$$

The matrix σ_z is diagonal and the eigenvalues are ± 1. That is, the two eigenfunctions are

$$\Lambda_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Lambda_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{3.140}$$

and we can use the simplified notation

$$\Lambda_\pm = |\pm\rangle. \tag{3.141}$$

The Pauli matrices anticommute; that is, $\sigma_x\sigma_y + \sigma_y\sigma_x = 0$, etc. Also, $\sigma_x\sigma_y = i\sigma_z$, with other relations obtained from a cyclic permutation of x, y, z; in general, we can write

$$\boldsymbol{\sigma} \times \boldsymbol{\sigma} = 2i\boldsymbol{\sigma}. \tag{3.142}$$

Further, $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \boldsymbol{\sigma}/3 = 1$. The eigenfunctions are orthogonal: $\langle m'|m\rangle = \delta_{m'm}$. The operators σ_x, σ_y, and σ_z yield

$$\sigma_x |\pm\rangle = |\mp\rangle,$$
$$\sigma_y |\pm\rangle = \pm i |\mp\rangle, \tag{3.143}$$
$$\sigma_z |\pm\rangle = \pm |\pm\rangle.$$

---

[33]The minus sign means that the magnetic moment is opposite in direction to the spin.

Instead of the form (3.22) for the spin magnetic moment interaction (with a photon) Hamiltonian, we can rewrite it more simply as

$$S = i(2\pi\hbar\omega)^{1/2}\boldsymbol{\mu}_s \cdot \boldsymbol{\varepsilon}' \tag{3.144}$$

where $\boldsymbol{\varepsilon}'$ is a unit polarization vector in the direction of the photon's magnetic field (that is, $\boldsymbol{\varepsilon}' = \boldsymbol{k} \times \boldsymbol{\varepsilon}/k$, where $\boldsymbol{\varepsilon}$ is the polarization vector for $\boldsymbol{A}$ and the electric field); $\boldsymbol{\varepsilon}'$, like $\boldsymbol{\varepsilon}$, has two fundamental polarization directions perpendicular to $\boldsymbol{k}$. As in Section 3.5.1, we consider the action of perturbations V and H'_μ, where now V is independent of particle spin, so that the total perturbation Hamiltonian is $H' = V + H'_\mu$. We allow both V and H'_μ to act once on a single particle with spin-$\frac{1}{2}$ and magnitude μ_0 of its magnetic moment (see Figure 3.7). The particle is initially unpolarized, so we average over initial spin-polarization states and sum over final spin polarizations; we also sum over polarizations for the outgoing photon. The photon emission probability will be determined by

$$\bar{S} = \frac{1}{2}\sum_{\text{pol}}\sum_{m'}\sum_{m}\left|\langle m'|\boldsymbol{\varepsilon}' \cdot \boldsymbol{\sigma}|m\rangle\right|^2 = \sum_{\text{pol}}(\varepsilon'^2_x + \varepsilon'^2_y + \varepsilon'^2_z) = 2. \tag{3.145}$$

In the evaluation of the photon-emission probability, we again employ the basic equation (3.126) in terms of the effective Hamiltonian (3.116) for the combined process. The diagrams are of the form in Figure 3.7. In the perturbation Hamiltonian (3.116), the energy denominators can be approximated by ($\kappa = mc/\hbar$)

$$E_0 - E_I = \hbar ck(1 - \boldsymbol{k} \cdot \boldsymbol{k}_f/\kappa k),$$
$$E_0 - E_{I'} = -\hbar ck(1 - \boldsymbol{k} \cdot \boldsymbol{k}_0/\kappa k), \tag{3.146}$$

in terms of the initial and final wave vectors ($\boldsymbol{k}_0$ and $\boldsymbol{k}_f$) of the scattered particle (with magnetic moment). The matrix elements of the scattering potential V are

$$\langle \boldsymbol{k}_I|V|\boldsymbol{k}_0\rangle = \langle \boldsymbol{k}_f + \boldsymbol{k}|V|\boldsymbol{k}_0\rangle,$$
$$\langle \boldsymbol{k}_f|V|\boldsymbol{k}_{I'}\rangle = \langle \boldsymbol{k}_f|V|\boldsymbol{k}_0 - \boldsymbol{k}\rangle, \tag{3.147}$$

which are the same in Born approximation. The combination of energy denominators is given by

$$\frac{1}{E_0 - E_I} + \frac{1}{E_0 - E_{I'}} \cong \frac{1}{\hbar ck}\left(1 + \frac{\boldsymbol{k} \cdot \boldsymbol{k}_f}{\kappa k} - 1 - \frac{\boldsymbol{k} \cdot \boldsymbol{k}_0}{\kappa k}\right)$$
$$= -\frac{\boldsymbol{k} \cdot \Delta\boldsymbol{k}_\mu}{mc^2k^2}, \tag{3.148}$$

where $\Delta\boldsymbol{k}_\mu = \boldsymbol{k}_0 - \boldsymbol{k}_f$ is the change in the wave vector of the scattered particle (having a magnetic moment). Employing the above results, including Equation (3.145), we obtain

$$dw_\mu = 2\mu_0^2(2\pi\hbar\omega)\left(\frac{\boldsymbol{k} \cdot \Delta\boldsymbol{k}_\mu}{mc^2k^2}\right)^2\frac{d^3\boldsymbol{k}}{(2\pi)^3}. \tag{3.149}$$

Setting $d^3\mathbf{k} = k^2 dk \, d\Omega$ and integrating over angles of emission, we have

$$dw_\mu = \frac{2}{3\pi} \frac{\mu_0^2}{m^2 c^5} (\Delta k_\mu)^2 \hbar\omega \, d\omega. \tag{3.150}$$

For the case of scattering of an electron ($\mu_0 = e\hbar/2mc$), the result can be written

$$dw_\mu = \tfrac{1}{4} (\hbar\omega/mc^2)^2 dw_q, \tag{3.151}$$

where dw_q is the corresponding emission probability associated with photon inter-actions with the charge [see Equation (3.134) with $\beta_f \to \Delta\beta$]. The emission from spin interactions is small for non-relativistic particles. Moreover, we note that there is no infrared divergence effect as $\omega \to 0$; in fact, $dw_\mu/d\omega \to 0$ in this limit. Finally, it should be mentioned that the results (3.149)–(3.151) are not confined to the soft-photon limit.

3.5.3 Relativistic Particles without Spin

Consideration of soft-photon emission associated with processes involving relativis-tic particles provides our first application of the methods of covariant QED outlined in Section 3.4. First, for scalar charged particles, we derive an expression for the photon-emission probability corresponding to the non-relativistic formulae (3.134) and (3.150). As in the derivation of those results, two perturbations are allowed to act on the particle (see Figure 3.7); one can correspond to, say, a scattering potential or some perturbation that creates the charge at high energy, while the other is the purely electromagnetic perturbation (photon-emission vertex). In our elementary sketch of covariant QED, we concerned ourselves with the establishment of the form of the perturbation amplitudes and Feynman propagators corresponding to the perturbation Hamiltonian and energy denominator in the non-covariant, non-relativisitic theory. Although the multiplying factors for these terms can be derived formally in a systematic exposition of covariant perturbation theory, we shall merely leave them as undetermined constants. This is not a serious problem; in fact, the undetermined factors can usually be established through a comparison with known results in some special case such as the non-relativistic limit or the classical limit.

