2 Lorentz invariance and second quantization

In the previous chapter, we saw that by treating each mode of electromagnetic radiation in a cavity as a simple harmonic oscillator, we can derive Einstein's relation between the coefficients of induced and spontaneous emission without resorting to statistical mechanics. This was our first calculation in quantum electrodynamics (QED). It is not a coincidence that the harmonic oscillator played an important role. After all, electromagnetic waves oscillate harmonically. In this chapter we will review special relativity and the simple harmonic oscillator and show how they are connected. This leads naturally to the notion of **second quantization**, which is a poorly chosen phrase used to describe the canonical quantization of relativistic fields.

It is worth mentioning at this point that there are two ways commonly used to quantize a field theory, both of which are covered in depth in this book. The first is canonical quantization. This is historically how quantum field theory was understood, and closely follows what you learned in quantum mechanics. The second way is called the *Feynman path integral*. Path integrals are more concise, more general, and certainly more formal, but when using path integrals it is sometimes hard to understand physically what you are calculating. It really is necessary to understand both ways. Some calculations, such as the LSZ formula which relates scattering amplitudes to correlation function (see Chapter 6), require the canonical approach, while other calculations, such as non-perturbative quantum chromodynamics (see Chapter 25), require path integrals. There are other ways to perform quantum field theory calculations, for example using old-fashioned perturbation theory (Chapter 4), or using Schwinger proper time (Chapter 33). Learning all of these approaches will give you a comprehensive picture of how and why quantum field theory works. We start with canonical quantization, as it provides the gentlest introduction to quantum field theory.

From now on we will set $\hbar = c = 1$. This gives all quantities dimensions of mass to some power (see Appendix A).

2.1 Lorentz invariance

Quantum field theory is the result of combining quantum mechanics with special relativity. Special relativity is relevant when velocities are a reasonable fraction of the speed of light, $v \sim 1$. In this limit, a new symmetry emerges: Lorentz invariance. A system is Lorentz invariant if it is symmetric under the Lorentz group, which is the generalization of the rotation group to include both rotations and boosts.

Normally, the more symmetric a system, the easier it is to solve problems. For example, solving the Schrödinger equation with a spherically symmetric potential (as in the hydrogen atom) is much easier than solving it with a cylindrically symmetric potential (such as for the hydrogen molecule). So why is quantum field theory so much harder than quantum mechanics? The answer, as Sidney Coleman put it, is because $E = mc^2$. This famous relation holds for particles at rest. When particles move relativistically, their kinetic energy is comparable to or exceeds their rest mass, $E_{\text{kin}} \gtrsim m$, which is only a factor of 2 away from the threshold for producing two particles. Thus, there is no regime in which the relativistic corrections of order v/c are relevant, but the effect from producing new particles is not.

2.1.1 Rotations

Lorentz invariance is symmetry under rotations and boosts. If you get confused, focus on perfecting your understanding of rotations alone. Then, consider boosts as a generalization.

Rotations should be extremely familiar to you, and they are certainly more intuitive than boosts. Under two-dimensional (2D) rotations, a vector (x, y) transforms as

$$x \to x\cos\theta + y\sin\theta, \tag{2.1}$$

$$y \to -x\sin\theta + y\cos\theta. \tag{2.2}$$

We can write this as

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x\cos\theta + y\sin\theta \\ -x\sin\theta + y\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{2.3}$$

or as

$$x_i \to R_{ij}x_j, \quad x_i = \begin{pmatrix} x \\ y \end{pmatrix}, \quad i = 1, 2. \tag{2.4}$$

When an index appears twice, as in $R_{ij}x_j$, that index should be summed over (the **Einstein summation convention**), so $R_{ij}x_j = R_{i1}x_1 + R_{i2}x_2$. This is known as a **contraction**.

Technically, we should write $x_i = R_i{}^j x_j$. However, having upper and lower indices on the same object makes expressions difficult to read, so we will often just lower or raise all the indices. We will be careful about the index position if it is ever ambiguous. For the row vector,

$$x^i = \begin{pmatrix} x & y \end{pmatrix} \to \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = x^j \left(R^T\right)^{ji}. \tag{2.5}$$

Note that $R^T = R^{-1}$. That is,

$$\left(R^T\right)_{ij} R_{jk} = \delta_{ik} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{ij} = \mathbb{1}_{ij} \tag{2.6}$$

or equivalently,

$$R^T R = \mathbb{1}. \tag{2.7}$$

This property (orthogonality) along with R preserving orientation ($\det R = 1$) is enough to characterize R as a rotation. This algebraic characterization in Eq. (2.7) is a much more useful definition of the group than the explicit form of the rotation matrices as a function of θ. The group of 2D rotations is also called the **special orthogonal group** SO(2). The group of 3D rotations is called SO(3).

If we contract the upper and lower indices of a vector, we find

$$x^i x_i = (x, y) \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2. \tag{2.8}$$

This is just the norm of the vector x_i and is invariant under rotations. To see that, note that under a rotation

$$x^i x_i \to \left(x^i R_{ij}^T\right)\left(R_{jk} x_k\right) = x^i \delta_{ik} x_k = x^i x_i, \tag{2.9}$$

since $R^T = R^{-1}$. In fact, another way to define the rotation group is as the set of linear transformations on $\mathbb{R}^n$ preserving the inner product $x^i x_i = \delta_{ij} x^i x^j$:

$$R_{ki} R_{lj} \delta_{kl} = [(R^T)\mathbb{1}(R)]_{ij} = (R^T R)_{ij} = \delta_{ij}, \tag{2.10}$$

which you can check explicitly using Eq. (2.3).

2.1.2 Lorentz transformations

Lorentz transformations work exactly like rotations, except with some minus signs here and there. Instead of preserving $r^2 = x^2 + y^2 + z^2$ they preserve $s^2 \equiv t^2 - x^2 - y^2 - z^2$. Instead of 3-vectors $v^i = (x, y, z)$ we use 4-vectors $x^\mu = (t, x, y, z)$. We generally use Greek indices for 4-vectors and Latin indices for 3-vectors. We write x^0 for the time component of a 4-vector.

