The other five diagrams in Fig. 1.4 involve intermediate states of several virtual particles rather than just a single virtual photon. In each of these diagrams there will be one virtual particle whose momentum is not determined by conservation of momentum at the vertices. Since perturbation theory requires us to sum over all possible intermediate states, we must integrate over all possible values of this momentum. At this step, however, a new difficulty appears: The loop-momentum integrals in the first three diagrams, when performed naively, turn out to be infinite. We will provide a fix for this problem, so that we get finite results, by the end of Part I. But the question of the physical origin of these divergences cannot be dismissed so lightly; that will be the main subject of Part II of this book.

We have discussed Feynman diagrams as an algorithm for performing computations. The chapters that follow should amply illustrate the power of this tool. As we expose more applications of the diagrams, though, they begin to take on a life and significance of their own. They indicate unsuspected relations between different physical processes, and they suggest intuitive arguments that might later be verified by calculation. We hope that this book will enable you, the reader, to take up this tool and apply it in novel and enlightening ways.

Chapter 2

The Klein-Gordon Field

2.1 The Necessity of the Field Viewpoint

Quantum field theory is the application of quantum mechanics to dynamical systems of *fields*, in the same sense that the basic course in quantum mechanics is concerned mainly with the quantization of dynamical systems of *particles*. It is a subject that is absolutely essential for understanding the current state of elementary particle physics. With some modification, the methods we will discuss also play a crucial role in the most active areas of atomic, nuclear, and condensed-matter physics. In Part I of this book, however, our primary concern will be with elementary particles, and hence *relativistic* fields.

Given that we wish to understand processes that occur at very small (quantum-mechanical) scales and very large (relativistic) energies, one might still ask why we must study the quantization of *fields*. Why can't we just quantize relativistic particles the way we quantized nonrelativistic particles?

This question can be answered on a number of levels. Perhaps the best approach is to write down a single-particle relativistic wave equation (such as the Klein-Gordon equation or the Dirac equation) and see that it gives rise to negative-energy states and other inconsistencies. Since this discussion usually takes place near the end of a graduate-level quantum mechanics course, we will not repeat it here. It is easy, however, to understand why such an approach cannot work. We have no right to assume that any relativistic process can be explained in terms of a single particle, since the Einstein relation $E = mc^2$ allows for the creation of particle-antiparticle pairs. Even when there is not enough energy for pair creation, multiparticle states appear, for example, as intermediate states in second-order perturbation theory. We can think of such states as existing only for a very short time, according to the uncertainty principle $\Delta E \cdot \Delta t = \hbar$. As we go to higher orders in perturbation theory, arbitrarily many such "virtual" particles can be created.

The necessity of having a multiparticle theory also arises in a less obvious way, from considerations of causality. Consider the amplitude for a free particle to propagate from $\mathbf{x}_0$ to $\mathbf{x}$:

$$U(t) = \langle \mathbf{x} | \, e^{-iHt} \, | \mathbf{x}_0 \rangle \, .$$

In nonrelativistic quantum mechanics we have $E = \mathbf{p}^2/2m$, so

$$\begin{aligned} U(t) &= \langle \mathbf{x}|\, e^{-i(\mathbf{p}^2/2m)t}\, |\mathbf{x}_0\rangle \\ &= \int d^3p\ \langle \mathbf{x}|\, e^{-i(\mathbf{p}^2/2m)t}\, |\mathbf{p}\rangle\, \langle \mathbf{p}|\, \mathbf{x}_0\rangle \\ &= \frac{1}{(2\pi)^3}\int d^3p\, e^{-i(\mathbf{p}^2/2m)t}\cdot e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_0)} \\ &= \left(\frac{m}{2\pi i t}\right)^{3/2} e^{im(\mathbf{x}-\mathbf{x}_0)^2/2t}. \end{aligned}$$

This expression is nonzero for all x and t, indicating that a particle can propagate between any two points in an arbitrarily short time. In a relativistic theory, this conclusion would signal a violation of causality. One might hope that using the relativistic expression $E = \sqrt{p^2 + m^2}$ would help, but it does not. In analogy with the nonrelativistic case, we have

$$\begin{aligned} U(t) &= \langle \mathbf{x}|\, e^{-it\sqrt{\mathbf{p}^2+m^2}}\, |\mathbf{x}_0\rangle \\ &= \frac{1}{(2\pi)^3}\int d^3p\, e^{-it\sqrt{\mathbf{p}^2+m^2}}\cdot e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_0)} \\ &= \frac{1}{2\pi^2|\mathbf{x}-\mathbf{x}_0|}\int_0^\infty dp\, p\, \sin(p|\mathbf{x}-\mathbf{x}_0|)e^{-it\sqrt{p^2+m^2}}. \end{aligned}$$

This integral can be evaluated explicitly in terms of Bessel functions.* We will content ourselves with looking at its asymptotic behavior for $x^2 \gg t^2$ (well outside the light-cone), using the method of stationary phase. The phase function $px - t\sqrt{p^2+m^2}$ has a stationary point at $p = imx/\sqrt{x^2-t^2}$. We may freely push the contour upward so that it goes through this point. Plugging in this value for p, we find that, up to a rational function of x and t,

$$U(t) \sim e^{-m\sqrt{x^2-t^2}}.$$

Thus the propagation amplitude is small but nonzero outside the light-cone, and causality is still violated.

Quantum field theory solves the causality problem in a miraculous way, which we will discuss in Section 2.4. We will find that, in the multiparticle field theory, the propagation of a particle across a spacelike interval is indistinguishable from the propagation of an *antiparticle* in the opposite direction (see Fig. 2.1). When we ask whether an observation made at point x_0 can affect an observation made at point x, we will find that the amplitudes for particle and antiparticle propagation exactly cancel—so causality is preserved.

Quantum field theory provides a natural way to handle not only multiparticle states, but also transitions between states of different particle number. It solves the causality problem by introducing antiparticles, then goes on to

*See Gradshteyn and Ryzhik (1980), #3.914.

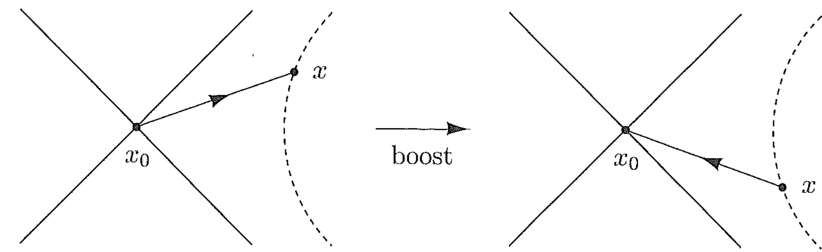


Figure 2.1. Propagation from x_0 to x in one frame looks like propagation from x to x_0 in another frame.

explain the relation between spin and statistics. But most important, it provides the tools necessary to calculate innumerable scattering cross sections, particle lifetimes, and other observable quantities. The experimental confirmation of these predictions, often to an unprecedented level of accuracy, is our real reason for studying quantum field theory.

2.2 Elements of Classical Field Theory

In this section we review some of the formalism of classical field theory that will be necessary in our subsequent discussion of quantum field theory.

Lagrangian Field Theory

The fundamental quantity of classical mechanics is the action, S, the time integral of the Lagrangian, L. In a local field theory the Lagrangian can be written as the spatial integral of a Lagrangian density, denoted by $\mathcal{L}$, which is a function of one or more fields $\phi(x)$ and their derivatives $\partial_\mu\phi$. Thus we have

$$S = \int L\, dt = \int \mathcal{L}(\phi, \partial_\mu\phi)\, d^4x. \tag{2.1}$$

Since this is a book on field theory, we will refer to $\mathcal{L}$ simply as the Lagrangian.