The fundamental factors that are needed in a covariant perturbtion calculation are the Feynman propagator (3.86), $P_F(\text{scalar}) = (p^2 + m^2 c^2)^{-1}$, and the term associated with the photon-emission vertex. For such a vertex (see Figure 3.2), the matrix element between initial state a and final state b has the form (3.73) in the momentum representation:

$$M_{ba} = a_s q \varepsilon_\mu (p_a + p_b)_\mu, \tag{3.152}$$

where a_s is some numerical factor for this scalar-charge interaction; $p_{a\mu}$ and $p_{b\mu}$ are the four-momenta for the incoming and outgoing charge (q), and ε_μ is the unit polarization four-vector. The amplitude M_{ba} is, like P_F, a Lorentz invariant and is the analog of the non-relativistic formula (3.20) associated with the same type of single-photon vertex. With the combined interactions (3.152) plus the perturbation

V, the matrix element for the process between initial state 0 and final state f will be (see Figure 3.7)

$$
\begin{aligned}
M_{f0} = {} & a_s q \varepsilon_\mu (p_f + p_I)_\mu (p_I^2 + m^2 c^2)^{-1} V_{I0} \\
& + V_{fI'} (p_{I'}^2 + m^2 c^2)^{-1} a_s q \varepsilon_\mu (p_{I'} + p_0)_\mu .
\end{aligned} \tag{3.153}
$$

At the vertices, there is four-momentum conservation so that

$$
p_{I\mu} = p_{f\mu} + k_\mu , \qquad p_{I'\mu} = p_{0\mu} - k_\mu , \tag{3.154}
$$

k_μ referring to the photon. Since $p_f^2 = (p_\mu p_\mu)_f = p_0^2 = (p_\mu p_\mu)_0 = -m^2 c^2$ and $k_\mu k_\mu = 0$, the Feynman propagators for the two intermediate states are, very simply,

$$
P_I = \frac{1}{2(p_f \cdot k)}, \qquad P_{I'} = -\frac{1}{2(p_0 \cdot k)}, \tag{3.155}
$$

where the notation $(p \cdot k) = p_\mu k_\mu$ is employed for the four-dimensional scalar products. Further, because of the relations (3.154),

$$
\begin{aligned}
(p_f + p_I)_\mu &= (2p_f + k)_\mu , \\
(p_{I'} + p_0)_\mu &= (2p_0 - k)_\mu .
\end{aligned} \tag{3.156}
$$

Also, the Lorentz gauge condition

$$
\varepsilon_\mu k_\mu = 0 \tag{3.157}
$$

holds. In the soft-photon limit

$$
V_{I0} \approx V_{fI'} \approx V_{f0}, \tag{3.158}
$$

which is the matrix element in the radiationless problem. The matrix element for the combined process involving V and the (soft-) photon perturbation is then the invariant

$$
M_{f0} \approx a_s q V_{f0} \left[\frac{(\varepsilon \cdot p_f)}{(k \cdot p_f)} - \frac{(\varepsilon \cdot p_0)}{(k \cdot p_0)} \right]. \tag{3.159}
$$

The photon-emission probability is [see Equation (3.126)]

$$
dw = \frac{|M_{f0}|^2}{|V_{f0}|^2} dN', \tag{3.160}
$$

where dN' is now the invariant phase-space factor (3.113) or number of photon final states divided by $2E$. That is, with k designating the magnitude of the energy or momentum of the outgoing photon,[34]

$$
dN' \propto k^2 dk \, d\Omega / k = k \, dk \, d\Omega . \tag{3.161}
$$

---

[34]The appearance of the additional factor of $1/k$ could be understood in another way. Instead of employing the (more natural) invariant phase-space factor, the photon-emission invariant amplitude (3.152) could be employed with an additional factor $k^{-1/2}$. This perturbation amplitude would then be essentially a relativistic generalization of the corresponding non-relativistic interaction Hamiltonian (3.20). Alternately, the factor could be regarded as associated with the relativistic normalization [see Equation (3.108)] of the photon function.

If, in addition, a sum over photon final-state polarizations is performed, the probability (3.160) is given by an expression of the form

$$dw = Kq^2 \sum_{\text{pol}} \left| \frac{(\varepsilon \cdot p_f)}{(k \cdot p_f)} - \frac{(\varepsilon \cdot p_0)}{(k \cdot p_0)} \right|^2 k \, dk \, d\Omega, \qquad (3.162)$$

where K is a constant.

Although the photon polarization unit four-vector ε_μ has four components, only two contribute in the polarization sum. There is a covariant (Lorentz) transversality condition (3.158), but within this class of gauges there can be a (convenient) choice of gauge (see Section 2.1.1). The transformed vector potential $A'_\mu = A_\mu + \partial_\mu \chi$ could be used if χ satisfies $\Box^2 \chi = 0$. This is satisfied by, for example, the choice $\chi = a \exp(ik_\nu x_\nu)$, where $k_\nu = (ik, \boldsymbol{k})$ is the photon four-momentum. The transformed polarization four-vector would then be

$$\varepsilon'_\mu = \varepsilon_\mu + iak_\mu, \qquad (3.163)$$

and the choice $a = \varepsilon_0/k$ could be made so that $\varepsilon'_0 = 0$. We assume that this gauge is employed so that (dropping the primes) ε_μ has a vanishing time component, the relation (3.157) giving the usual transversality for the space part of ε_μ:

$$\varepsilon_\mu k_\mu = (\boldsymbol{\varepsilon} \cdot \boldsymbol{k}) = 0. \qquad (3.164)$$

For general $\boldsymbol{p}_f$ and $\boldsymbol{p}_0$, the expression (3.162) does not yield a simple formula on integrating over $d\Omega$ and summing over polarizations. Only for a frame with $\boldsymbol{p}_0 = 0$ (or $\boldsymbol{p}_f = 0$) does the result simplify. Designating either of these as $\boldsymbol{p}$ (with the other zero), we have

$$\sum_{\text{pol}} |\varepsilon \cdot p|^2 = \sum_{\text{pol}} (\boldsymbol{\varepsilon} \cdot \boldsymbol{p})^2 = \boldsymbol{p}^2 - (\boldsymbol{p} \cdot \boldsymbol{k})^2 / k^2, \qquad (3.165)$$

and

$$\sum_{\text{pol}} \left| \frac{(\varepsilon \cdot p)}{(k \cdot p)} \right|^2 = \frac{\beta^2}{k^2} \frac{1 - \cos^2 \theta}{(1 - \beta \cos \theta)^2}, \qquad (3.166)$$

in terms of $\beta \, (= v/c)$ for the outgoing charge and the angle θ between $\boldsymbol{p}$ and $\boldsymbol{k}$. This expression (3.166) gives the angular distribution $(dw/d\Omega)$ of the emission. Integrating over $d\Omega = 2\pi \sin \theta \, d\theta$, we get the emission probability in any direction:

$$dw = 2\pi Kq^2 \frac{dk}{k} \beta^2 \int_{-1}^{1} \frac{1 - x^2}{(1 - \beta x)^2} dx. \qquad (3.167)$$

The integral is elementary [in fact, it was already encountered in Equation (2.146)] and we obtain

$$dw = 4\pi Kq^2 \frac{dk}{k} \left(\frac{1}{\beta} \ln \frac{1 + \beta}{1 - \beta} - 2 \right). \qquad (3.168)$$

For $\beta \ll 1$, the parenthesis above is

$$\frac{1}{\beta} \ln \frac{1 + \beta}{1 - \beta} - 2 = \frac{2}{3} \beta^2 \left(1 + \frac{3}{5} \beta^2 + \cdots \right), \qquad (3.169)$$

which allows a comparison with the non-relativistic formula (3.134). For $q = ze$, we then have $K = (4\pi^2\hbar c)^{-1}$ and

$$dw = \frac{\alpha}{\pi}z^2\left(\frac{1}{\beta}\ln\frac{1+\beta}{1-\beta} - 2\right)\frac{dk}{k}, \tag{3.170}$$

which is identical to the result obtained in Equation (2.149). The formula exhibits the usual infrared divergence factor (dk/k). It is valid in the soft-photon limit, that is, for $\hbar ck \ll (\gamma - 1)mc^2$, *except* when $\beta \ll 1$, for which, as we have seen in Section 5.1, it holds for general photon energy.