Lorentz transformations acting on 4-vectors are matrices Λ satisfying

$$\Lambda^T g \Lambda = g = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \tag{2.11}$$

In this and future matrices, empty entries are 0. $g_{\mu\nu}$ is known as the **Minkowski metric**. Sometimes we write $\eta_{\mu\nu}$ for this metric, with $g_{\mu\nu}$ reserved for a general metric, as in general relativity. But outside of quantum gravity contexts, which will be clear when we encounter them, taking $g_{\mu\nu} = \eta_{\mu\nu}$ will cause no confusion in quantum field theory. Equation (2.11) says that Lorentz transformations preserve the Minkowskian inner product:

$$x^\mu x_\mu = g_{\mu\nu} x^\mu x^\nu = t^2 - x^2 - y^2 - z^2. \tag{2.12}$$

A rotation around the z axis leaves $x^2 + y^2$ invariant while a boost in the z direction leaves $t^2 - z^2$ invariant. So, instead of being sines and cosines, which satisfy $\cos^2\theta + \sin^2\theta = 1$, boosts are made from hyperbolic sines and cosines, which satisfy $\cosh^2\beta - \sinh^2\beta = 1$.

The Lorentz group is the most general set of transformations preserving the Minkowski metric. Up to some possible discrete transformations (see Section 2.1.3 below), a general Lorentz transformation can be written as a product of rotations around the x, y or z axes:

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cos\theta_x & \sin\theta_x \\ & & -\sin\theta_x & \cos\theta_x \end{pmatrix}, \begin{pmatrix} 1 & & & \\ & \cos\theta_y & & -\sin\theta_y \\ & & 1 & \\ & \sin\theta_y & & \cos\theta_y \end{pmatrix}, \begin{pmatrix} 1 & & & \\ & \cos\theta_z & \sin\theta_z & \\ & -\sin\theta_z & \cos\theta_z & \\ & & & 1 \end{pmatrix} \tag{2.13}$$

and boosts in the x, y or z direction:

$$\begin{pmatrix} \cosh\beta_x & \sinh\beta_x & & \\ \sinh\beta_x & \cosh\beta_x & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \begin{pmatrix} \cosh\beta_y & & \sinh\beta_y & \\ & 1 & & \\ \sinh\beta_y & & \cosh\beta_y & \\ & & & 1 \end{pmatrix} \cdot \begin{pmatrix} \cosh\beta_z & & & \sinh\beta_z \\ & 1 & & \\ & & 1 & \\ \sinh\beta_z & & & \cosh\beta_z \end{pmatrix}. \tag{2.14}$$

The θ_i are ordinary rotation angles around the i axis, with $0 \leq \theta_i < 2\pi$, and the β_i are hyperbolic angles sometimes called **rapidities**, with $-\infty < \beta_i < \infty$. Note that these matrices do not commute, so the order in which we do the rotations and boosts is important. We will rarely need an actual matrix representation of the group elements like this, but it is helpful to see.

To relate the β_i to something useful, such as velocity, recall that for velocities $v \ll 1$ well below the speed of light, a boost should reduce to a Galilean transformation $x \to x + vt$. The unique transformations that preserve $t^2 - x^2$ and reduce to the Galilean transformations at small v are

$$x \to \frac{x + vt}{\sqrt{1 - v^2}}, \quad t \to \frac{t + vx}{\sqrt{1 - v^2}}. \tag{2.15}$$

Thus we can identify

$$\cosh\beta_x = \frac{1}{\sqrt{1 - v^2}}, \quad \sinh\beta_x = \frac{v}{\sqrt{1 - v^2}}. \tag{2.16}$$

These equations relate boosts to ordinary velocity. In particular, $\beta_x = v$ to leading order in v.

Scalar fields are functions of space-time that are Lorentz invariant. That is, under an arbitrary Lorentz transformation the field does not change:

$$\phi(x) \to \phi(x). \tag{2.17}$$

Sometimes the notation $\phi(x^\mu) \to \phi((\Lambda^{-1})^\mu_\nu\, x^\nu)$ is used, which makes it seem like the scalar field is changing in some way. It is not. While our definitions of x^μ change in different frames $x^\mu \to \Lambda^\mu_\nu x^\nu$, the space-time point labeled by x^μ is fixed. That equations are invariant under relabeling of coordinates tells us absolutely nothing about nature. The physical content of Lorentz invariance is that nature has a symmetry under which scalar fields do not transform. Take, for example, the temperature of a fluid, which can vary from point to point. If we change reference frames, the labels for the points change, but the temperature at each point stays the same. A **scalar** (not scalar field) is just a number. For example, $\hbar$ and 7 and the electric charge e are scalars.

Under Lorentz transformations Λ^μ_ν, **4-vectors** V_μ transform as

$$V^\mu \to \Lambda^\mu_\nu V^\nu. \tag{2.18}$$

This transformation law is the defining property of a 4-vector. If V^μ is not just a number but depends on x, we write $V^\mu(x)$ and call it a **vector field**. Under Lorentz transformations, vector fields transform just like 4-vectors. For a vector field, as for a scalar field, the coordinates of x transform but the space-time point to which they refer is invariant. The difference from a scalar field is that the components of a vector field at the point x transform into each other as well. If you need a concrete example, think about how the components of the electric field $\vec{E}(\vec{x})$ rotate into each other under 3D rotations, while a scalar potential $\phi(\vec{x})$ for which $\vec{E}(\vec{x}) = \vec{\nabla}\phi(\vec{x})$ is rotationally invariant.

A vector field $V_\mu(x)$ is a set of four functions of space-time. A Lorentz-invariant theory constructed with vector fields has a symmetry: the result of calculations will be the same if the four functions are mixed up according to Eq. (2.18). For example, $g^{\mu\nu}\partial_\mu V_\nu(x)$ is Lorentz invariant at each space-time point x if and only if $V_\mu(x)$ transforms as a vector field under Lorentz transformations. If $V_\mu(x)$ were just a collection of four scalar fields, $g^{\mu\nu}\partial_\mu V_\nu(x)$ would be frame-dependent.

Some important 4-vectors are position:

$$x^\mu = (t, x, y, z), \tag{2.19}$$

derivatives with respect to x^μ:

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = (\partial_t, \partial_x, \partial_y, \partial_z), \tag{2.20}$$

and momentum:

$$p^\mu = (E, p_x, p_y, p_z). \tag{2.21}$$

Tensors transform as

$$T^{\mu\nu} \to \Lambda^\mu_\alpha \Lambda^\nu_\beta T^{\alpha\beta}. \tag{2.22}$$

Tensor fields are functions of space-time, such as the energy-momentum tensor $T^{\mu\nu}(x)$ or the metric $g^{\mu\nu}(x)$ in general relativity. If you add more indices, such as $Z^{\mu\nu\alpha\beta}$, we still call it a tensor. The number of indices is the **rank** of a tensor, so $T^{\mu\nu}$ is rank 2, $Z^{\mu\nu\alpha\beta}$ is rank 4, etc.