The principle of least action states that when a system evolves from one given configuration to another between times t_1 and t_2, it does so along the "path" in configuration space for which S is an extremum (normally a minimum). We can write this condition as

$$\begin{aligned} 0 &= \delta S \\ &= \int d^4x \left\{ \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi) \right\} \\ &= \int d^4x \left\{ \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\delta\phi + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) \right\}. \end{aligned} \tag{2.2}$$

The last term can be turned into a surface integral over the boundary of the four-dimensional spacetime region of integration. Since the initial and final field configurations are assumed given, $\delta\phi$ is zero at the temporal beginning and end of this region. If we restrict our consideration to deformations $\delta\phi$ that vanish on the spatial boundary of the region as well, then the surface term is zero. Factoring out the $\delta\phi$ from the first two terms, we note that, since the integral must vanish for arbitrary $\delta\phi$, the quantity that multiplies $\delta\phi$ must vanish at all points. Thus we arrive at the Euler-Lagrange equation of motion for a field,

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0. \tag{2.3}$$

If the Lagrangian contains more than one field, there is one such equation for each.

Hamiltonian Field Theory

The Lagrangian formulation of field theory is particularly suited to relativistic dynamics because all expressions are explicitly Lorentz invariant. Nevertheless we will use the Hamiltonian formulation throughout the first part of this book, since it will make the transition to quantum mechanics easier. Recall that for a discrete system one can define a conjugate momentum $p \equiv \partial L/\partial\dot{q}$ (where $\dot{q} = \partial q/\partial t$) for each dynamical variable q. The Hamiltonian is then $H \equiv \sum p\dot{q} - L$. The generalization to a continuous system is best understood by pretending that the spatial points $\mathbf{x}$ are discretely spaced. We can define

$$\begin{aligned} p(\mathbf{x}) \equiv \frac{\partial L}{\partial\dot{\phi}(\mathbf{x})} &= \frac{\partial}{\partial\dot{\phi}(\mathbf{x})}\int \mathcal{L}(\phi(\mathbf{y}), \dot{\phi}(\mathbf{y}))\, d^3y \\ &\sim \frac{\partial}{\partial\dot{\phi}(\mathbf{x})}\sum_{\mathbf{y}} \mathcal{L}(\phi(\mathbf{y}), \dot{\phi}(\mathbf{y}))\, d^3y \\ &= \pi(\mathbf{x})\, d^3x, \end{aligned}$$

where

$$\pi(\mathbf{x}) \equiv \frac{\partial\mathcal{L}}{\partial\dot{\phi}(\mathbf{x})} \tag{2.4}$$

is called the *momentum density* conjugate to $\phi(\mathbf{x})$. Thus the Hamiltonian can be written

$$H = \sum_{\mathbf{x}} p(\mathbf{x})\dot{\phi}(\mathbf{x}) - L.$$

Passing to the continuum, this becomes

$$H = \int d^3x\, \left[\pi(\mathbf{x})\dot{\phi}(\mathbf{x}) - \mathcal{L}\right] \equiv \int d^3x\, \mathcal{H}. \tag{2.5}$$

We will rederive this expression for the Hamiltonian density $\mathcal{H}$ near the end of this section, using a different method.

As a simple example, consider the theory of a single field $\phi(x)$, governed by the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}\dot{\phi}^2 - \tfrac{1}{2}(\nabla\phi)^2 - \tfrac{1}{2}m^2\phi^2 \tag{2.6}$$

For now we take ϕ to be a real-valued field. The quantity m will be interpreted as a mass in Section 2.3, but for now just think of it as a parameter. From this Lagrangian the usual procedure gives the equation of motion

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\phi = 0 \qquad \text{or} \qquad (\partial^\mu\partial_\mu + m^2)\phi = 0, \tag{2.7}$$

which is the well-known Klein-Gordon equation. (In this context it is a *classical* field equation, like Maxwell's equations—not a quantum-mechanical wave equation.) Noting that the canonical momentum density conjugate to $\phi(x)$ is $\pi(x) = \dot{\phi}(x)$, we can also construct the Hamiltonian:

$$H = \int d^3x\, \mathcal{H} = \int d^3x\, \left[\tfrac{1}{2}\pi^2 + \tfrac{1}{2}(\nabla\phi)^2 + \tfrac{1}{2}m^2\phi^2\right]. \tag{2.8}$$

We can think of the three terms, respectively, as the energy cost of "moving" in time, the energy cost of "shearing" in space, and the energy cost of having the field around at all. We will investigate this Hamiltonian much further in Sections 2.3 and 2.4.

Noether's Theorem

Next let us discuss the relationship between symmetries and conservation laws in classical field theory, summarized in *Noether's theorem*. This theorem concerns continuous transformations on the fields ϕ, which in infinitesimal form can be written

$$\phi(x) \to \phi'(x) = \phi(x) + \alpha\Delta\phi(x), \tag{2.9}$$

where α is an infinitesimal parameter and $\Delta\phi$ is some deformation of the field configuration. We call this transformation a symmetry if it leaves the equations of motion invariant. This is insured if the action is invariant under (2.9). More generally, we can allow the action to change by a surface term, since the presence of such a term would not affect our derivation of the Euler-Lagrange equations of motion (2.3). The Lagrangian, therefore, must be invariant under (2.9) up to a 4-divergence:

$$\mathcal{L}(x) \to \mathcal{L}(x) + \alpha\partial_\mu\mathcal{J}^\mu(x), \tag{2.10}$$

for some $\mathcal{J}^\mu$. Let us compare this expectation for $\Delta\mathcal{L}$ to the result obtained by varying the fields:

$$\begin{aligned} \alpha\Delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi}(\alpha\Delta\phi) + \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\partial_\mu(\alpha\Delta\phi) \\ &= \alpha\,\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Delta\phi\right) + \alpha\left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right]\Delta\phi. \end{aligned} \tag{2.11}$$

The second term vanishes by the Euler-Lagrange equation (2.3). We set the remaining term equal to $\alpha\partial_\mu\mathcal{J}^\mu$ and find

$$\partial_\mu j^\mu(x) = 0, \qquad \text{for} \quad j^\mu(x) = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Delta\phi - \mathcal{J}^\mu. \tag{2.12}$$

(If the symmetry involves more than one field, the first term of this expression for $j^\mu(x)$ should be replaced by a sum of such terms, one for each field.) This result states that the current $j^\mu(x)$ is conserved. For each continuous symmetry of $\mathcal{L}$, we have such a conservation law.

The conservation law can also be expressed by saying that the charge

$$Q \equiv \int_{\text{all space}} j^0\, d^3x \tag{2.13}$$

is a constant in time. Note, however, that the formulation of field theory in terms of a local Lagrangian density leads directly to the local form of the conservation law, Eq. (2.12).

The easiest example of such a conservation law arises from a Lagrangian with only a kinetic term: $\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2$. The transformation $\phi \to \phi + \alpha$, where α is a constant, leaves $\mathcal{L}$ unchanged, so we conclude that the current $j^\mu = \partial^\mu\phi$ is conserved. As a less trivial example, consider the Lagrangian

$$\mathcal{L} = |\partial_\mu\phi|^2 - m^2|\phi|^2, \tag{2.14}$$

where ϕ is now a *complex*-valued field. You can easily show that the equation of motion for this Lagrangian is again the Klein-Gordon equation, (2.7). This Lagrangian is invariant under the transformation $\phi \to e^{i\alpha}\phi$; for an infinitesimal transformation we have

$$\alpha\Delta\phi = i\alpha\phi; \qquad \alpha\Delta\phi^* = -i\alpha\phi^*. \tag{2.15}$$

(We treat ϕ and ϕ^* as independent fields. Alternatively, we could work with the real and imaginary parts of ϕ.) It is now a simple matter to show that the conserved Noether current is

$$j^\mu = i\big[(\partial^\mu\phi^*)\phi - \phi^*(\partial^\mu\phi)\big]. \tag{2.16}$$

(The overall constant has been chosen arbitrarily.) You can check directly that the divergence of this current vanishes by using the Klein-Gordon equation. Later we will add terms to this Lagrangian that couple ϕ to an electromagnetic field. We will then interpret j^μ as the electromagnetic current density carried by the field, and the spatial integral of j^0 as its electric charge.