3.5.4 Relativistic Spin-$\frac{1}{2}$ Particles

Here we wish to derive an expression for the probability that soft-photon emission accompanies some process involving relativistic particles of spin-$\frac{1}{2}$. This case is far more important than that for spin-0 considered in the previous subsection, since it applies to electrons. It will provide a first illustration of the techniques of modern QED and is perhaps the simplest example of methods in the theory. Employing the basic principles, the application is elementary, yielding a general formula that can be compared with that for spin-0 and with the classical expression.

In the covariant theory, the probability of a process will be proportional to the square of a matrix element $\mathcal{M}_{f0}$. This amplitude is a Lorentz invariant and is constructed from the one-particle Dirac wave functions ψ_0 and ψ_f for the initial and final states and a product (M) of factors associated with interaction vertices and Feynman propagators. In the momentum representation, the invariant amplitude is constructed from

$$\mathcal{M}_{f0} = \bar{u}_f M u_0, \tag{3.171}$$

where u_0 and u_f are the momentum-space Fourier amplitudes for the initial and final states, respectively. That is, u is related to ψ by

$$\psi(\mathbf{r}, t) = u\, e^{i(p\cdot x)}, \tag{3.172}$$

where, again, $(p \cdot x) = p_\mu x_\mu$. Both ψ and u are four-component (spinor) wave functions satisfying the free-particle Dirac equation in the position and momentum representation, respectively.

Now, however, we have to develop the basic theory a little further. For example, what, precisely, is the conjugate wave function $\bar{u}$? While ψ is the column wave function (3.94) with four components, $\bar{\psi}$ is *not* the row $\psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$, which is the Hermitian conjugate[35] wave function, but is defined by[36]

$$\bar{\psi} = \psi^\dagger\beta = \psi^\dagger\gamma_0; \tag{3.173}$$

---

[35]The Hermitian conjugate or adjoint of a matrix is formed by interchanging rows and columns and taking the complex conjugate of its elements: $(A^\dagger)_{jk} = (A_{kj})^*$. A matrix is Hermitian if it is equal to its Hermitian conjugate, that is, if $A^\dagger = A$. While β and the α matrices are Hermitian, of the γ matrices only $\gamma_0\,(= \beta)$ is Hermitian; γ_1, γ_2, and γ_3 are anti-Hermitian.

[36]The operation on the right in Equation (3.173) is a matrix multiplication of the single-row matrix $\psi^\dagger$ on $\beta\,(= \gamma_0)$ and yields another single-row matrix $(\bar{\psi})$. This can be easily seen from the matrix multiplication rule $(AB)_{kj} = A_{kl}B_{lj}$. Here $A_{kl} = 0$ unless $k = 1$ and $(AB)_{kj} = (AB)_{1j}$, that is, a single-row matrix.

also, of course, $\overline{u} = u^\dagger \gamma_0$. The reason for introducing the functions $\overline{\psi}$ and $\overline{u}$ (rather than $\psi^\dagger$ and $u^\dagger$) to form the invariant (3.171) can be seen through consideration of the continuity equation for probability density and current in the Dirac theory. To obtain this relation, we first derive the equations that $\psi^\dagger$ and $\overline{\psi}$ satisfy. These manipulations will make use of the Hermitian character of the α and β matrices: $\alpha_j^\dagger = \alpha_j$; $\beta^\dagger = \beta$. That β and $\boldsymbol{\alpha}$ are Hermitian is expected because of the Hermitian character of the (four-momentum) operators $i\partial_\mu$ in the Dirac equation (3.93) (multiplied by i). Alternatively, the character can be seen from the specific representation

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \tag{3.174}$$

where the σ_j are the three Pauli spin matrices (3.139). The representation (3.174), in terms of the three anti-commuting Pauli matrices and the identity matrix, satisfies the identities (3.95).

Returning now to the problem of the equations for $\psi^\dagger$ and $\overline{\psi}$, we take the Hermitian conjugate of the Dirac equation (3.93), making use of the Hermitian property of α_j and β:

$$(1/c)\partial_t \psi^\dagger + \partial_j \psi^\dagger \alpha_j - i\kappa \psi^\dagger \beta = 0. \tag{3.175}$$

Multiplying from the right by β and making use of the anti-commutation relation $\alpha_j \beta = -\beta \alpha_j$, we have, in terms of the definition (3.173) for $\overline{\psi}$,

$$(1/c)\partial_t \overline{\psi} - \partial_j \overline{\psi} \alpha_j - i\kappa \overline{\psi} \beta = 0. \tag{3.176}$$

Again multiplying on the right by β and using the definitions (3.97) and (3.100), there results an equation for $\overline{\psi}$ in covariant form:

$$\overline{\psi}(\not{p} + imc) = 0, \tag{3.177}$$

where the gradient ∂_μ in $\not{p}$ is meant to operate on the function $\overline{\psi}$ to the left. This equation is the free-particle Dirac equation for the conjugate function $\overline{\psi}$, and the sign difference in parentheses should be noted when comparison is made with the Dirac equation (3.99) for ψ. On the other hand, the conjugate momentum component amplitude $\overline{u}$ is given by

$$\overline{\psi} = \overline{u} e^{-i(p \cdot x)}, \tag{3.178}$$

with a sign difference in the exponent (as if $p_\mu \to -p_\mu$). That is, $\overline{u}_{(-p)}$ satisfies $\overline{u}_{(-p)}(\not{p} + imc)$ and $\overline{u}_p = \overline{u}$ satisfies

$$\overline{u}(\not{p} - imc) = 0, \tag{3.179}$$

the parenthesis having the same sign combination as in

$$(\not{p} - imc)u = 0. \tag{3.180}$$

The sign difference should also be noted in the terms in Equation (3.175) for $\psi^\dagger$ and in Equation (3.176) for $\overline{\psi}$, indicating how $\overline{\psi}$ is a more natural conjugate function than $\psi^\dagger$ in a covariant theory.

The continuity equation for probability density and current can be obtained by multiplying equation (3.93) on the left by $\psi^\dagger$ and adding the result to Equation (3.175) multiplied on the right by ψ. We then have

$$\frac{\partial}{\partial t}(\psi^\dagger \psi) + \frac{\partial}{\partial x_j}(c\psi^\dagger \alpha_j \psi) = 0, \tag{3.181}$$

which is a conservation equation with density $\rho = \psi^\dagger \psi$ and current density $\boldsymbol{j} = c\psi^\dagger \boldsymbol{\alpha} \psi$. However, in terms of the covariant Dirac matrices γ_μ, since $\overline{\psi} = \psi^\dagger \beta$ and so $\psi^\dagger = \overline{\psi}\beta$, while $\gamma_j = i\alpha_j \beta = -i\beta\alpha_j$, the continuity equation (3.181) can be written in covariant form in terms of $\overline{\psi}$, ψ, and the γ_μ:

$$\partial_\mu(\overline{\psi}\gamma_\mu \psi) = 0. \tag{3.182}$$

Thus, we see how, in a covariant theory, the function $\overline{\psi}$ is a more natural conjugate function than $\psi^\dagger$.

Finally, for the basic formulation, we derive a relation between the normalization for $\overline{u}u$ and that for $u^\dagger u$. If we write the Dirac equation (3.93) in terms of the amplitude u defined in Equation (3.172) and then multiply the equation from the left by $\overline{u}$ $(= u^\dagger \beta)$, we obtain

$$Eu^\dagger \beta u = cu^\dagger \beta \boldsymbol{\alpha} \cdot \boldsymbol{p}u + mc^2 u^\dagger u. \tag{3.183}$$

Taking the Hermitian conjugate of this equation, and making use of the anti-Hermitian character of the operator $\beta\boldsymbol{\alpha}$, there results an equation identical to the relation (3.183) except that the first term on the right has a minus sign. Adding Equation (3.183) then yields the result

$$\overline{u}u = (mc^2/E)u^\dagger u, \tag{3.184}$$

which will be referred back to later.