When the same index appears twice, it is contracted, just as for rotations. Contractions implicitly involve the Minkowski metric and are Lorentz invariant. For example:

$$V^\mu W_\mu = V_\mu g^{\mu\nu} W_\nu = V_0 W_0 - V_1 W_1 - V_2 W_2 - V_3 W_3. \tag{2.23}$$

Such a contraction is Lorentz invariant and transforms like a scalar (just as the dot product of two 3-vectors $\vec{V} \cdot \vec{W}$, which is a contraction with δ_{ij}, is rotationally invariant). So, under a Lorentz transformation,

$$V^\mu W_\mu = VgW \to (V\Lambda^T) g (\Lambda W) = VgW = V^\mu W_\mu. \tag{2.24}$$

When writing contractions this way, you can usually just pretend g is the identity matrix. You will only need to distinguish g from δ when you write out components. This is one of the reasons the 4-vector notation is very powerful. Contracting indices is just a notational convention, not a deep property of mathematics.

It is worth adding a few more words about raising and lowering indices in field theory. In general relativity, it is important to be careful about distinguishing vectors with lower indices (**covariant** vectors) and vectors with upper indices (**contravariant** vectors). When an index appears twice (in a contraction) the technically correct approach is for one index to be upper and one to be lower. However, that can make the notation very cumbersome. For example, if the indices are ordered, you must write $V^\mu(x) \to \Lambda^\mu{}_\nu V^\nu(x)$, which is different from $V^\mu(x) \to \Lambda_\nu{}^\mu V^\nu(x)$. It is easier just to write $V^\mu \to \Lambda^{\mu\nu} V^\nu$ where the index order is clear. In special relativity, we always contract with the Minkowski metric $g_{\mu\nu} = \eta_{\mu\nu}$. So, we will often forget about which indices are upper and which are lower and just use the modern contraction convention for which all contractions are equivalent:

$$V_\mu W^\mu = V^\mu W_\mu = V_\mu W_\mu = V^\mu W^\mu. \tag{2.25}$$

Index position is important only when we plug in explicit vectors or matrices.

Although the index position is not important for us, the actual indices are. You should never have anything such as

$$V_\mu W_\mu X_\mu \tag{2.26}$$

with three (or more) of the same indices. To avoid this, be very careful about relabeling. For example, do not write

$$(V^2)(W^2) = V_\mu V_\mu W_\mu W_\mu; \tag{2.27}$$

instead write

$$(V^2)(W^2) = V_\mu^2 W_\nu^2 = V_\mu V_\mu W_\nu W_\nu = g_{\mu\alpha} g_{\nu\beta} V_\mu V_\alpha W_\nu W_\beta. \tag{2.28}$$

You will quickly get the hang of all this contracting.

The simplest Lorentz-invariant operator that we can write down involving derivatives is the **d'Alembertian**:

$$\Box = \partial_\mu^2 = \partial_t^2 - \partial_x^2 - \partial_y^2 - \partial_z^2. \tag{2.29}$$

This is the relativistic generalization of the **Laplacian**:

$$\triangle = \vec{\nabla}^2 = \partial_x^2 + \partial_y^2 + \partial_z^2. \tag{2.30}$$

Finally, it is worth keeping the terminology straight. We say that objects such as

$$V^2 = V_\mu V^\mu, \quad \phi, \quad 1, \quad \partial_\mu V^\mu \tag{2.31}$$

are **Lorentz invariant**, meaning they do not depend on our Lorentz frame at all, while objects such as

$$V_\mu, \quad F_{\mu\nu}, \quad \partial_\mu, \quad x_\mu \tag{2.32}$$

are **Lorentz covariant**, meaning they *do* change in different frames, but precisely as the Lorentz transformation dictates. Something such as energy density is neither Lorentz invariant nor Lorentz covariant; it is instead the 00 component of a Lorentz tensor $\mathcal{T}_{\mu\nu}$.

2.1.3 Discrete transformations

Lorentz transformations are defined to be those that preserve the Minkowski metric:

$$\Lambda^T g \Lambda = g. \tag{2.33}$$

Equivalently, they are those that leave inner products such as

$$V_\mu W^\mu = V_0 W_0 - V_1 W_1 - V_2 W_2 - V_3 W_3 \tag{2.34}$$

invariant. By this definition, the transformations

$$P : (t, x, y, z) \to (t, -x, -y, -z) \tag{2.35}$$

known as **parity** and

$$T : (t, x, y, z) \to (-t, x, y, z) \tag{2.36}$$

known as **time reversal** are also Lorentz transformations. They can be written as

$$P = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}. \tag{2.37}$$

Parity and time reversal are special because they cannot be written as the product of rotations and boosts, Eqs. (2.13) and (2.14). Discrete transformations play an important role in quantum field theory (see Chapter 11).

We say that a vector is **timelike** when

$$V^\mu V_\mu > 0 \quad \text{(timelike)} \tag{2.38}$$

and **spacelike** when

$$V^\mu V_\mu < 0 \quad \text{(spacelike)}. \tag{2.39}$$

Naturally, time $= (t, 0, 0, 0)$ is timelike and space $= (0, x, 0, 0)$ is spacelike. Whether something is timelike or spacelike is preserved under Lorentz transformations since the norm is preserved. If a vector has zero norm we say it is **lightlike**:

$$V^\mu V_\mu = 0 \quad \text{(lightlike)}. \tag{2.40}$$

If p^μ is a 4-momentum, then (since $p^2 = m^2$) it is lightlike if and only if it is massless. Photons are massless, which is the origin of the term *lightlike*.

Many more details of the mathematical structure of the Lorentz group (such as its unitary representations) will be covered in Chapters 8 and 10.