Noether's theorem can also be applied to spacetime transformations such as translations and rotations. We can describe the infinitesimal translation

$$x^\mu \to x^\mu - a^\mu$$

alternatively as a transformation of the field configuration

$$\phi(x) \to \phi(x + a) = \phi(x) + a^\mu\partial_\mu\phi(x).$$

The Lagrangian is also a scalar, so it must transform in the same way:

$$\mathcal{L} \to \mathcal{L} + a^\mu\partial_\mu\mathcal{L} = \mathcal{L} + a^\nu\partial_\mu(\delta^\mu{}_\nu\mathcal{L}).$$

Comparing this equation to (2.10), we see that we now have a nonzero $\mathcal{J}^\mu$. Taking this into account, we can apply the theorem to obtain four separately conserved currents:

$$T^\mu{}_\nu \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi - \mathcal{L}\delta^\mu{}_\nu. \tag{2.17}$$

This is precisely the *stress-energy tensor*, also called the *energy-momentum tensor*, of the field ϕ. The conserved charge associated with time translations is the Hamiltonian:

$$H = \int T^{00}\, d^3x = \int \mathcal{H}\, d^3x. \tag{2.18}$$

By computing this quantity for the Klein-Gordon field, one can recover the result (2.8). The conserved charges associated with spatial translations are

$$P^i = \int T^{0i}\, d^3x = -\int \pi\partial_i\phi\, d^3x, \tag{2.19}$$

and we naturally interpret this as the (physical) momentum carried by the field (not to be confused with the canonical momentum).

2.3 The Klein-Gordon Field as Harmonic Oscillators

We begin our discussion of *quantum* field theory with a rather formal treatment of the simplest type of field: the real Klein-Gordon field. The idea is to start with a classical field theory (the theory of a classical scalar field governed by the Lagrangian (2.6)) and then "quantize" it, that is, reinterpret the dynamical variables as operators that obey canonical commutation relations.† We will then "solve" the theory by finding the eigenvalues and eigenstates of the Hamiltonian, using the harmonic oscillator as an analogy.

The classical theory of the real Klein-Gordon field was discussed briefly (but sufficiently) in the previous section; the relevant expressions are given in Eqs. (2.6), (2.7), and (2.8). To quantize the theory, we follow the same procedure as for any other dynamical system: We promote ϕ and π to operators, and impose suitable commutation relations. Recall that for a discrete system of one or more particles the commutation relations are

$$[q_i, p_j] = i\delta_{ij};$$
$$[q_i, q_j] = [p_i, p_j] = 0.$$

†This procedure is sometimes called *second quantization*, to distinguish the resulting Klein-Gordon equation (in which ϕ is an operator) from the old one-particle Klein-Gordon equation (in which ϕ was a wavefunction). In this book we never adopt the latter point of view; we start with a classical equation (in which ϕ is a classical field) and quantize it exactly once.

For a continuous system the generalization is quite natural; since $\pi(\mathbf{x})$ is the momentum *density*, we get a Dirac delta function instead of a Kronecker delta:

$$\begin{aligned}\left[\phi(\mathbf{x}), \pi(\mathbf{y})\right] &= i\delta^{(3)}(\mathbf{x}-\mathbf{y});\\ \left[\phi(\mathbf{x}), \phi(\mathbf{y})\right] &= \left[\pi(\mathbf{x}), \pi(\mathbf{y})\right] = 0.\end{aligned} \tag{2.20}$$

(For now we work in the Schrödinger picture where ϕ and π do not depend on time. When we switch to the Heisenberg picture in the next section, these "equal time" commutation relations will still hold provided that both operators are considered at the same time.)

The Hamiltonian, being a function of ϕ and π, also becomes an operator. Our next task is to find the spectrum from the Hamiltonian. Since there is no obvious way to do this, let us seek guidance by writing the Klein-Gordon equation in Fourier space. If we expand the classical Klein-Gordon field as

$$\phi(\mathbf{x}, t) = \int \frac{d^3p}{(2\pi)^3}\, e^{i\mathbf{p}\cdot\mathbf{x}}\, \phi(\mathbf{p}, t)$$

(with $\phi^*(\mathbf{p}) = \phi(-\mathbf{p})$ so that $\phi(\mathbf{x})$ is real), the Klein-Gordon equation (2.7) becomes

$$\left[\frac{\partial^2}{\partial t^2} + \left(|\mathbf{p}|^2 + m^2\right)\right]\phi(\mathbf{p}, t) = 0. \tag{2.21}$$

This is the same as the equation of motion for a simple harmonic oscillator with frequency

$$\omega_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2 + m^2}. \tag{2.22}$$

The simple harmonic oscillator is a system whose spectrum we already know how to find. Let us briefly recall how it is done. We write the Hamiltonian as

$$H_{\text{SHO}} = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2\phi^2.$$

To find the eigenvalues of H_{SHO}, we write ϕ and p in terms of ladder operators:

$$\phi = \frac{1}{\sqrt{2\omega}}(a + a^\dagger); \qquad p = -i\sqrt{\frac{\omega}{2}}(a - a^\dagger). \tag{2.23}$$

The canonical commutation relation $[\phi, p] = i$ is equivalent to

$$\left[a, a^\dagger\right] = 1. \tag{2.24}$$

The Hamiltonian can now be rewritten

$$H_{\text{SHO}} = \omega(a^\dagger a + \tfrac{1}{2}).$$

The state $|0\rangle$ such that $a\,|0\rangle = 0$ is an eigenstate of H with eigenvalue $\frac{1}{2}\omega$, the zero-point energy. Furthermore, the commutators

$$\left[H_{\text{SHO}}, a^\dagger\right] = \omega a^\dagger, \qquad \left[H_{\text{SHO}}, a\right] = -\omega a$$

make it easy to verify that the states

$$|n\rangle = (a^\dagger)^n\,|0\rangle$$

are eigenstates of H_{SHO} with eigenvalues $(n+\frac{1}{2})\omega$. These states exhaust the spectrum.

We can find the spectrum of the Klein-Gordon Hamiltonian using the same trick, but now each Fourier mode of the field is treated as an independent oscillator with its own a and $a^\dagger$. In analogy with (2.23) we write

$$\phi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\left(a_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} + a^\dagger_{\mathbf{p}}e^{-i\mathbf{p}\cdot\mathbf{x}}\right); \tag{2.25}$$

$$\pi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}(-i)\sqrt{\frac{\omega_{\mathbf{p}}}{2}}\left(a_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} - a^\dagger_{\mathbf{p}}e^{-i\mathbf{p}\cdot\mathbf{x}}\right). \tag{2.26}$$

The inverse expressions for $a_{\mathbf{p}}$ and $a^\dagger_{\mathbf{p}}$ in terms of ϕ and π are easy to derive but rarely needed. In the calculations below we will find it useful to rearrange (2.25) and (2.26) as follows:

$$\phi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\left(a_{\mathbf{p}} + a^\dagger_{-\mathbf{p}}\right)e^{i\mathbf{p}\cdot\mathbf{x}}; \tag{2.27}$$

$$\pi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}(-i)\sqrt{\frac{\omega_{\mathbf{p}}}{2}}\left(a_{\mathbf{p}} - a^\dagger_{-\mathbf{p}}\right)e^{i\mathbf{p}\cdot\mathbf{x}}. \tag{2.28}$$

The commutation relation (2.24) becomes

$$\left[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}'}\right] = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{p}'), \tag{2.29}$$

from which you can verify that the commutator of ϕ and π works out correctly:

$$\begin{aligned}\left[\phi(\mathbf{x}), \pi(\mathbf{x}')\right] &= \int \frac{d^3p\,d^3p'}{(2\pi)^6}\frac{-i}{2}\sqrt{\frac{\omega_{\mathbf{p}'}}{\omega_{\mathbf{p}}}}\left(\left[a^\dagger_{-\mathbf{p}}, a_{\mathbf{p}'}\right] - \left[a_{\mathbf{p}}, a^\dagger_{-\mathbf{p}'}\right]\right)e^{i(\mathbf{p}\cdot\mathbf{x}+\mathbf{p}'\cdot\mathbf{x}')}\\ &= i\delta^{(3)}(\mathbf{x}-\mathbf{x}').\end{aligned} \tag{2.30}$$

(If computations such as this one and the next are unfamiliar to you, please work them out carefully; they are quite easy after a little practice, and are fundamental to the formalism of the next two chapters.)