The free-particle Dirac equations (3.183) and (3.184) in the momentum representation are necessary to prove the important expression for the soft-photon emission probability associated with a process involving relativistic electrons. We again consider the combined action of a perturbation V and a photon-emission perturbation on a spin-$\frac{1}{2}$ charge (see Figure 3.7). The photon-emission vertex is associated with a coupling of the form $\not{\epsilon}$ [the $(q/c)\gamma_\mu A_\mu$ term in Equation (3.105)]. There are two types of intermediate states, just as in the spin-0 case in Section 5.3, for which the Feynman propagator is of the form (3.106) or (3.107). The matrix element for the process is given by the form (3.171) in terms of an invariant amplitude M. According to the Feynman rules for spin-$\frac{1}{2}$ charges, the amplitude corresponding to the spin-0 case (3.153) is now given by

$$M = \not{\epsilon}(\not{p}_I - imc)^{-1}V(I, 0) + V(f, I')(\not{p}_{I'} - imc)^{-1}\not{\epsilon}. \tag{3.185}$$

As in the relativistic spin-0 formulation, there is four-momentum conservation at the photon-emission vertex and $p_{I\mu}$ and $p_{I'\mu}$ are again given by the relations (3.154). The V-perturbation amplitudes are associated with particle "transitions" $0 \rightarrow I$

and $I' \to f$ as indicated in the simplified notation in Equation (3.185). With the propagators expressed in the form (3.107), the amplitude M is given by

$$M = \frac{\not{\varepsilon}(\not{p}_f + \not{k} + imc)}{2(k \cdot p_f)} V(I, 0) - V(f, I') \frac{(\not{p}_0 - \not{k} + imc)}{2(k \cdot p_0)} \not{\varepsilon}. \tag{3.186}$$

In the soft-photon limit, we neglect the terms $\not{k}$ in the numerators and we also approximate

$$V(I, 0) \approx V(f, 0),$$
$$V(f, I') \approx V(f, 0). \tag{3.187}$$

Employing the identity (3.103), we rewrite $\not{\varepsilon}\not{p}_f$ and $\not{p}_0\not{\varepsilon}$:

$$\not{\varepsilon}\not{p}_f = -\not{p}_f\not{\varepsilon} + 2(\varepsilon \cdot p_f),$$
$$\not{p}_0\not{\varepsilon} = -\not{\varepsilon}\not{p}_0 + 2(\varepsilon \cdot p_0). \tag{3.188}$$

Since $\bar{u}_f \not{p}_f = imc\bar{u}_f$ and $\not{p}_0 u_0 = imcu_0$, we have

$$\mathcal{M}_{f0} = \bar{u}_f M u_0 \approx \bar{u}_f V(f, 0) u_0 \left[\frac{(\varepsilon \cdot p_f)}{(k \cdot p_f)} - \frac{(\varepsilon \cdot p_0)}{(k \cdot p_0)} \right]. \tag{3.189}$$

This modified amplitude for the combined process has the same relation to that for the radiationless problem as in the spin-0 case. Thus, the photon-emission probability is identical to the formulas for that case. In other words, the angular distribution for the soft photons is again given by the form (3.166), and the probability for emission in all directions is given by the result (3.170). The establishment of the multiplying constant requires, as in the spin-0 case, a comparison with the non-relativistic formulas that were given a detailed systematic derivation.

Spin seems to be unimportant for emission in the soft-photon limit. It should be emphasized that in the covariant spin-$\frac{1}{2}$ formulation employed in this subsection, the spin effects are included in the single electromagnetic perturbation $\not{\varepsilon}$ (or γ_μ). That is, as we have emphasized near the end of Section 4.3, this perturbation includes the three perturbation terms $\boldsymbol{A} \cdot \boldsymbol{p}$, $\boldsymbol{A} \cdot \boldsymbol{A}$, and $\boldsymbol{\mu} \cdot \text{curl } \boldsymbol{A}$ in the non-covariant theory. The result that spin is unimportant in soft-photon emission involving relativistic (and non-relativistic) charges can perhaps be understood in terms of the result (3.150) for spin transitions as obtained in a non-covariant formulation. Finally, it should be emphasized again that all of these results have been derived in the Born approximation. The results will be employed later in various applications.

There is one feature of the soft-photon emission formula (3.170) for relativistic particles that should perhaps be noted at this point. In the limit of ultrarelativistic energies where $\beta \to 1$, the argument of the logarithm approaches the large number $(2E/mc^2)^2$. The emission probability then approaches

$$dw \to \frac{2z^2\alpha}{\pi} \left(\ln \frac{2E}{mc^2} - 1 \right) \frac{dk}{k}. \tag{3.190}$$

Further, if we integrate over photon energies or wave numbers from k_1 to k_2, the total photon-emission probability is approximately

$$w(k_1 < k < k_2) \to \frac{2z^2\alpha}{\pi} \ln \frac{E}{mc^2} \ln \frac{k_2}{k_1}. \tag{3.191}$$

Although $\alpha = 1/137$, the logarithmic factors can compensate to make the photon emission probability appreciable (that is, not small). This does not occur in the non-relativistic case where instead of the first logarithmic factor in Equation (3.191), there is a factor v^2/c^2 [see Equation (3.134)].

3.6 SPECIAL FEATURES OF ELECTROMAGNETIC PROCESSES

There are several characteristic of electromagnetic processes that can be described in a general way and that represent features that are of great importance in determining the details and qualitative aspects of various processes. Some of these, like radiative corrections and renormalization techniques, are really beyond the scope of this book for a substantive discussion; we only mention these fundamental and subtle topics. However, certain physical processes are very closely related and can be described in a unified way with the help of a property called *crossing symmetry*, which will be discussed later in this section. Also, for some kinds of processes there are special kinematic invariants that are useful in relating cross sections through crossing symmetry, and these will be introduced. Of great importance in categorizing processes is their "order" or characteristic magnitude of the associated cross sections. This can be done quite simply, and it is useful to be aware of the elementary ideas that allow a determination of the characteristic magnitudes of various cross sections.

3.6.1 "Order" of a Process

One of the basic properties of the electromagnetic coupling is that it is a fairly weak interaction—not *very* weak like that (called "weak," in fact) associated with neutrino processes, etc., but weak enough that a perturbation-theory approach is useful. We have discussed this aspect of the coupling in Section 3.5.1 in connection with soft-photon emission. In fact, treatment of that phenomenon yielded results that exhibit the characteristic strength of the coupling very clearly. We have seen [see Equations (3.134), (3.170), (3.190)] that the differential probability for soft-photon emission (energy within $\hbar\,d\omega$) to accompany some process involving a charged particle is of the form

$$dw \sim \begin{cases} \alpha\beta^2(d\omega/\omega) & \text{(NR)}, \\ \alpha\ln\gamma\,(d\omega/\omega) & \text{(ER)}, \end{cases} \tag{3.192}$$

in the non-relativistic (NR) and extreme-relativistic (ER) limits. Aside from the factor $d\omega/\omega$ and the kinematic factors β^2 and $\ln\gamma$, the probability is determined by the dimensionless coupling constant $\alpha = e^2/\hbar c$. The factor α can be considered the square of a dimensionless charge:

$$\bar{q} = e/(\hbar c)^{1/2} = \alpha^{1/2}. \tag{3.193}$$

The probability [Equation (3.192), for example] is the square of an amplitude M corresponding to some perturbation, in this case associated with a photon-emission vertex. For a single-photon vertex, the perturbation (interaction Hamiltonian or