2.1.4 Solving problems with Lorentz invariance

Special relativity in quantum field theory is much easier than the special relativity you learned in your introductory physics course. We never need to talk about putting long cars in small garages or engineers with flashlights on trains. These situations are all designed

to make your non-relativistic intuition mislead you. In quantum field theory, other than the perhaps unintuitive notion that energy can turn into matter through $E = mc^2$, your non-relativistic intuition will serve you perfectly well.

For field theory, all you really need from special relativity is the one equation that defines Lorentz transformations:

$$\Lambda^T g \Lambda = g. \tag{2.41}$$

This implies that contractions such as $p^2 \equiv p^\mu p_\mu$ are Lorentz invariant. For problems that involve changing frames, usually you know everything in one frame and are interested in some quantity in another frame. For example, you may know momenta p_1^μ and p_2^μ of two incoming particles that collide and are interested in the energy of an outgoing particle E_3 in the center-of-mass frame (the center-of-mass frame is defined as the frame in which the total 3-momenta, $\vec{p}_{\text{tot}} = 0$). For such problems, it is best to first calculate a Lorentz-invariant quantity such as $p_{\text{tot}}^2 = (p_1^\mu + p_2^\mu)^2$ in the first frame, then go to the second frame, and solve for the unknown quantity. Since p_{tot}^2 is Lorentz invariant, it has the same value in both frames. Usually, when you input everything you know about the second frame (e.g. $\vec{p}_{\text{tot}} = 0$ if it is the center-of-mass frame), you can solve for the remaining unknowns. If you find yourself plugging in explicit boost and rotation matrices, you are probably solving the problem the hard way. This trick is especially useful for situations in which there are many particles, say $p_1^\mu, \ldots, p_5^\mu$, and therefore many Lorentz-invariant quantities, such as $p_1^\mu p_{4\mu}$ or $(p_5^\mu + p_4^\mu)^2$.

2.2 Classical plane waves as oscillators

We next review the simple harmonic oscillator and discuss the connection to special relativity.

2.2.1 Simple harmonic oscillator

Anything with a linear restoring potential (any potential is linear close enough to equilibrium), such as a spring, or a string with tension, or a wave, is a harmonic oscillator. For example, a spring has

$$m\frac{d^2x}{dt^2} + kx = 0, \tag{2.42}$$

which is satisfied by $x(t) = \cos\left(\sqrt{\frac{k}{m}}t\right)$, so it oscillates with frequency

$$\omega = \sqrt{\frac{k}{m}}. \tag{2.43}$$

A more general solution is

$$x(t) = c_1 e^{i\omega t} + c_2 e^{-i\omega t}. \tag{2.44}$$

The classical Hamiltonian for this system is the sum of kinetic and potential energies:

$$H = \frac{1}{2}\frac{p^2}{m} + \frac{1}{2}m\omega^2 x^2. \tag{2.45}$$

To quantize the harmonic oscillator, we promote x and p to operators and impose the canonical commutation relations

$$[x, p] = i. \tag{2.46}$$

Analysis of the harmonic oscillator spectrum is simplest if we change variables to

$$a = \sqrt{\frac{m\omega}{2}}\left(x + \frac{ip}{m\omega}\right), \quad a^\dagger = \sqrt{\frac{m\omega}{2}}\left(x - \frac{ip}{m\omega}\right), \tag{2.47}$$

which satisfy

$$[a, a^\dagger] = 1, \tag{2.48}$$

so that

$$H = \omega\left(a^\dagger a + \frac{1}{2}\right). \tag{2.49}$$

Thus, energy eigenstates are eigenstates of the number operator

$$\hat{N} = a^\dagger a, \tag{2.50}$$

which is Hermitian. The results we derived in Section 1.3:

$$\hat{N}|n\rangle = n|n\rangle, \tag{2.51}$$

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \tag{2.52}$$

$$a|n\rangle = \sqrt{n}|n-1\rangle, \tag{2.53}$$

follow from these definitions. We can also calculate how the operators evolve in time (in the Heisenberg picture):

$$i\frac{d}{dt}a = [a, H] = \left[a, \omega\left(a^\dagger a + \frac{1}{2}\right)\right] = \omega(aa^\dagger a - a^\dagger aa) = \omega[a, a^\dagger]a = \omega a. \tag{2.54}$$

This equation is solved by

$$a(t) = e^{-i\omega t}a(0). \tag{2.55}$$

2.2.2 Connection to special relativity

To connect special relativity to the simple harmonic oscillator we note that the simplest possible Lorentz-invariant equation of motion that a field can satisfy is $\Box\phi = 0$. That is,

$$\Box\phi = (\partial_t^2 - \vec{\nabla}^2)\phi = 0. \tag{2.56}$$

The classical solutions are plane waves. For example, one solution is

$$\phi(x) = a_p(t)e^{i\vec{p}\cdot\vec{x}}, \tag{2.57}$$

where

$$(\partial_t^2 + \vec{p}\cdot\vec{p})a_p(t) = 0. \tag{2.58}$$

This is exactly the equation of motion of a harmonic oscillator. A general solution is

$$\phi(x,t) = \int \frac{d^3p}{(2\pi)^3}\left[a_p(t)\, e^{i\vec{p}\cdot\vec{x}} + a_p^\star(t)\, e^{-i\vec{p}\cdot\vec{x}}\right], \tag{2.59}$$

with $(\partial_t^2 + \vec{p}\cdot\vec{p})a_p(t) = 0$, which is just a Fourier decomposition of the field into plane waves. Or more simply

$$\phi(x,t) = \int \frac{d^3p}{(2\pi)^3}\left(a_p e^{-ipx} + a_p^\star e^{ipx}\right), \tag{2.60}$$

with a_p and $a_p^\star$ now just numbers and $p_\mu \equiv (\omega_p, \vec{p})$ with $\omega_p \equiv |\vec{p}|$. To be extra clear about notation, px contains an implicit 4-vector contraction: $px = p^\mu x_\mu = \omega_p x_0 - \vec{p}\cdot\vec{x}$.