We are now ready to express the Hamiltonian in terms of ladder operators. Starting from its expression (2.8) in terms of ϕ and π, we have

$$\begin{aligned}H &= \int d^3x \int \frac{d^3p\,d^3p'}{(2\pi)^6}e^{i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}}\left\{-\frac{\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}{4}\left(a_{\mathbf{p}} - a^\dagger_{-\mathbf{p}}\right)\left(a_{\mathbf{p}'} - a^\dagger_{-\mathbf{p}'}\right)\right.\\ &\qquad\qquad \left. + \frac{-\mathbf{p}\cdot\mathbf{p}' + m^2}{4\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}\left(a_{\mathbf{p}} + a^\dagger_{-\mathbf{p}}\right)\left(a_{\mathbf{p}'} + a^\dagger_{-\mathbf{p}'}\right)\right\}\\ &= \int \frac{d^3p}{(2\pi)^3}\,\omega_{\mathbf{p}}\left(a^\dagger_{\mathbf{p}}a_{\mathbf{p}} + \tfrac{1}{2}\left[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}}\right]\right).\end{aligned} \tag{2.31}$$

The second term is proportional to $\delta(0)$, an infinite c-number. It is simply the sum over all modes of the zero-point energies $\omega_{\mathbf{p}}/2$, so its presence is completely expected, if somewhat disturbing. Fortunately, this infinite energy shift cannot be detected experimentally, since experiments measure only energy *differences* from the ground state of H. We will therefore ignore this infinite constant term in all of our calculations. It is possible that this energy shift of the ground state could create a problem at a deeper level in the theory; we will discuss this matter in the Epilogue.

Using this expression for the Hamiltonian in terms of $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$, it is easy to evaluate the commutators

$$[H, a_{\mathbf{p}}^\dagger] = \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger; \qquad [H, a_{\mathbf{p}}] = -\omega_{\mathbf{p}} a_{\mathbf{p}}. \tag{2.32}$$

We can now write down the spectrum of the theory, just as for the harmonic oscillator. The state $|0\rangle$ such that $a_{\mathbf{p}}|0\rangle = 0$ for all $\mathbf{p}$ is the ground state or *vacuum*, and has $E = 0$ after we drop the infinite constant in (2.31). All other energy eigenstates can be built by acting on $|0\rangle$ with creation operators. In general, the state $a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger \cdots |0\rangle$ is an eigenstate of H with energy $\omega_{\mathbf{p}} + \omega_{\mathbf{q}} + \cdots$. These states exhaust the spectrum.

Having found the spectrum of the Hamiltonian, let us try to interpret its eigenstates. From (2.19) and a calculation similar to (2.31) we can write down the total momentum operator,

$$\mathbf{P} = -\int d^3x\, \pi(\mathbf{x}) \nabla\phi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\, \mathbf{p}\, a_{\mathbf{p}}^\dagger a_{\mathbf{p}}. \tag{2.33}$$

So the operator $a_{\mathbf{p}}^\dagger$ creates momentum $\mathbf{p}$ and energy $\omega_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2 + m^2}$. Similarly, the state $a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger \cdots |0\rangle$ has momentum $\mathbf{p} + \mathbf{q} + \cdots$. It is quite natural to call these excitations *particles*, since they are discrete entities that have the proper relativistic energy-momentum relation. (By a *particle* we do not mean something that must be localized in space; $a_{\mathbf{p}}^\dagger$ creates particles in momentum eigenstates.) From now on we will refer to $\omega_{\mathbf{p}}$ as $E_{\mathbf{p}}$ (or simply E), since it really is the energy of a particle. Note, by the way, that the energy is always positive: $E_{\mathbf{p}} = +\sqrt{|\mathbf{p}|^2 + m^2}$.

This formalism also allows us to determine the statistics of our particles. Consider the two-particle state $a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger |0\rangle$. Since $a_{\mathbf{p}}^\dagger$ and $a_{\mathbf{q}}^\dagger$ commute, this state is identical to the state $a_{\mathbf{q}}^\dagger a_{\mathbf{p}}^\dagger |0\rangle$ in which the two particles are interchanged. Moreover, a single mode $\mathbf{p}$ can contain arbitrarily many particles (just as a simple harmonic oscillator can be excited to arbitrarily high levels). Thus we conclude that Klein-Gordon particles obey *Bose-Einstein statistics*.

We naturally choose to normalize the vacuum state so that $\langle 0|0\rangle = 1$. The one-particle states $|\mathbf{p}\rangle \propto a_{\mathbf{p}}^\dagger |0\rangle$ will also appear quite often, and it is worthwhile to adopt a convention for their normalization. The simplest normalization $\langle \mathbf{p}|\mathbf{q}\rangle = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q})$ (which many books use) is not Lorentz invariant, as we can demonstrate by considering the effect of a boost in the 3-direction. Under such a boost we have $p_3' = \gamma(p_3 + \beta E)$, $E' = \gamma(E + \beta p_3)$. Using the delta function identity

$$\delta(f(x) - f(x_0)) = \frac{1}{|f'(x_0)|}\delta(x - x_0), \tag{2.34}$$

we can compute

$$\begin{aligned}
\delta^{(3)}(\mathbf{p} - \mathbf{q}) &= \delta^{(3)}(\mathbf{p}' - \mathbf{q}') \cdot \frac{dp_3'}{dp_3} \\
&= \delta^{(3)}(\mathbf{p}' - \mathbf{q}')\gamma\Big(1 + \beta\frac{dE}{dp_3}\Big) \\
&= \delta^{(3)}(\mathbf{p}' - \mathbf{q}')\frac{\gamma}{E}(E + \beta p_3) \\
&= \delta^{(3)}(\mathbf{p}' - \mathbf{q}')\frac{E'}{E}.
\end{aligned}$$

The problem is that volumes are not invariant under boosts; a box whose volume is V in its rest frame has volume V/γ in a boosted frame, due to Lorentz contraction. But from the above calculation, we see that the quantity $E_{\mathbf{p}}\delta^{(3)}(\mathbf{p} - \mathbf{q})$ *is* Lorentz invariant. We therefore define

$$|\mathbf{p}\rangle = \sqrt{2E_{\mathbf{p}}}\, a_{\mathbf{p}}^\dagger |0\rangle, \tag{2.35}$$

so that

$$\langle \mathbf{p}|\mathbf{q}\rangle = 2E_{\mathbf{p}}(2\pi)^3\delta^{(3)}(\mathbf{p} - \mathbf{q}). \tag{2.36}$$

(The factor of 2 is unnecessary, but is convenient because of the factor of 2 in Eq. (2.25).)

On the Hilbert space of quantum states, a Lorentz transformation Λ will be implemented as some unitary operator $U(\Lambda)$. Our normalization condition (2.35) then implies that

$$U(\Lambda)\,|\mathbf{p}\rangle = |\Lambda\mathbf{p}\rangle\,. \tag{2.37}$$

If we prefer to think of this transformation as acting on the operator $a_{\mathbf{p}}^\dagger$, we can also write

$$U(\Lambda)\, a_{\mathbf{p}}^\dagger\, U^{-1}(\Lambda) = \sqrt{\frac{E_{\Lambda\mathbf{p}}}{E_{\mathbf{p}}}}\, a_{\Lambda\mathbf{p}}^\dagger. \tag{2.38}$$

With this normalization we must divide by $2E_{\mathbf{p}}$ in other places. For example, the completeness relation for the one-particle states is

$$(\mathbf{1})_{\text{1-particle}} = \int \frac{d^3p}{(2\pi)^3}\, |\mathbf{p}\rangle \frac{1}{2E_{\mathbf{p}}} \langle\mathbf{p}|\,, \tag{2.39}$$

where the operator on the left is simply the identity within the subspace of one-particle states, and zero in the rest of the Hilbert space. Integrals of this form will occur quite often; in fact, the integral

$$\int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} = \int \frac{d^4p}{(2\pi)^4}(2\pi)\delta(p^2 - m^2)\Big|_{p^0>0} \tag{2.40}$$

is a Lorentz-invariant 3-momentum integral, in the sense that if $f(p)$ is Lorentz-invariant, so is $\int d^3p\, f(p)/(2E_{\mathbf{p}})$. The integration can be thought of

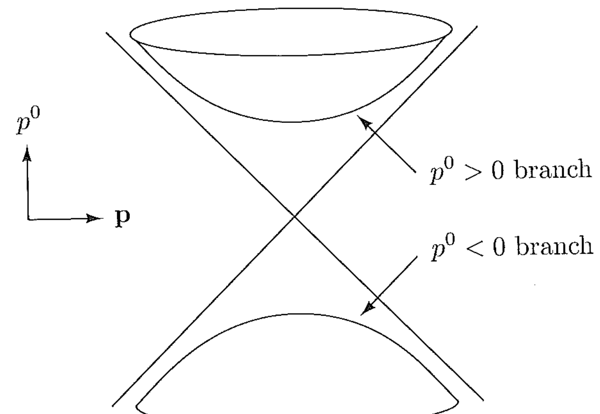


Figure 2.2. The Lorentz-invariant 3-momentum integral is over the upper branch of the hyperboloid $p^2 = m^2$.

as being over the $p^0 > 0$ branch of the hyperboloid $p^2 = m^2$ in 4-momentum space (see Fig. 2.2).