Lagrangian) is proportional to the charge q (or the dimensionless $\overline{q}$), and the probability is

$$W = |M|^2 \propto \overline{q}^2 = \alpha. \tag{3.194}$$

A general relation can be written down for the total amplitude and probability associated with a purely electromagnetic process that can be represented by a collection of diagrams representing the actions of perturbations and the roles of propagators (or energy denominators). Let us now consider processes and diagrams that allow both charged particles and/or photons in intermediate states as well as in initial and final states; up to now we have regarded the photons to be "real" and occurring in intermediate states only if they were the same as in initial or final states. That is, we now consider processes that can correspond to "virtual" photons—photons that are only part of intermediate states. In this more general classification of electromagnetic processes the *order* of the various diagrams can be designated by the corresponding number of vertices they contain. This designation is sufficient in a covariant formulation for spin-$\frac{1}{2}$ charges for which there is only one kind of vertex, namely that representing the γ_μ interaction [see Equation (3.105)]. This is the vertex of Figure 3.2 (or 3.3) with one photon line and two fermion lines, and the strength of the perturbation is proportional to q (or $\overline{q}$). In a non-relativistic formulation, there are two types of vertices (Figures 3.2 and 3.4) associated with the interactions (3.20) and (3.21), respectively (H'_q and H'_{q^2}), as well as the magentic-moment interaction H'_μ given by the form (3.22), but this latter perturbation is usually not important [see Equation (3.25)]. In fact, as we have seen, for spin-0 charges a covariant formulation also introduces two types of vertices [see Equation (3.73)]. However, for these cases we find that the two-photon perturbation H'_{q^2} (and its relativistic counterpart) is proportional to q^2 or to $\overline{q}^2$, so it is higher order by the same factor $\alpha^{1/2}$ in its amplitude. That is, the two-photon vertex yields a factor α^2 in the squared amplitude—like two single-photon vertices.

In classifying the order of a general electromagnetic process, we can say that its probability is of the form

$$W \propto \alpha^n, \tag{3.195}$$

where

$$n = n_1 + 2n_2, \tag{3.196}$$

being the number (n_1) of single-photon vertices plus twice the number (n_2) of two-photon vertices associated with the process. Processes that are complex and require a larger number of interactions, or vertices in a diagrammatic representation, are less probable and have smaller rates. Usually the important parameter associated with some process is a cross section and, if n is its order, the cross section can be written

$$\sigma \sim \alpha^n l^2, \tag{3.197}$$

where l is some characteristic length. We can think of four such lengths: the (electron) Compton wavelength ($\Lambda = \hbar/mc$), the de Broglie wavelength ($\lambdabar = \hbar/p$),

the classical electron radius ($r_0 = e^2/mc^2$), and the Bohr radius ($\alpha_0 = \hbar^2/me^2$). Of these, the only possible candidates for l in Equation (3.197) are Λ and $\lambdabar$ if we are considering free-particle processes,[37] for both r_0 and a_0 involve the electronic charge e and all such dependence should be contained within the factor α^n. The squared ratio $(\Lambda/\lambdabar)^2$ approaches

$$(\Lambda/\lambdabar)^2 \rightarrow \begin{cases} \beta^2 & \text{(NR)} \\ \gamma^2 & \text{(ER)} \end{cases} \tag{3.198}$$

in the non-relativistic and extreme-relativistic limits. The factor β^2 and γ^2 are kinematic in nature and in the intermediate domain where c.m. energies are of the order mc^2 the ratio $\Lambda/\lambdabar$ is of order unity. At this "characteristic" energy, the appropriate choice for l is the constant Λ, and so the characteristic cross section is of the order

$$\sigma \sim \alpha^n \Lambda^2. \tag{3.199}$$

Figure 3.8 exhibits Feynman diagrams for a variety of electromagnetic processes, with those on each row having the same order and so the same characteristic cross section. For each process, only one diagram is shown; some processes have a number of other diagrams that contribute to the total amplitude. The top three rows exhibit diagrams with only one kind of vertex, corresponding to the charged spin-$\frac{1}{2}$ case. In the bottom three rows are diagrams for processes involving spin-0 charges for which there are two kinds of vertices. For spin-0 charges, there would also be the diagrams on the top half contributing to the total amplitude for the corresponding processes. Also, in a non-relativistic formulation, there would be both the single- and double-photon vertex in a diagrammatic representation—even for spin-$\frac{1}{2}$ charges. Finally, it should be noted that, for simplicity, the arrows have been left off the diagrams in Figure 3.8. It is clear, for example, that a V-like vertex involves production of a particle-antiparticle pair, etc. The description of the processes below the diagrams will also clarify the nature of the processes represented. In the processes (bremsstrahlung, for example) where there is an "extra" photon produced, the differential cross section will always be proportional to $d\omega/\omega$, as is indicated, for these cases, in parentheses for the corresponding cross section on the right.

We can be a little more sophisticated in expressing the cross section for an electromagnetic process. Formally, at least, the origin of the energy dependence of a cross section can be indicated to give, for an individual process, an expression modifying the characteristic value (3.199). The energy dependence can come from three different factors in the cross section. For a process, the product of the cross section and the incident particle flux (essentially its incident velocity v_0) gives the transition probability per unit time ($\Delta W/\Delta t$). This quantity is, in turn, given by the square of the matrix element (M_{f0}) for the process times the final state phase space per unit energy (Φ_f, see Section 1.3.3). Thus, we write

$$\sigma v_0 \propto |M_{f0}|^2 \Phi_f. \tag{3.200}$$

---

[37]For processes involving bound electrons, a_0 might be a candidate. In fact, the characteristic cross section for photoionization is of the order αa_0^2.

$$\sigma \sim \alpha^2 \Lambda^2$$

Coulomb scatt. Compton scatt. pair prod. in γ–γ pair annihilation

$$\sigma \sim \alpha^3 \Lambda^2 (d\omega/\omega)$$

Coulomb bremsung. pair prod. in Coul. field radiative Compton radiative γ–γ pair prod

$$\sigma \sim \alpha^4 \Lambda^2$$

γ–γ scatt. pair prod. in Coul. scatt. double pair prod in γ–γ

$$\sigma \sim \alpha^2 \Lambda^2$$

Compton scatt. pair prod. in γ–γ pair annihilation

$$\sigma \sim \alpha^3 \Lambda^2 (d\omega/\omega)$$

Coulomb bremsung. pair prod. in Coul. field radiative Compton radiative γ–γ pair prod.

$$\sigma \sim \alpha^4 \Lambda^2$$

pair prod. in Coul. scatt. double pair prod in γ–γ

Figure 3.8 A collection of diagrams associated with various processes. There are other diagrams for the particular processes. The top part is for the covariant spin-$\frac{1}{2}$ formulation. The bottom diagrams would be for the covariant spin-0 formulation or a non-relativistic formulation and include the two-photon vertex.

The energy dependence in the cross section for a process is contained in v_0, M_{f0}, and Φ_f. If we express these in dimensionless form by dividing by their values at the characteristic energy $E_c \sim mc^2$, the cross section can be written

$$\sigma \sim \alpha^n \Lambda^2 \overline{\eta}, \tag{3.201}$$

where

$$\overline{\eta} = \overline{\mu}_{f0}^2 \overline{\Phi}_f / \overline{\beta}_0, \tag{3.202}$$

with $(\overline{\beta}_0 = v_0/c)$

$$\overline{\mu}_{f0}^2 = \left|M_{f0}(E)\right|^2 / \left|M_{f0}(E_c)\right|^2 , \tag{3.203}$$

$$\overline{\Phi}_f = \Phi_f(E)/\Phi_f(E_c). \tag{3.204}$$

The phase-space factor Φ_f always increases with increasing energy; for a single-particle phase space, $\Phi_f \propto v_f$ and E_f^2 in the NR and ER limits, respectively. The matrix element M_{f0} generally decreases with increasing energy at high energies. The energy dependence of $\overline{\eta}$ thus depends on the nature of the process, that is, its amplitude, number of final-state particles, etc.