Not only is $\Box\phi = 0$ the simplest Lorentz-invariant field equation possible, it is one of the equations that free massless fields will always satisfy (up to some exotic exceptions). For example, recall that there is a nice Lorentz-covariant treatment of electromagnetism using

$$F_{\mu\nu} \equiv \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{2.61}$$

This $F_{\mu\nu}$ transforms covariantly as a tensor under Lorentz transformations and thus concisely encodes how $\vec{E}$ and $\vec{B}$ rotate into each other under boosts. In terms of $F_{\mu\nu}$, Maxwell's equations in empty space have the simple forms

$$\partial_\mu F_{\mu\nu} = 0, \qquad \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0. \tag{2.62}$$

Any field satisfying these equations can be written as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.63}$$

Although not necessary, we can also require $\partial_\mu A_\mu = 0$, which is a gauge choice (**Lorenz gauge**). We will discuss gauge invariance in great detail in Chapters 8 and 25. For now, it is enough to know that the physical $\vec{E}$ and $\vec{B}$ fields *can* be combined into an antisymmetric tensor $F_{\mu\nu}$, which is determined by a 4-vector A_μ satisfying $\partial_\mu A_\mu = 0$. In Lorenz gauge, Maxwell's equations reduce to

$$\partial_\mu F_{\mu\nu} = \Box A_\nu - \partial_\nu(\partial_\mu A_\mu) = \Box A_\nu = 0. \tag{2.64}$$

Thus, each component of A_ν satisfies the minimal Lorentz-invariant equation of motion.

That $\Box\phi = 0$ for a scalar field and $\Box A_\mu = 0$ for a vector field have the same form is not a coincidence. The electromagnetic field is made up of particles of spin 1 called photons. The polarizations of the field are encoded in the four fields $A_\nu(x)$. In fact, massless particles of *any* spin will satisfy $\Box\chi_i = 0$ where the different fields, indexed by i, encode different polarizations of that particle. This is not obvious, and we are not ready to prove it, so let us focus simply on the electromagnetic field. For simplicity, we will ignore polarizations for now and just treat A_ν as a scalar field ϕ (such approximations were used in some of the earliest QED papers, e.g. [Born *et al.*, 1926]). A general solution to Maxwell's equations in Lorenz gauge is therefore given by Eq. (2.60) for each polarization (polarizations will

be explained in Chapter 8). Such a solution simply represents the Fourier decomposition of electromagnetic fields into plane waves. The oscillation of the waves is the same as the oscillation of a harmonic oscillator for each value of $\vec{p}$.

2.3 Second quantization

Since the modes of an electromagnetic field have the same classical equations as a simple harmonic oscillator, we can quantize them in the same way. We introduce an annihilation operator a_p and its conjugate creation operator $a_p^\dagger$ for each wavenumber $\vec{p}$ and integrate over them to get the Hamiltonian for the free theory:

$$H_0 = \int \frac{d^3p}{(2\pi)^3}\omega_p\left(a_p^\dagger a_p + \frac{1}{2}\right), \tag{2.65}$$

with

$$\omega_p = |\vec{p}|\,. \tag{2.66}$$

This is known as **second quantization**. At the risk of oversimplifying things a little, that is all there is to quantum field theory. The rest is just quantum mechanics.

First quantization refers to the discrete modes, for example, of a particle in a box. Second quantization refers to the integer numbers of excitations of each of these modes. However, this is somewhat misleading – the fact that there are discrete modes is a classical phenomenon. The two steps really are (1) interpret these modes as having energy $E = \hbar\omega$ and (2) quantize each mode as a harmonic oscillator. In that sense we are only quantizing once. Whether second quantization is a good name for this procedure is semantics, not physics.

There are two new features in second quantization:

1. We have many quantum mechanical systems – one for each $\vec{p}$ – all at the same time.
2. We interpret the nth excitation of the $\vec{p}$ harmonic oscillator as having n **particles**.

Let us take a moment to appreciate this second point. Recall the old simple harmonic oscillator: the electron in a quadratic potential. We would never interpret the states $|n\rangle$ of this system as having n electrons. The fact that a pointlike electron in a quadratic potential has analogous equations of motion to a Fourier component of the electromagnetic field is just a coincidence. Do not let it confuse you. Both are just the simplest possible dynamical systems, with linear restoring forces.[1]

In second quantization, the Hilbert space is promoted to a **Fock space**, which is defined at each time as a direct sum,

$$\mathcal{F} = \oplus_n \mathcal{H}_n, \tag{2.67}$$

[1] To set up a proper analogy we need to first treat the electron as a classical field (we do not know how to do that yet), and find a set of solutions (such as the discrete frequencies of the electromagnetic waves). Then we would quantize each of those solutions, allowing $|n\rangle$ excitations. However, if we did this, electrons would have Bose–Einstein statistics. Instead, they must have Fermi–Dirac statistics, so we would have to restrict n to 0 or 1. The second quantization of electrons will be discussed in Chapters 10 through 12, and the interpretation of an electron as a classical field, which requires Grassmann numbers, in Chapter 14.

of Hilbert spaces, $\mathcal{H}_n$, of physical n-particle states. If there is one particle type, states in $\mathcal{H}_n$ are linear combinations of states $\{|p_1^\mu, \ldots, p_n^\mu\rangle\}$ of all possible momenta satisfying $p_i^2 = m^2$ with $p_i^0 > 0$. If there are many different particle types, the Fock space is the direct sum of the Hilbert spaces associated with each particle. The Fock space is the same at all times, by time-translation invariance, and in any frame, by Lorentz invariance. Note that the Fock space is *not* a sum over Hilbert spaces defined with arbitrary 4-vectors, since the energy for a physical state is determined by its 3-momentum $\vec{p}_i$ and its mass m_i as $p_i^0 = \sqrt{\vec{p}_i^{\,2} + m_i^2}$. We thus write $|\vec{p}\rangle$, $|p^\mu\rangle$ and $|p\rangle$ interchangeably.

2.3.1 Field expansion

Now let us get a little more precise about what the Hamiltonian in Eq. (2.65) means. The natural generalizations of

$$[a, a^\dagger] = 1 \tag{2.68}$$

are the equal-time commutation relations

$$[a_k, a_p^\dagger] = (2\pi)^3 \delta^3(\vec{p} - \vec{k}). \tag{2.69}$$

The factors of 2π are a convention, stemming from our convention for Fourier transforms (see Appendix A). These $a_p^\dagger$ operators create particles with momentum p:

$$a_p^\dagger |0\rangle = \frac{1}{\sqrt{2\omega_p}} |\vec{p}\rangle, \tag{2.70}$$

where $|\vec{p}\rangle$ is a state with a single particle of momentum $\vec{p}$. This factor of $\sqrt{2\omega_p}$ is just another convention, but it will make some calculations easier. Its nice Lorentz transformation properties are studied in Problem 2.6.