Finally let us consider the interpretation of the state $\phi(\mathbf{x})|0\rangle$. From the expansion (2.25) we see that

$$\phi(\mathbf{x})|0\rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} e^{-i\mathbf{p}\cdot\mathbf{x}} |\mathbf{p}\rangle \tag{2.41}$$

is a linear superposition of single-particle states that have well-defined momentum. Except for the factor $1/2E_{\mathbf{p}}$, this is the same as the familiar nonrelativistic expression for the eigenstate of position $|\mathbf{x}\rangle$; in fact the extra factor is nearly constant for small (nonrelativistic) $\mathbf{p}$. We will therefore put forward the same interpretation, and claim that the operator $\phi(\mathbf{x})$, acting on the vacuum, *creates a particle at position* $\mathbf{x}$. This interpretation is further confirmed when we compute

$$\begin{aligned}\langle 0|\phi(\mathbf{x})|\mathbf{p}\rangle &= \langle 0| \int \frac{d^3p'}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}'}}} \left(a_{\mathbf{p}'} e^{i\mathbf{p}'\cdot\mathbf{x}} + a^\dagger_{\mathbf{p}'} e^{-i\mathbf{p}'\cdot\mathbf{x}}\right) \sqrt{2E_{\mathbf{p}}}\, a^\dagger_{\mathbf{p}} |0\rangle \\ &= e^{i\mathbf{p}\cdot\mathbf{x}}.\end{aligned} \tag{2.42}$$

We can interpret this as the position-space representation of the single-particle wavefunction of the state $|\mathbf{p}\rangle$, just as in nonrelativistic quantum mechanics $\langle \mathbf{x}|\mathbf{p}\rangle \propto e^{i\mathbf{p}\cdot\mathbf{x}}$ is the wavefunction of the state $|\mathbf{p}\rangle$.

2.4 The Klein-Gordon Field in Space-Time

In the previous section we quantized the Klein-Gordon field in the Schrödinger picture, and interpreted the resulting theory in terms of relativistic particles. In this section we will switch to the Heisenberg picture, where it will be easier to discuss time-dependent quantities and questions of causality. After a few preliminaries, we will return to the question of acausal propagation raised in Section 2.1. We will also derive an expression for the *Klein-Gordon propagator*, a crucial part of the Feynman rules to be developed in Chapter 4.

In the Heisenberg picture, we make the operators ϕ and π time-dependent in the usual way:

$$\phi(x) = \phi(\mathbf{x}, t) = e^{iHt}\phi(\mathbf{x})e^{-iHt}, \tag{2.43}$$

and similarly for $\pi(x) = \pi(\mathbf{x}, t)$. The Heisenberg equation of motion,

$$i\frac{\partial}{\partial t}\mathcal{O} = [\mathcal{O}, H], \tag{2.44}$$

allows us to compute the time dependence of ϕ and π:

$$\begin{aligned} i\frac{\partial}{\partial t}\phi(\mathbf{x}, t) &= \left[\phi(\mathbf{x}, t), \int d^3x' \left\{\tfrac{1}{2}\pi^2(\mathbf{x}', t) + \tfrac{1}{2}\left(\nabla\phi(\mathbf{x}', t)\right)^2 + \tfrac{1}{2}m^2\phi^2(\mathbf{x}', t)\right\}\right] \\ &= \int d^3x' \left(i\delta^{(3)}(\mathbf{x} - \mathbf{x}')\pi(\mathbf{x}', t)\right) \\ &= i\pi(\mathbf{x}, t); \\ i\frac{\partial}{\partial t}\pi(\mathbf{x}, t) &= \left[\pi(\mathbf{x}, t), \int d^3x' \left\{\tfrac{1}{2}\pi^2(\mathbf{x}', t) + \tfrac{1}{2}\phi(\mathbf{x}', t)(-\nabla^2 + m^2)\phi(\mathbf{x}', t)\right\}\right] \\ &= \int d^3x' \left(-i\delta^{(3)}(\mathbf{x} - \mathbf{x}')(-\nabla^2 + m^2)\phi(\mathbf{x}', t)\right) \\ &= -i(-\nabla^2 + m^2)\phi(\mathbf{x}, t). \end{aligned}$$

Combining the two results gives

$$\frac{\partial^2}{\partial t^2}\phi = (\nabla^2 - m^2)\phi, \tag{2.45}$$

which is just the Klein-Gordon equation.

We can better understand the time dependence of $\phi(x)$ and $\pi(x)$ by writing them in terms of creation and annihilation operators. First note that

$$H a_{\mathbf{p}} = a_{\mathbf{p}}(H - E_{\mathbf{p}}),$$

and hence

$$H^n a_{\mathbf{p}} = a_{\mathbf{p}}(H - E_{\mathbf{p}})^n,$$

for any n. A similar relation (with $-$ replaced by $+$) holds for $a^\dagger_{\mathbf{p}}$. Thus we have derived the identities

$$e^{iHt}a_{\mathbf{p}}e^{-iHt} = a_{\mathbf{p}}e^{-iE_{\mathbf{p}}t}, \qquad e^{iHt}a^\dagger_{\mathbf{p}}e^{-iHt} = a^\dagger_{\mathbf{p}}e^{iE_{\mathbf{p}}t}, \tag{2.46}$$

which we can use on expression (2.25) for $\phi(\mathbf{x})$ to find the desired expression for the Heisenberg operator $\phi(x)$, according to (2.43). (We will always use the symbols $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$ to represent the time-independent, Schrödinger-picture ladder operators.) The result is

$$\begin{aligned}\phi(\mathbf{x},t) &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\left(a_{\mathbf{p}}e^{-ip\cdot x}+a_{\mathbf{p}}^\dagger e^{ip\cdot x}\right)\bigg|_{p^0=E_{\mathbf{p}}};\\ \pi(\mathbf{x},t) &= \frac{\partial}{\partial t}\phi(\mathbf{x},t).\end{aligned} \tag{2.47}$$

It is worth mentioning that we can perform the same manipulations with $\mathbf{P}$ instead of H to relate $\phi(\mathbf{x})$ to $\phi(0)$. In analogy with (2.46), one can show

$$e^{-i\mathbf{P}\cdot\mathbf{x}}a_{\mathbf{p}}e^{i\mathbf{P}\cdot\mathbf{x}} = a_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}}, \qquad e^{-i\mathbf{P}\cdot\mathbf{x}}a_{\mathbf{p}}^\dagger e^{i\mathbf{P}\cdot\mathbf{x}} = a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}, \tag{2.48}$$

and therefore

$$\begin{aligned}\phi(x) &= e^{i(Ht-\mathbf{P}\cdot\mathbf{x})}\phi(0)e^{-i(Ht-\mathbf{P}\cdot\mathbf{x})}\\ &= e^{iP\cdot x}\phi(0)e^{-iP\cdot x},\end{aligned} \tag{2.49}$$

where $P^\mu = (H,\mathbf{P})$. (The notation here is confusing but standard. Remember that $\mathbf{P}$ is the momentum operator, whose eigenvalue is the total momentum of the system. On the other hand, $\mathbf{p}$ is the momentum of a single Fourier mode of the field, which we interpret as the momentum of a particle in that mode. For a one-particle state of well-defined momentum, $\mathbf{p}$ is the eigenvalue of $\mathbf{P}$.)