3.6.2 Radiative Corrections and Renormalization

In Section 3.1, brief mention was made of the problems encountered in QED when attempts were made to compute processes higher order in perturbation theory. After the initial development of QED in 1927, in the early 1930s it was noticed that in the calculation of perturbation amplitudes to higher order, divergences sometimes appeared in the theory when summations over intermediate states were performed. These problems occurred, in particular, when the higher-order amplitude involved perturbations that could be described in terms of diagrams in which a virtual photon is "emitted and reabsorbed" by the same charge (in loose terminology). These "photons" are thus not observed but contribute to the corresponding correction (M_1) to the lowest-order amplitude (M_0). Then, for a process in which the final state is specified, the total amplitude is

$$M = M_0 + M_1 + \cdots . \tag{3.205}$$

Here, for each M_k, the final state is the same, so that the probability for the process is

$$W = |M_0 + M_1 + \cdots|^2 = |M_0|^2 + M_0 M_1^* + M_1 M_0^* + \cdots . \tag{3.206}$$

The type of amplitudes that would be part of M_1 can be indicated in terms of the diagrams[38] in Figure 3.9. These are the types[39] modifying the lowest-order amplitude for which M_0 is part of the total diagram for the complete process. The characteristic magnitude for each amplitude is determined by $\overline{e}^n$, where $\overline{e} = \alpha^{1/2}$ is the dimensionless (electron) charge and n is the number of vertices. Moreover, there is a special terminology for these types of diagram parts, as is indicated. Because of its role in giving rise to the electron self-energy due to electromagnetic effects, parts with virtual photons being emitted and reabsorbed before another interaction (vertex) are called "self-energy parts." Then there is the "vertex part" of the same order in $\overline{e}$. The "photon self-energy part" has a virtual electron-positron pair bubble; this type of virtual state is also connected with a phenomenon known as *vacuum polarization*. At the bottom of Figure 3.9 are what we shall call "radiative parts." These perturbation amplitudes (M_r) are associated with a different process, namely,

---

[38] Here we consider only one kind of basic vertex as in the covariant formulation for spin-$\frac{1}{2}$ charges.

[39] For a complete process with associated diagrams, there can be other kinds of higher-order diagrams. For example, in Coulomb scattering (see Figure 3.8), there can be the higher-order amplitude associated with the exchange of two photons.

M_0
$(\bar{e})$ (lowest-order part)

M_l

$(\bar{e}^3)$ (electron self-
                                                                              energy part)

                                                                              (photon self-
                                                                              energy part)

M_r
$(\bar{e}^2)$ (radiative parts)
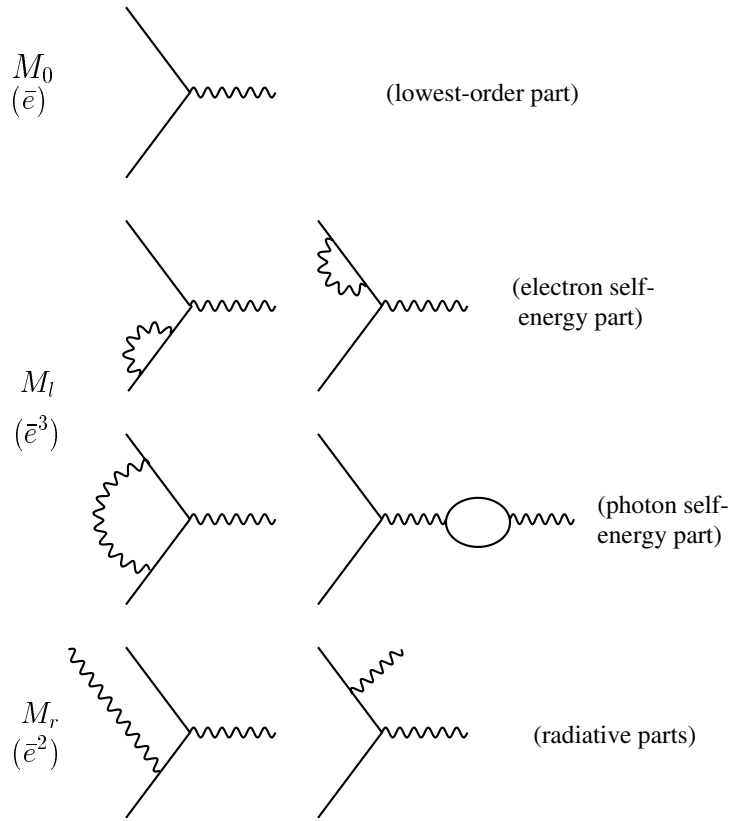
Figure 3.9 Self-energy diagrams.

that in which there is an additional photon in the final state. Thus, they do not contribute to the total amplitude (3.205), and the associated probability is

$$W_r = |M_r|^2 . \tag{3.207}$$

We see that, while $|M_0|^2$ is of order α (times something from the rest of the diagram), the lowest-order correction term is of order α^2; W_r is also of order α^2.

When the corrected amplitude M_1 is computed by the standard methods, it is found, for example, in the integration over the four-dimensional volume d^4k' associated with the virtual photon state, that there are divergences. Divergences occur, in different instances, at both the low and high end of the k' integration and came to be called, respectively, the "infrared catastrophe" and the "ultraviolet catastrophe." The former is really no catastrophe at all, as first shown by Bloch and Nordsieck in 1937. They explained that for any process where we consider some particular final state, our ability to observe the details of the process depends on the experimental energy resolution and detector sensitivities. In particular, if an additional soft photon is emitted in the process, our measuring apparatus may not be capable of detecting it. Suppose that the apparatus has a threshold ε_t so that photons of energy below ε_t

cannot be detected. Then, experimentally we are looking at the process without an extra photon *plus* the radiative process with an extra photon of any energy *less than* ε_t. The associated probability would be

$$E_{\text{expt}} = W + W_r = W_0 + W_1 + \cdots + W_r, \tag{3.208}$$

with $W_0 = |M_0|^2$ and $W_1 = M_0 M_1^* + M_1 M_0^*$ from Equation (3.206). If we write

$$W_{\text{expt}} = W_0(1 + \xi), \tag{3.209}$$

we find that the relative correction ξ is of order $\bar{e}^2 = \alpha$ and composed of several parts. In the two terms

$$\xi = W_1/W_0 + W_r/W_0, \tag{3.210}$$

there is a cancellation of the infrared divergence. This can be shown to order α when explicit processes are considered, such as bremsstrahlung or Compton scattering, and also for a general process.[40] That is, if we take ε_0 as the lower limit in the (extra) photon spectrum in W_r and also in the lower limit of the virtual photon spectrum in the (infrared) divergent part of W_1, we find that both integrals have the same coefficient except for a sign difference. In W_r there is a term of the form $\ln(\varepsilon_t/\varepsilon_0)$ (see Section 3.5) and in W_1 there is one of the form $\ln(\varepsilon_0/\varepsilon_{\text{char}})$, where $\varepsilon_{\text{char}}$ is some characteristic energy in the problem (such as mc^2 or some charged particle energy). The terms have the same coefficient and the sum is then of the form $\ln(\varepsilon_t/\varepsilon_{\text{char}})$ and contains no divergence for $\varepsilon_0 \to 0$.

The occurrence of infrared divergences is a consequence of the particular mathematical formulation of the subject. When the results are expressed in a form more suited to experience or measurement, the divergence disappears. Not that the soft photons emitted are unreal; there *are* an infinite number of infinitesimally soft photons emitted during, say, the scattering of a charged particle. However, the very soft photons have no physical effect on the process. In fact, in the case of very soft photons when the wavelength is very large, there is little distinction between real and virtual photons. The real photons are eventually reabsorbed someplace and in that sense are virtual in their temporary existence. Thus, it is understandable that there is a cancellation of the infrared divergences, their occurrence being simply a mathematical artifact. In fact, in most problems involving soft photons, it is actually preferable to treat phenomena by classical electrodynamics, which is a limiting domain of QED. In the classical theory, the photon concept does not appear and the mathemtical formulation does not result in an infrared "catastrophe."