To compute the normalization of one-particle states, we start with

$$\langle 0|0\rangle = 1, \tag{2.71}$$

which leads to

$$\langle \vec{p}\,|\vec{k}\rangle = 2\sqrt{\omega_p \omega_k} \langle 0|a_p a_k^\dagger|0\rangle = 2\omega_p (2\pi)^3 \delta^3(\vec{p} - \vec{k}). \tag{2.72}$$

The identity operator for one-particle states is

$$\mathbb{1} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} |\vec{p}\rangle \langle \vec{p}\,|, \tag{2.73}$$

which we can check with

$$|\vec{k}\rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} |\vec{p}\rangle \langle \vec{p}\,|\vec{k}\rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} 2\omega_p (2\pi)^3 \delta^3(\vec{p} - \vec{k}) |\vec{p}\rangle = |\vec{k}\rangle. \tag{2.74}$$

We then define quantum fields as integrals over creation and annihilation operators for each momentum:

$$\phi_0(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \left(a_p e^{i\vec{p}\vec{x}} + a_p^\dagger e^{-i\vec{p}\vec{x}}\right), \tag{2.75}$$

where the subscript 0 indicates this is a free field. The factor of $\sqrt{2\omega_p}$ is included for later convenience.

This equation looks just like the classical free-particle solutions, Eq. (2.59), to Maxwell's equations (ignoring polarizations) but instead of a_p and $a_p^\dagger$ being *functions*, they are now the annihilation and creation *operators* for that mode. Sometimes we say the classical a_p is c-number valued and the quantum one is q-number valued. The connection with Eq. (2.59) is only suggestive. The quantum equation, Eq. (2.75), should be taken as the definition of a field operator $\phi_0(\vec{x})$ constructed from the creation and annihilation operators a_p and $a_p^\dagger$.

To get a sense of what the operator ϕ_0 does, we can act with it on the vacuum and project out a momentum component:

$$\begin{aligned}\langle\vec{p}|\phi_0(\vec{x})|0\rangle &= \langle 0|\sqrt{2\omega_p}a_p \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\left(a_k e^{i\vec{k}\vec{x}} + a_k^\dagger e^{-i\vec{k}\vec{x}}\right)|0\rangle \\ &= \int \frac{d^3k}{(2\pi)^3}\sqrt{\frac{\omega_p}{\omega_k}}\left[e^{i\vec{k}\vec{x}}\langle 0|a_p a_k|0\rangle + e^{-i\vec{k}\vec{x}}\langle 0|a_p a_k^\dagger|0\rangle\right] \\ &= e^{-i\vec{p}\vec{x}}. \end{aligned} \tag{2.76}$$

This is the same thing as the projection of a position state on a momentum state in one-particle quantum mechanics:

$$\langle\vec{p}|\vec{x}\rangle = e^{-i\vec{p}\vec{x}}. \tag{2.77}$$

So, $\phi_0(\vec{x})|0\rangle = |\vec{x}\rangle$, that is, $\phi_0(\vec{x})$ creates a particle at position $\vec{x}$. This should not be surprising, since $\phi_0(x)$ in Eq. (2.75) is very similar to $x = a + a^\dagger$ in the simple harmonic oscillator. Since ϕ_0 is Hermitian, $\langle\vec{x}| = \langle 0|\phi_0(\vec{x})$ as well.

By the way, there are many states $|\psi\rangle$ in the Fock space that satisfy $\langle\vec{p}|\psi\rangle = e^{-i\vec{p}\vec{x}}$. Since $\langle\vec{p}|$ only has non-zero matrix elements with one-particle states, adding to $|x\rangle$ a two- or zero-particle state, as in $\phi_0^2(\vec{x})|0\rangle$, has no effect on $\langle\vec{p}|\vec{x}\rangle$. That is, $|\psi\rangle = \left(\phi_0(\vec{x}) + \phi_0^2(\vec{x})\right)|0\rangle$ also satisfies $\langle\vec{p}|\psi\rangle = e^{-i\vec{p}\vec{x}}$. The state $|\vec{x}\rangle \equiv \phi_0(\vec{x})|0\rangle$ is the unique *one-particle* state with $\langle\vec{p}|\psi\rangle = e^{-i\vec{p}\vec{x}}$.

2.3.2 Time dependence

In quantum field theory, we generally work in the Heisenberg picture, where all the time dependence is in operators such as ϕ and a_p. For free fields, the creation and annihilation operators for each momentum $\vec{p}$ in the quantum field are just those of a simple harmonic oscillator. These operators should satisfy Eq. (2.55), $a_p(t) = e^{-i\omega_p t}a_p$, and its conjugate $a_p^\dagger(t) = e^{i\omega_p t}a_p^\dagger$, where a_p and $a_p^\dagger$ (without an argument) are time independent. Then, we can *define* a quantum scalar field as

$$\phi_0(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left(a_p e^{-ipx} + a_p^\dagger e^{ipx}\right), \tag{2.78}$$

with $p^\mu \equiv (\omega_p, \vec{p}\,)$ and $\omega_p = |\vec{p}\,|$ as in Eq. (2.60). The 0 subscript still indicates that these are free fields.

To be clear, there is no physical content in Eq. (2.78). It is just a definition. The physical content is in the algebra of a_p and $a_p^\dagger$ and in the Hamiltonian H_0. Nevertheless, we will see that collections of a_p and $a_p^\dagger$ in the form of Eq. (2.78) are very useful in quantum field theory. For example, you may note that while the integral is over only three components of p_μ, the phases have combined into a manifestly Lorentz-invariant form. This field now automatically satisfies $\Box\phi(x) = 0$. If a scalar field had mass m, we could still write it in exactly the same way but with a massive dispersion relation: $\omega_p \equiv \sqrt{\vec{p}^2 + m^2}$. Then the quantum field still satisfies the classical equation of motion: $(\Box + m^2)\phi(x) = 0$.

Let us check that our free Hamiltonian is consistent with the expectation for time evolution. Commuting the free fields with H_0 we find

$$\begin{aligned}
[H_0, \phi_0(\vec{x}, t)] &= \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_k}} \left[\omega_p \left(a_p^\dagger a_p + \frac{1}{2} \right), a_k e^{-ikx} + a_k^\dagger e^{ikx} \right] \\
&= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \left[-\omega_p a_p \, e^{-ipx} + \omega_p a_p^\dagger e^{ipx} \right] \\
&= -i\partial_t \phi_0(\vec{x}, t),
\end{aligned} \tag{2.79}$$

which is exactly the expected result.