Equation (2.47) makes explicit the dual particle and wave interpretations of the quantum field $\phi(x)$. On the one hand, $\phi(x)$ is written as a Hilbert space operator, which creates and destroys the particles that are the quanta of field excitation. On the other hand, $\phi(x)$ is written as a linear combination of solutions ($e^{ip\cdot x}$ and $e^{-ip\cdot x}$) of the Klein-Gordon equation. Both signs of the time dependence in the exponential appear: We find both e^{-ip^0t} and e^{+ip^0t}, although p^0 is always positive. If these were single-particle wavefunctions, they would correspond to states of positive and negative energy; let us refer to them more generally as *positive-* and *negative-frequency* modes. The connection between the particle creation operators and the waveforms displayed here is always valid for free quantum fields: A positive-frequency solution of the field equation has as its coefficient the operator that *destroys* a particle in that single-particle wavefunction. A negative-frequency solution of the field equation, being the Hermitian conjugate of a positive-frequency solution, has as its coefficient the operator that *creates* a particle in that positive-energy single-particle wavefunction. In this way, the fact that relativistic wave equations have both positive- and negative-frequency solutions is reconciled with the requirement that a sensible quantum theory contain only positive excitation energies.

Causality

Now let us return to the question of causality raised at the beginning of this chapter. In our present formalism, still working in the Heisenberg picture, the amplitude for a particle to propagate from y to x is $\langle 0|\,\phi(x)\phi(y)\,|0\rangle$. We will call this quantity $D(x-y)$. Each operator ϕ is a sum of a and $a^\dagger$ operators, but only the term $\langle 0|\,a_{\mathbf{p}}a_{\mathbf{q}}^\dagger\,|0\rangle = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})$ survives in this expression. It is easy to check that we are left with

$$D(x-y) = \langle 0|\,\phi(x)\phi(y)\,|0\rangle = \int\frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}e^{-ip\cdot(x-y)}. \tag{2.50}$$

We have already argued in (2.40) that integrals of this form are Lorentz invariant. Let us now evaluate this integral for some particular values of $x-y$.

First consider the case where the difference $x-y$ is purely in the time-direction: $x^0-y^0=t$, $\mathbf{x}-\mathbf{y}=0$. (If the interval from y to x is timelike, there is always a frame in which this is the case.) Then we have

$$\begin{aligned}D(x-y) &= \frac{4\pi}{(2\pi)^3}\int_0^\infty dp\,\frac{p^2}{2\sqrt{p^2+m^2}}e^{-i\sqrt{p^2+m^2}t}\\ &= \frac{1}{4\pi^2}\int_m^\infty dE\,\sqrt{E^2-m^2}\,e^{-iEt}\\ &\underset{t\to\infty}{\sim} e^{-imt}.\end{aligned} \tag{2.51}$$

Next consider the case where $x-y$ is purely spatial: $x^0-y^0=0$, $\mathbf{x}-\mathbf{y}=\mathbf{r}$. The amplitude is then

$$\begin{aligned}D(x-y) &= \int\frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}e^{i\mathbf{p}\cdot\mathbf{r}}\\ &= \frac{2\pi}{(2\pi)^3}\int_0^\infty dp\,\frac{p^2}{2E_{\mathbf{p}}}\frac{e^{ipr}-e^{-ipr}}{ipr}\\ &= \frac{-i}{2(2\pi)^2r}\int_{-\infty}^\infty dp\,\frac{p\,e^{ipr}}{\sqrt{p^2+m^2}}.\end{aligned}$$

The integrand, considered as a complex function of p, has branch cuts on the imaginary axis starting at $\pm im$ (see Fig. 2.3). To evaluate the integral we push the contour up to wrap around the upper branch cut. Defining $\rho=-ip$, we obtain

$$\frac{1}{4\pi^2r}\int_m^\infty d\rho\,\frac{\rho\,e^{-\rho r}}{\sqrt{\rho^2-m^2}} \underset{r\to\infty}{\sim} e^{-mr}. \tag{2.52}$$

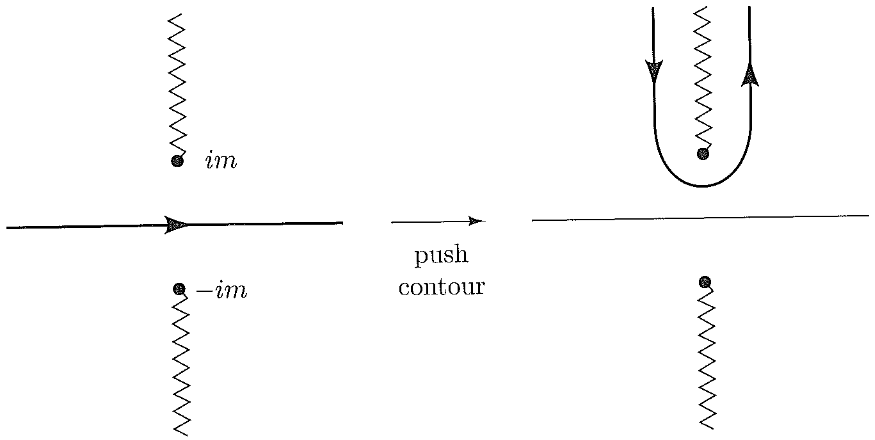


Figure 2.3. Contour for evaluating propagation amplitude $D(x-y)$ over a spacelike interval.

So again we find that outside the light-cone, the propagation amplitude is exponentially vanishing but nonzero.

To really discuss causality, however, we should ask not whether particles can propagate over spacelike intervals, but whether a *measurement* performed at one point can affect a measurement at another point whose separation from the first is spacelike. The simplest thing we could try to measure is the field $\phi(x)$, so we should compute the commutator $[\phi(x), \phi(y)]$; if this commutator vanishes, one measurement cannot affect the other. In fact, if the commutator vanishes for $(x-y)^2 < 0$, causality is preserved quite generally, since commutators involving any function of $\phi(x)$, including $\pi(x) = \partial\phi/\partial t$, would also have to vanish. Of course we know from Eq. (2.20) that the commutator vanishes for $x^0 = y^0$; now let's do the more general computation:

$$\begin{aligned}
[\phi(x), \phi(y)] &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \int \frac{d^3q}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{q}}}} \\
&\qquad \times \Big[\big(a_{\mathbf{p}} e^{-ip\cdot x} + a_{\mathbf{p}}^\dagger e^{ip\cdot x}\big), \big(a_{\mathbf{q}} e^{-iq\cdot y} + a_{\mathbf{q}}^\dagger e^{iq\cdot y}\big) \Big] \\
&= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \big(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)} \big) \\
&= D(x-y) - D(y-x). \qquad (2.53)
\end{aligned}$$

When $(x-y)^2 < 0$, we can perform a Lorentz transformation on the second term (since each term is separately Lorentz invariant), taking $(x-y) \to -(x-y)$, as shown in Fig. 2.4. The two terms are therefore equal and cancel to give zero; causality is preserved. Note that if $(x-y)^2 > 0$ there is no continuous Lorentz transformation that takes $(x-y) \longrightarrow -(x-y)$. In this case, by Eq. (2.51), the amplitude is (fortunately) nonzero, roughly $(e^{-imt} - e^{imt})$ for the special case $\mathbf{x} - \mathbf{y} = 0$. Thus we conclude that no measurement in the

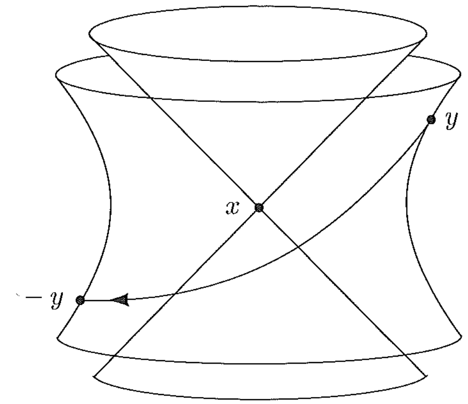


Figure 2.4. When $x-y$ is spacelike, a continuous Lorentz transformation can take $(x-y)$ to $-(x-y)$.

Klein-Gordon theory can affect another measurement outside the light-cone.