The ultraviolet divergence is a more serious problem in QED and was handled in calculations of phenomena for the first time in 1947—twenty years after the subject was originally developed. The first such calculation was by Bethe in an evaluation of the Lamb shift, and the techniques of the method have been developed fully and incorporated into modern relativistic QED. The resolution of this type of divergence problem with the theory is not clean like that for the infrared divergence, and it is generally regarded as a rather unsatisfactory procedure even though it seems to allow

---

[40]See J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons*, 2nd ed., Berlin: Springer-Verlag, 1976. This book gives an excellent discussion of the divergences in QED and, in particular, of the infrared divergence.

accurate higher-order calculations of any observable electromagnetic phenomenon. In the end, perhaps the ultraviolet divergence will, in some future formulation of QED, like the infrared problem, be shown to be a mathematical artifact. This has yet to be achieved, and the present theory still suffers from an uncomfortable subtraction of infinities.

The method for handling the ultraviolet divergence makes use of *renormalization* techniques. This idea, already mentioned in Section 2.6 of the previous chapter on classical electrodynamics, prescribes that results be expressed in terms of the observed or phenomenological mass (m_0) and charge (e_0) of the particles involved in the process. The values m_0 and e_0 contain electromagnetic parts (δm and δe) due to emission and absorption of virtual photons and charged pairs, and these contributions should, in principle, be capable of computation by the theory. The problem is that the theory gives infinity for δm and δe when it is applied. However, when radiative corrections to processes are computed and then formally expressed in terms of m_0 and e_0, the ultraviolet divergences can be handled by introducing covariant cutoff or convergence factors, so that the whole computational procedure is made systematic. In a practical sense, the whole technique seems to work very well. However, the parameters m_0 and e_0 are measured quantities, and the theory really gives infinity for both when it is applied.

We go no further into this topic. Generally, the coefficients of the higher-order corrections in α^n have numerical values of order unity when the theory is handled as prescribed above. The nature of the infrared and ultraviolet divergences seem to be quite different. The former difficulty is completely understood and resolved, while the latter is handled only with some embarrassment. For a more extensive discussion, the reader is referred to the standard textbooks on the subject.

3.6.3 Kinematic Invariants

In covariant perturbation theory, the amplitude $\mathcal{M}_{f0}$ for a process is itself a Lorentz invariant, being a function of invariants involving the energies and momenta of the various particles that are present in the initial and final states. These are the so-called kinematic invariants constructed from the four-momenta of the incoming and outgoing particles. In particular, for the case where there are two incoming and two outgoing particles, certain of these invariants are particularly convenient and have some physical significance. For the process

$$a + b \rightarrow c + d, \tag{3.211}$$

kinematic invariants can be constructed from the individual four-momenta $(p_\mu)_\alpha = (iE/c, \boldsymbol{p})_\alpha$, and there are four such quantities. Designating an individual four-momentum simply as p_α, we can express the conservation of four-momentum in the reaction (3.211) as

$$\sum_{\alpha=1}^{4} q_\alpha = 0, \tag{3.212}$$

if $q_1 = p_a$, $q_2 = p_b$, $q_3 = -p_c$, and $q_4 = -p_d$.

In addition to the squares of the individual q_α $[(q_\mu q_\mu)_\alpha = q_\alpha^2 = -m_\alpha^2 c^4]$, which are just constants, six invariants $q_\alpha \cdot q_\beta$ can be formed from the six combinations of α and β. These quantities are kinematic invariants, but only two are independent

because of the four conservation relations (3.212) for each component ($\mu = 0, 1, 2, 3$) of four-momentum. However, instead of the $q_\alpha \cdot q_\beta$, it is more convenient to introduce the following three invariants:

$$s = (q_1 + q_2)^2 = (q_3 + q_4)^2,$$
$$t = (q_1 + q_3)^2 = (q_2 + q_4)^2, \tag{3.213}$$
$$u = (q_1 + q_4)^2 = (q_2 + q_3)^2.$$

Here the squares represent four-dimensional scalar products and, since, as is easily seen,

$$s + t + u = -\sum_\alpha m_\alpha^2 c^4, \tag{3.214}$$

only two of the three invariants are independent.

The quantities s, t, and u are what we shall be referring to when we speak of kinematic invariants. Although only two are independent, it is convenient to introduce all three. This is because in addition to the process (3.211) for specific particles a, b, c, and d, there are also reactions involving the corresponding anti-particles. If, for example, b and c are transferred to the other sides of the arrow and made to represent their anti-particles (designated by a bar), the reaction is $a + \bar{c} \to \bar{b} + d$. In fact, designating the particles by numbers instead of letters, we can have the following three "channels" for reactions with two incoming and two outgoing particles[41]:

$$1 + 2 \to 3 + 4 \quad (s),$$
$$1 + \bar{3} \to \bar{2} + 4 \quad (t), \tag{3.215}$$
$$1 + \bar{4} \to 3 + \bar{2} \quad (u).$$

The channels are sometimes referred to as the "s", "t", and "u" channels, as indicated above. The reason for this is that s, t, and u have a very physical significance for the corresponding channels. They are, for the channels, just minus the total c.m. energies (divided by c^2).

In the case of elastic scattering, the expressions for s, t, and u are particularly simple. Then the type of outgoing particle is the same as the incoming ones: $m_1 = m_3$ and $m_2 = m_4$. In terms of c.m. quantities, the q's are

$$q_1 = (i E_1/c, \mathbf{p}_s), \qquad q_2 = (i E_2/c, -\mathbf{p}_s),$$
$$q_3 = -(i E_3/c, \mathbf{p}'_s), \qquad q_4 = -(i E_4/c, -\mathbf{p}'_s), \tag{3.216}$$

where $\mathbf{p}_s$ and $\mathbf{p}'_s$ (or their negative) refer to the initial and final momenta, respectively. Since the scattering is elastic, $|\mathbf{p}_s| = |\mathbf{p}'_s| \equiv p_s$ and so $E_1 = E_3$ and $E_2 = E_4$. The values for the three invariants can then be written

$$s = -(E_1 + E_2)^2/c^2,$$
$$t = 2p_s^2(1 - \cos\theta_s), \tag{3.217}$$
$$u = -(E_1 - E_2)^2/c^2 + 2p_s^2(1 + \cos\theta_s),$$

where θ_s is the c.m. scattering angle between $\mathbf{p}_s$ and $\mathbf{p}'_s$.

---

[41]The reactions can also procede in either direction, these additional processes being simply the time reverse of the others.

3.6.4 Crossing Symmetry

The kinematic invariants are particularly useful in connection with a fundamental property of covariant perturbation theory. Because the amplitude $\mathcal{M}_{f0}$ introduced in Equation (3.109) is an invariant, it must be a function of invariants. These invariants must be associated with the observable characteristics of the process, that is, with parameters of the initial and final states. For a reaction of the type (3.211), these would be the invariants s, t, and u. More properly, for the case where the particles have spin, if we are not interested in the spin states of the incoming and outgoing particles, we should average over initial spin states and sum over final states. Then the only remaining kinematic parameters for these states are the particle four-momenta, and the invariant squared amplitude can be written

$$\sum_{\text{all spins}} \left| \mathcal{M}_{f0} \right|^2 = f(s, t, u). \tag{3.218}$$

One of the significant and useful features of covariant QED is the basic symmetry in its formulation in terms of particle and antiparticle states. The unified description of processes involving, in particular, electrons and positrons has been examined only in a superficial manner in this chapter. A brief discussion was given in Section 4.2, indicating how the theory encompasses, in a convenient and simplified way, the inclusion of positron and pair-production effects in terms of a backward-in-time description of the antiparticle states. As we have seen, it is useful to speak of "entry" and "exit" states in evaluating the total amplitude for the process. The unified treatment is also appropriate in the description of initial and final states for processes, that is, in the diagrams, for the meaning of "external lines." In this formulation, the cases corresponding to an initial-state particle and a final-state antiparticle are equivalent in that a single expression for an amplitude can refer to either case. The symmetric covariant theory yields a useful result that allows a convenient way of obtaining the squared amplitude (3.218), spin-averaged, for any cross channel (3.215) in terms of that for one of the other channels. That is, as a result of the unified formulation of the theory, a single function f yields the squared amplitude for each of the three channels (3.215). The invariants s, t, and u will be different for each channel, but the functional form (f) will be the same. This is what is known by *crossing symmetry*.