For any Hamiltonian, quantum fields satisfy the Heisenberg equations of motion:

$$i\partial_t \phi(x) = [\phi, H]\,. \tag{2.80}$$

In a free theory, $H = H_0$, and this is consistent with Eq. (2.78). In an interacting theory, that is, one whose Hamiltonian H differs from the free Hamiltonian H_0, the Heisenberg equations of motion are still satisfied, but we will rarely be able to solve them exactly. To study interacting theories, it is often useful to use the same notation for interacting fields as for free fields:

$$\phi(\vec{x}, t) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \left[a_p(t) e^{-ipx} + a_p^\dagger(t) e^{ipx} \right]. \tag{2.81}$$

At any *fixed time*, the full interacting creation and annihilation operators $a_p^\dagger(t)$ and $a_p(t)$ satisfy the same algebra as in the free theory – the Fock space is the same at every time, due to time-translation invariance. We can therefore define the exact creation operators $a_p(t)$ to be equal to the free creation operators a_p at any given fixed time, $a_p(t_0) = a_p$ and so $\phi(\vec{x}, t_0) = \phi_0(\vec{x}, t_0)$. However, the operators that create particular momentum states $|p\rangle$ in the interacting theory mix with each other as time evolves. We generally will not be able to solve the dynamics of an interacting theory exactly. Instead, we will expand $H = H_0 + H_{\rm int}$ and calculate amplitudes using time-dependent perturbation theory with $H_{\rm int}$, just as in quantum mechanics. In Chapter 7, we use this approach to derive the Feynman rules.

The first-quantized (quantum mechanics) limit of the second-quantized theory (quantum field theory) comes from restricting to the one-particle states, which is appropriate in the non-relativistic limit. A basis of these states is given by the vectors $\langle x| = \langle \vec{x}, t|$:

$$\langle x| = \langle 0|\, \phi(\vec{x}, t). \tag{2.82}$$

Then, a Schrödinger picture wavefunction is

$$\psi(x) = \langle x|\psi\rangle\,, \tag{2.83}$$

which satisfies

$$i\partial_t\psi(x) = i\partial_t\langle 0|\phi(\vec{x},t)|\psi\rangle = i\,\langle 0|\partial_t\phi(\vec{x},t)|\psi\rangle\,. \tag{2.84}$$

In the massive case, the free quantum field $\phi_0(x)$ satisfies $\partial_t^2\phi_0 = \left(\vec{\nabla}^2 - m^2\right)\phi_0$ and we have from Eq. (2.79) (with the massive dispersion relation $\omega_p = \sqrt{\vec{p}^2 + m^2}$):

$$\begin{aligned} i\langle 0|\partial_t\phi(\vec{x},t)|\psi\rangle &= \langle 0|\int \frac{d^3p}{(2\pi)^3}\frac{\sqrt{\vec{p}^2+m^2}}{\sqrt{2\omega_p}}\left(a_p e^{-ipx} - a_p^\dagger e^{ipx}\right)|\psi\rangle \\ &= \langle 0|\sqrt{m^2 - \vec{\nabla}^2}\phi_0(x)|\psi\rangle\,. \end{aligned} \tag{2.85}$$

So,

$$i\partial_t\psi(x) = \sqrt{m^2 - \vec{\nabla}^2}\psi(x) = \left(m - \frac{\vec{\nabla}^2}{2m} + \mathcal{O}\left(\frac{1}{m^2}\right)\right)\psi(x). \tag{2.86}$$

The final form is the low-energy (large-mass) expansion. We can then define the non-relativistic Hamiltonian by subtracting off the mc^2 contribution to the energy, which is irrelevant in the non-relativistic limit. This gives

$$i\partial_t\psi(x) = -\frac{\vec{\nabla}^2}{2m}\psi(x), \tag{2.87}$$

which is the non-relativistic Schrödinger equation for a free theory. Another way to derive the quantum mechanics limit of quantum field theory is discussed in Section 33.6.2.

2.3.3 Commutation relations

We will occasionally need to use the equal-time commutation relations of the second-quantized field and its time derivative. The commutator of a field at two different points is

$$\begin{aligned} [\phi(\vec{x}),\phi(\vec{y})] &= \int\frac{d^3p}{(2\pi)^3}\int\frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p 2\omega_q}}\left[\left(a_p e^{i\vec{p}\vec{x}} + a_p^\dagger e^{-i\vec{p}\vec{x}}\right),\left(a_q e^{i\vec{q}\vec{y}} + a_q^\dagger e^{-i\vec{q}\vec{y}}\right)\right] \\ &= \int\frac{d^3p}{(2\pi)^3}\int\frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p 2\omega_q}}\left(e^{i\vec{p}\vec{x}}e^{-i\vec{q}\vec{y}}\left[a_p, a_q^\dagger\right] + e^{-i\vec{p}\vec{x}}e^{i\vec{q}\vec{y}}\left[a_p^\dagger, a_q\right]\right). \end{aligned} \tag{2.88}$$

Using Eq. (2.69), $[a_k, a_p^\dagger] = (2\pi)^3\delta^3(\vec{p}-\vec{k})$, this becomes

$$\begin{aligned} [\phi(\vec{x}),\phi(\vec{y})] &= \int\frac{d^3p}{(2\pi)^3}\int\frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p 2\omega_q}}\left[\,e^{i\vec{p}\vec{x}}e^{-i\vec{q}\vec{y}} - e^{-i\vec{p}\vec{x}}e^{i\vec{q}\vec{y}}\,\right](2\pi)^3\,\delta^3(\vec{p}-\vec{q}) \\ &= \int\frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_p}\left[e^{i\vec{p}(\vec{x}-\vec{y})} - e^{-i\vec{p}(\vec{x}-\vec{y})}\right]. \end{aligned} \tag{2.89}$$

Since the integral measure and $\omega_p = \sqrt{\vec{p}^2+m^2}$ are symmetric under $\vec{p}\to-\vec{p}$ we can flip the sign on the exponent of one of the terms to see that the commutator vanishes:

$$[\phi(\vec{x}), \phi(\vec{y})] = 0. \tag{2.90}$$

The equivalent calculation at different times is much more subtle (we discuss the general result in Section 12.6 in the context of the spin-statistics theorem).