Causality is maintained in the Klein-Gordon theory just as suggested at the end of Section 2.1. To understand this mechanism properly, however, we should broaden the context of our discussion to include a *complex* Klein-Gordon field, which has distinct particle and antiparticle excitations. As was mentioned in the discussion of Eq. (2.15), we can add a conserved charge to the Klein-Gordon theory by considering the field $\phi(x)$ to be complex- rather than real-valued. When the complex scalar field theory is quantized (see Problem 2.2), $\phi(x)$ will create positively charged particles and destroy negatively charged ones, while $\phi^\dagger(x)$ will perform the opposite operations. Then the commutator $[\phi(x), \phi^\dagger(y)]$ will have nonzero contributions, which must delicately cancel outside the light-cone to preserve causality. The two contributions have the spacetime interpretation of the two terms in (2.53), but with charges attached. The first term will represent the propagation of a negatively charged particle from y to x. The second term will represent the propagation of a positively charged particle from x to y. In order for these two processes to be present and give canceling amplitudes, both of these particles must exist, and they must have the same mass. In quantum field theory, then, causality requires that every particle have a corresponding antiparticle with the same mass and opposite quantum numbers (in this case electric charge). For the real-valued Klein-Gordon field, the particle is its own antiparticle.

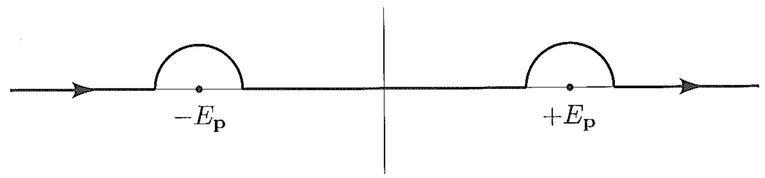
The Klein-Gordon Propagator

Let us study the commutator $[\phi(x), \phi(y)]$ a little further. Since it is a c-number, we can write $[\phi(x), \phi(y)] = \langle 0| [\phi(x), \phi(y)] |0\rangle$. This can be rewritten as a four-dimensional integral as follows, assuming for now that $x^0 > y^0$:

$$\langle 0| [\phi(x), \phi(y)] |0\rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \big(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)} \big)$$

$$
\begin{aligned}
&= \int \frac{d^3p}{(2\pi)^3} \left\{ \frac{1}{2E_\mathbf{p}} e^{-ip\cdot(x-y)}\Big|_{p^0=E_\mathbf{p}} \right. \\
&\qquad\qquad \left. + \frac{1}{-2E_\mathbf{p}} e^{-ip\cdot(x-y)}\Big|_{p^0=-E_\mathbf{p}} \right\} \\
&\underset{x^0>y^0}{=} \int \frac{d^3p}{(2\pi)^3} \int \frac{dp^0}{2\pi i} \frac{-1}{p^2-m^2} e^{-ip\cdot(x-y)}. \qquad (2.54)
\end{aligned}
$$

In the last step the p^0 integral is to be performed along the following contour:



For $x^0 > y^0$ we can close the contour below, picking up both poles to obtain the previous line of (2.54). For $x^0 < y^0$ we may close the contour above, giving zero. Thus the last line of (2.54), together with the prescription for going around the poles, is an expression for what we will call

$$D_R(x-y) \equiv \theta(x^0-y^0)\, \langle 0| \left[\phi(x), \phi(y)\right] |0\rangle . \qquad (2.55)$$

To understand this quantity better, let's do another computation:

$$
\begin{aligned}
(\partial^2+m^2)D_R(x-y) &= \left(\partial^2\theta(x^0-y^0)\right) \langle 0| \left[\phi(x),\phi(y)\right] |0\rangle \\
&\quad + 2\left(\partial_\mu\theta(x^0-y^0)\right)\left(\partial^\mu \langle 0| \left[\phi(x),\phi(y)\right] |0\rangle\right) \\
&\quad + \theta(x^0-y^0)\,(\partial^2+m^2)\, \langle 0| \left[\phi(x),\phi(y)\right] |0\rangle \\
&= -\delta(x^0-y^0)\, \langle 0| \left[\pi(x),\phi(y)\right] |0\rangle \\
&\quad + 2\delta(x^0-y^0)\, \langle 0| \left[\pi(x),\phi(y)\right] |0\rangle + 0 \\
&= -i\delta^{(4)}(x-y). \qquad (2.56)
\end{aligned}
$$

This says that $D_R(x-y)$ is a Green's function of the Klein-Gordon operator. Since it vanishes for $x^0 < y^0$, it is the *retarded* Green's function.

If we had not already derived expression (2.54), we could find it by Fourier transformation. Writing

$$D_R(x-y) = \int \frac{d^4p}{(2\pi)^4}\, e^{-ip\cdot(x-y)}\, \widetilde{D}_R(p), \qquad (2.57)$$

we obtain an algebraic expression for $\widetilde{D}_R(p)$:

$$(-p^2+m^2)\widetilde{D}_R(p) = -i.$$

Thus we immediately arrive at the result

$$D_R(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2-m^2}\, e^{-ip\cdot(x-y)}. \qquad (2.58)$$

The p^0-integral of (2.58) can be evaluated according to four different contours, of which that used in (2.54) is only one. In Chapter 4 we will find that a different pole prescription,

is extremely useful; it is called the *Feynman prescription.* A convenient way to remember it is to write

$$D_F(x-y) \equiv \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2-m^2+i\epsilon}\, e^{-ip\cdot(x-y)}, \qquad (2.59)$$

since the poles are then at $p^0 = \pm(E_\mathbf{p} - i\epsilon)$, displaced properly above and below the real axis. When $x^0 > y^0$ we can perform the p^0 integral by closing the contour below, obtaining exactly the propagation amplitude $D(x-y)$ (2.50). When $x^0 < y^0$ we close the contour above, obtaining the same expression but with x and y interchanged. Thus we have

$$
\begin{aligned}
D_F(x-y) &= \begin{cases} D(x-y) & \text{for } x^0 > y^0 \\ D(y-x) & \text{for } x^0 < y^0 \end{cases} \\
&= \theta(x^0-y^0)\, \langle 0| \phi(x)\phi(y) |0\rangle + \theta(y^0-x^0)\, \langle 0| \phi(y)\phi(x) |0\rangle \\
&\equiv \langle 0| T\phi(x)\phi(y) |0\rangle . \qquad (2.60)
\end{aligned}
$$

The last line defines the "time-ordering" symbol T, which instructs us to place the operators that follow in order with the latest to the left. By applying (∂^2+m^2) to the last line, you can verify directly that D_F is a Green's function of the Klein-Gordon operator.

Equations (2.59) and (2.60) are, from a practical point of view, the most important results of this chapter. The Green's function $D_F(x-y)$ is called the *Feynman propagator* for a Klein-Gordon particle, since it is, after all, a propagation amplitude. Indeed, the Feynman propagator will turn out to be part of the Feynman rules: $D_F(x-y)$ (or $\widetilde{D}_F(p)$) is the expression that we will attach to internal lines of Feynman diagrams, representing the propagation of virtual particles.

Nevertheless we are still a long way from being able to do any real calculations, since so far we have talked only about the *free* Klein-Gordon theory, where the field equation is linear and there are no interactions. Individual particles live in their isolated modes, oblivious to each others' existence and to the existence of any other species of particles. In such a theory there is no hope of making any observations, by scattering or any other means. On the other hand, the formalism we have developed is extremely important, since the free theory forms the basis for doing perturbative calculations in the interacting theory.

Particle Creation by a Classical Source

There is one type of interaction, however, that we are already equipped to handle. Consider a Klein-Gordon field coupled to an external, classical source field $j(x)$. That is, consider the field equation

$$(\partial^2 + m^2)\phi(x) = j(x), \tag{2.61}$$

where $j(x)$ is some fixed, known function of space and time that is nonzero only for a finite time interval. If we start in the vacuum state, what will we find after $j(x)$ has been turned on and off again?