Figure 3.10 indicates the replacements for s, t, and u in channels II and III as obtained from those for channel I. When an antiparticle is involved, the q_α has a minus sign as a result of the backward-in-time description. Also exhibited in Figure 3.10 are three examples of specific physical processes in the three corresponding channels. In addition to these, there are the time-reversed processes that have the same squared amplitude. We see, for example, that the processes of Compton scattering, pair annihilation into two photons, and pair production in photon-photon collisions are closely related. Note that, since the photon is identical to its antiparticle ($\gamma = \overline{\gamma}$), we do not distinguish the two. In transferring the photon to the opposite side of the reaction equation (3.211) the processes of absorption and emission are interchanged. We should also keep in mind that for antiparticles, even though they are indicated as going backward in time in a diagram, the actual momentum of the particle is in the opposite direction; in other words, $p_{\text{anti-particle}} = -p_{\text{diagram}}$. In

I: $\quad a + b \to c + d$

$\quad\quad 1 + 2 \to 3 + 4$

$\quad\quad \gamma + e \to \gamma + e$

$\quad\quad e + e \to e + e$

$\quad\quad e + \mu \to e + \mu$

$s_I = (q_1 + q_2)^2$

$t_I = (q_1 + q_3)^2$

$u_I = (q_1 + q_4)^2$

II: $\quad a + \bar{c} \to \bar{b} + d$

$\quad\quad 1 + \bar{3} \to \bar{2} + 4$

$\quad\quad \gamma + \gamma \to \bar{e} + e$

$\quad\quad e + \bar{e} \to \bar{e} + e$

$\quad\quad e + \bar{e} \to \bar{\mu} + \mu$

$s_{II} = (q_1 - q_3)^2$

$t_{II} = (q_1 - q_2)^2$

$u_{II} = (q_1 - q_2)^2$

III: $\quad a + \bar{d} \to c + \bar{b}$

$\quad\quad 1 + \bar{4} \to 3 + \bar{2}$

$\quad\quad \gamma + \bar{e} \to \bar{e} + \gamma$

$\quad\quad e + \bar{e} \to \bar{e} + e$

$\quad\quad e + \bar{\mu} \to e + \bar{\mu}$

$s_{III} = (q_1 - q_4)^2$

$t_{III} = (q_1 + q_3)^2$
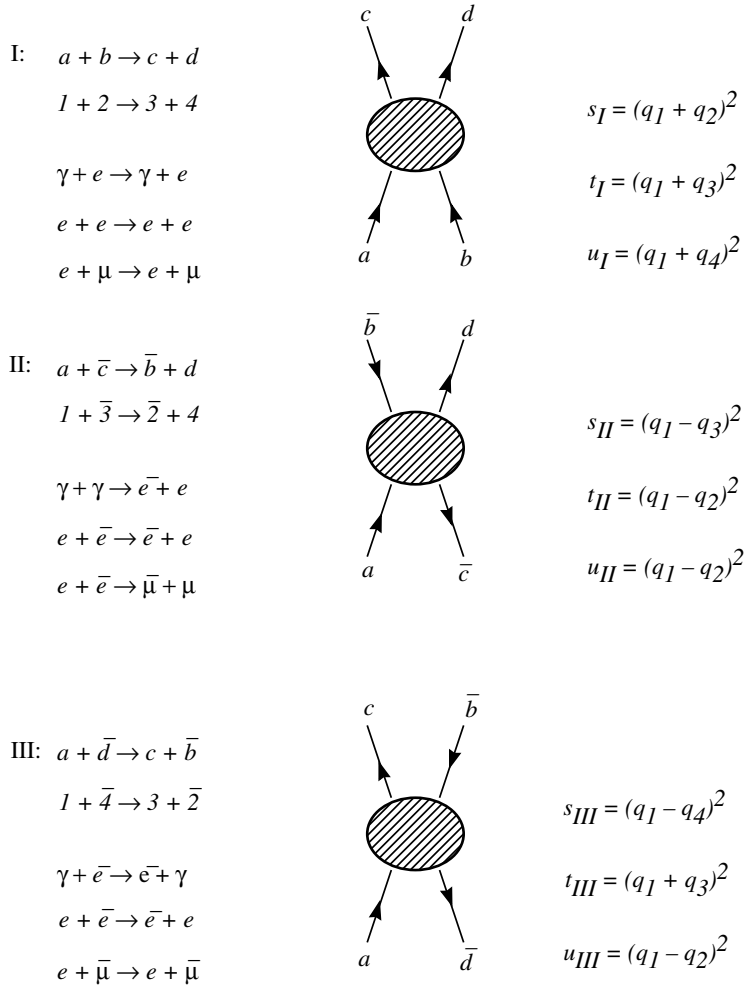
$u_{III} = (q_1 - q_2)^2$

Figure 3.10 Channels associated with related processes.

evaluating the kinematics of the process in terms of diagrams and Feynman rules, the four-momenta are handled as in the diagram.

There are other examples of crossing symmetry besides those associated with the processes in Figure 3.10. For example, bremsstrahlung and pair production are related by crossing symmetry (see Figure 3.8). In fact, this was noted already in 1934 by Bethe and Heitler. Actually, the existence of this type of symmetry is very general and not restricted to electromagnetic processes. Finally, it should be mentioned that the physical domains of s, t, and u for the three channels of a process would be different. Mathematically, it is said that the scattering amplitude, being an analytic function, is "analytically continued" from one domain to another for the three channels.

BIBLIOGRAPHICAL NOTES

A brief historical summary of quantum electrodynamics, along with a collection of fundamental papers may be found in

1. Schwinger, J., *Quantum Electrodynamics*, New York: Dover Publ., Inc., 1958.

To this collection of papers on QED the following could be added:

2. Fermi, E., *Rev. Mod. Phys.* **4**, 87 (1932); see also Reference 8 of Chapter 1.

There are a number of excellent textbooks on modern covariant QED, especially

3. Jauch, J. M. and Rohrlich, F., *The Theory of Photons and Electrons*, 2nd ed., Berlin: Springer-Verlag, 1976.

4. Berestetskii, V. B., Lifshitz, E. M., and Pitaevskii, L. P., *Relativistic Quantum Theory, Parts 1 and 2*, Reading, MA: Addison-Wesley Publ. Co., 1979.

5. Bjorken, J. D. and Drell, S., *Relativistic Quantum Mechanics*, New York: McGraw-Hill, 1964.

6. Feynman, R. P., *Quantum Electrodynamics*, New York: W. A. Benjamin, Inc., 1961.

7. Feynman, R. P., *Theory of Fundamental Processes*, New York: W. A. Benjamin, Inc., 1962.

8. Dyson, F. J., *Advanced Quantum Mechanics* (mimeographed lecture notes), Laboratory of Nuclear Studies, Cornell University, 1951.

Not so modern, but excellent, is

9. Heitler, W., *The Quantum Theory of Radiation*, Oxford, UK: Oxford University Press, 1954.

For the material in Sections 3.5 and 3.6, References 3 and 4 are especially good.