The field equation (2.61) follows from the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2 + j(x)\phi(x). \tag{2.62}$$

But if $j(x)$ is turned on for only a finite time, it is easiest to solve the problem using the field equation directly. Before $j(x)$ is turned on, $\phi(x)$ has the form

$$\phi_0(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\left(a_{\mathbf{p}}e^{-ip\cdot x} + a_{\mathbf{p}}^\dagger e^{ip\cdot x}\right).$$

If there were no source, this would be the solution for all time. With a source, the solution of the equation of motion can be constructed using the retarded Green's function:

$$\begin{aligned}\phi(x) &= \phi_0(x) + i\int d^4y\, D_R(x-y)j(y)\\ &= \phi_0(x) + i\int d^4y\int\frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}\,\theta(x^0-y^0)\\ &\qquad\times\left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right)j(y).\end{aligned} \tag{2.63}$$

If we wait until all of j is in the past, the theta function equals 1 in the whole domain of integration. Then $\phi(x)$ involves only the Fourier transform of j,

$$\tilde{j}(p) = \int d^4y\, e^{ip\cdot y}j(y),$$

evaluated at 4-momenta p such that $p^2 = m^2$. It is natural to group the positive-frequency terms together with $a_{\mathbf{p}}$ and the negative-frequency terms with $a_{\mathbf{p}}^\dagger$; this yields the expression

$$\phi(x) = \int\frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\left\{\left(a_{\mathbf{p}} + \frac{i}{\sqrt{2E_{\mathbf{p}}}}\tilde{j}(p)\right)e^{-ip\cdot x} + \text{h.c.}\right\}. \tag{2.64}$$

You can now guess (or compute) the form of the Hamiltonian after $j(x)$ has acted: Just replace $a_{\mathbf{p}}$ with $(a_{\mathbf{p}} + i\tilde{j}(p)/\sqrt{2E_{\mathbf{p}}})$ to obtain

$$H = \int\frac{d^3p}{(2\pi)^3}E_{\mathbf{p}}\left(a_{\mathbf{p}}^\dagger - \frac{i}{\sqrt{2E_{\mathbf{p}}}}\tilde{j}^*(p)\right)\left(a_{\mathbf{p}} + \frac{i}{\sqrt{2E_{\mathbf{p}}}}\tilde{j}(p)\right).$$

The energy of the system after the source has been turned off is

$$\langle 0|\,H\,|0\rangle = \int\frac{d^3p}{(2\pi)^3}\,\frac{1}{2}|\tilde{j}(p)|^2, \tag{2.65}$$

where $|0\rangle$ still denotes the ground state of the free theory. We can interpret these results in terms of particles by identifying $|\tilde{j}(p)|^2/2E_{\mathbf{p}}$ as the probability density for creating a particle in the mode p. Then the total number of particles produced is

$$\int dN = \int\frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}|\tilde{j}(p)|^2. \tag{2.66}$$

Only those Fourier components of $j(x)$ that are in resonance with on-mass-shell (i.e., $p^2 = m^2$) Klein-Gordon waves are effective at creating particles.

We will return to this subject in Problem 4.1. In Chapter 6 we will study the analogous problem of photon creation by an accelerated electron (bremsstrahlung).

Problems

2.1 Classical electromagnetism (with no sources) follows from the action

$$S = \int d^4x\left(-\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}\right), \qquad \text{where } F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

(a) Derive Maxwell's equations as the Euler-Lagrange equations of this action, treating the components $A_\mu(x)$ as the dynamical variables. Write the equations in standard form by identifying $E^i = -F^{0i}$ and $\epsilon^{ijk}B^k = -F^{ij}$.

(b) Construct the energy-momentum tensor for this theory. Note that the usual procedure does not result in a symmetric tensor. To remedy that, we can add to $T^{\mu\nu}$ a term of the form $\partial_\lambda K^{\lambda\mu\nu}$, where $K^{\lambda\mu\nu}$ is antisymmetric in its first two indices. Such an object is automatically divergenceless, so

$$\widehat{T}^{\mu\nu} = T^{\mu\nu} + \partial_\lambda K^{\lambda\mu\nu}$$

is an equally good energy-momentum tensor with the same globally conserved energy and momentum. Show that this construction, with

$$K^{\lambda\mu\nu} = F^{\mu\lambda}A^\nu,$$

leads to an energy-momentum tensor $\widehat{T}$ that is symmetric and yields the standard formulae for the electromagnetic energy and momentum densities:

$$\mathcal{E} = \tfrac{1}{2}(E^2 + B^2); \qquad \mathbf{S} = \mathbf{E}\times\mathbf{B}.$$

2.2 **The complex scalar field.** Consider the field theory of a complex-valued scalar field obeying the Klein-Gordon equation. The action of this theory is

$$S = \int d^4x\,(\partial_\mu\phi^*\partial^\mu\phi - m^2\phi^*\phi).$$

It is easiest to analyze this theory by considering $\phi(x)$ and $\phi^*(x)$, rather than the real and imaginary parts of $\phi(x)$, as the basic dynamical variables.

(a) Find the conjugate momenta to $\phi(x)$ and $\phi^*(x)$ and the canonical commutation relations. Show that the Hamiltonian is

$$H = \int d^3x\,(\pi^*\pi + \nabla\phi^*\cdot\nabla\phi + m^2\phi^*\phi).$$

Compute the Heisenberg equation of motion for $\phi(x)$ and show that it is indeed the Klein-Gordon equation.

(b) Diagonalize H by introducing creation and annihilation operators. Show that the theory contains two sets of particles of mass m.

(c) Rewrite the conserved charge

$$Q = \int d^3x\,\frac{i}{2}(\phi^*\pi^* - \pi\phi)$$

in terms of creation and annihilation operators, and evaluate the charge of the particles of each type.

(d) Consider the case of two complex Klein-Gordon fields with the same mass. Label the fields as $\phi_a(x)$, where $a = 1, 2$. Show that there are now four conserved charges, one given by the generalization of part (c), and the other three given by

$$Q^i = \int d^3x\,\frac{i}{2}(\phi_a^*(\sigma^i)_{ab}\pi_b^* - \pi_a(\sigma^i)_{ab}\phi_b),$$

where σ^i are the Pauli sigma matrices. Show that these three charges have the commutation relations of angular momentum ($SU(2)$). Generalize these results to the case of n identical complex scalar fields.

2.3 Evaluate the function

$$\langle 0|\,\phi(x)\phi(y)\,|0\rangle = D(x-y) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}e^{-ip\cdot(x-y)},$$

for $(x-y)$ spacelike so that $(x-y)^2 = -r^2$, explicitly in terms of Bessel functions.

Chapter 3

The Dirac Field

Having exhaustively treated the simplest relativistic field equation, we now move on to the second simplest, the Dirac equation. You may already be familiar with the Dirac equation in its original incarnation, that is, as a single-particle quantum-mechanical wave equation.* In this chapter our viewpoint will be quite different. First we will rederive the Dirac equation as a *classical* relativistic field equation, with special emphasis on its relativistic invariance. Then, in Section 3.5, we will quantize the Dirac field in a manner similar to that used for the Klein-Gordon field.

3.1 Lorentz Invariance in Wave Equations

First we must address a question that we swept over in Chapter 2: What do we mean when we say that an equation is "relativistically invariant"? A reasonable definition is the following: If ϕ is a field or collection of fields and $\mathcal{D}$ is some differential operator, then the statement "$\mathcal{D}\phi = 0$ is relativistically invariant" means that if $\phi(x)$ satisfies this equation, and we perform a rotation or boost to a different frame of reference, then the transformed field, in the new frame of reference, satisfies the same equation. Equivalently, we can imagine physically rotating or boosting all particles or fields by a common angle or velocity; again, the equation $\mathcal{D}\phi = 0$ should be true after the transformation. We will adopt this "active" point of view toward transformations in the following analysis.

The Lagrangian formulation of field theory makes it especially easy to discuss Lorentz invariance. An equation of motion is automatically Lorentz invariant by the above definition if it follows from a Lagrangian that is a Lorentz *scalar*. This is an immediate consequence of the principle of least action: If boosts leave the Lagrangian unchanged, the boost of an extremum in the action will be another extremum.

*This subject is covered, for example, in Schiff (1968), Chapter 13; Baym (1969), Chapter 23; Sakurai (1967), Chapter 3. Although the present chapter is self-contained, we recommend that you also study the single-particle Dirac equation at some point.