Next, we note that the time derivative of the free field, at $t = 0$, has the form

$$\pi(\vec{x}) \equiv \partial_t \phi(x)\Big|_{t=0} = -i \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_p}{2}} \left(a_p e^{i\vec{p}\vec{x}} - a_p^\dagger e^{-i\vec{p}\vec{x}}\right), \tag{2.91}$$

where π is the operator canonically conjugate to ϕ. As $\phi(\vec{x})$ is the second-quantized analog of the $\hat{x}$ operator, $\pi(\vec{x})$ is the analog of the $\hat{p}$ operator. Note that $\pi(\vec{x})$ has nothing to do with the physical momentum of states in the Hilbert space: $\pi(\vec{x})\,|0\rangle$ is not a state of given momentum. Instead, it is a state also at position $\vec{x}$ created by the time derivative of $\phi(\vec{x})$.

Now we compute

$$\begin{aligned}
[\phi(\vec{x}), \pi(\vec{y})] &= -i \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3q}{(2\pi)^3} \sqrt{\frac{\omega_p}{2}} \frac{1}{\sqrt{2\omega_q}} \left(e^{i\vec{p}\vec{y}} e^{-i\vec{q}\vec{x}} \left[a_q^\dagger, a_p\right] - e^{i\vec{q}\vec{x}} e^{-i\vec{p}\vec{y}} \left[a_q, a_p^\dagger\right]\right) \\
&= \frac{i}{2} \int \frac{d^3p}{(2\pi)^3} \left[e^{i\vec{p}(\vec{x}-\vec{y})} + e^{-i\vec{p}(\vec{x}-\vec{y})}\right].
\end{aligned} \tag{2.92}$$

Both of these integrals give $\delta^3(\vec{x} - \vec{y})$, so we find

$$[\phi(\vec{x}), \pi(\vec{y})] = i\delta^3(\vec{x} - \vec{y}), \tag{2.93}$$

which is the analog of $[\hat{x}, \hat{p}] = i$ in quantum mechanics. It encapsulates the field theory version of the uncertainty principle: you cannot know the properties of the field and its rate of change at the same place at the same time.

In a general interacting theory, at any fixed time, $\phi(\vec{x})$ and $\pi(\vec{x})$ have expressions in terms of creation and annihilation operators whose algebra is identical to that of the free theory. Therefore, they satisfy the commutation relations in Eqs. (2.90) and (2.93) as well as $[\pi(\vec{x}), \pi(\vec{y})] = 0$. The Hamiltonian in an interacting theory should be expressed as a functional of the operators $\phi(\vec{x})$ and $\pi(\vec{x})$ with time evolution given by $\partial_t \mathcal{O} = i[H, \mathcal{O}]$. Any such Hamiltonian can then be expressed entirely in terms of creation and annihilation operators using Eqs. (2.75) and (2.91); thus it has a well-defined action on the associated Fock space. Conversely, it is sometimes more convenient (especially for non-relativistic or condensed matter applications) to derive the form of the Hamiltonian in terms of a_p and $a_p^\dagger$. We can then express a_p and $a_p^\dagger$ in terms of $\phi(\vec{x})$ and $\pi(\vec{x})$ by inverting Eqs. (2.75) and (2.91) for a_p and $a_p^\dagger$ (the solution is the field theory equivalent of Eq. (2.47)).

In summary, all we have done to quantize the electromagnetic field is to treat it as an infinite set of simple harmonic oscillators, one for each wavenumber $\vec{p}$. More generally:

Quantum field theory is just quantum mechanics with an infinite number of harmonic oscillators.

2.3.4 Einstein coefficients revisited

In quantum mechanics we usually study a single electron in a background potential $V(x)$. In quantum field theory, the background (e.g. the electromagnetic system) is dynamical, so all kinds of new phenomena can be explained. We already saw one example in Chapter 1. We can now be a little more explicit about what the relevant Hamiltonian should be for Dirac's calculation of the Einstein coefficients.

We can always write a Hamiltonian as

$$H = H_0 + H_{\text{int}}, \tag{2.94}$$

where H_0 describes some system that we can solve exactly. In the case of the two-state system discussed in Chapter 1, we can take H_0 to be the sum of the Hamiltonians for the atom and the photons:

$$H_0 = H_{\text{atom}} + H_{\text{photon}}. \tag{2.95}$$

The eigenstates of H_{atom} are the energy eigenstates $|\psi_n\rangle$ of the hydrogen atom, with energies E_n. H_{photon} is the Hamiltonian in Eq. (2.65) above:

$$H_{\text{photon}} = \int \frac{d^3k}{(2\pi)^3}\omega_k \left(a_k^\dagger a_k + \frac{1}{2}\right). \tag{2.96}$$

The remaining H_{int} is hopefully small enough to let us use perturbation theory.

Fermi's golden rule from quantum mechanics says the rate for transitions between two states is proportional to the square of the matrix element of the interaction between the two states:

$$\Gamma \propto |\langle f|H_{\text{int}}|i\rangle|^2 \delta(E_f - E_i), \tag{2.97}$$

and we can treat the interaction semi-classically:

$$H_{\text{int}} = \phi H_I. \tag{2.98}$$

As mentioned in Footnote 2 in Chapter 1, H_I can be derived from the $\frac{e}{m}\vec{p}\cdot\vec{A}$ interaction of the minimally coupled non-relativistic Hamiltonian, $H = \frac{1}{2m}(\vec{p} + e\vec{A})^2$. Since we are ignoring spin, it does not pay to be too precise about H_I; the important point being only that H_{int} has a quantum field ϕ in it, representing the photon, and H_I has non-zero matrix elements between different atomic states.

According to Fermi's golden rule, the transition probability is proportional to the matrix element of the interaction squared. Then,

$$\mathcal{M}_{1\to 2} = \langle \text{atom}^\star; n_k - 1|H_{\text{int}}|\text{atom}; n_k\rangle \propto \langle \text{atom}^\star|H_I|\text{atom}\rangle \sqrt{n_k}, \tag{2.99}$$

$$\mathcal{M}_{2\to 1} = \langle \text{atom}; n_k + 1|H_{\text{int}}|\text{atom}^\star; n_k\rangle \propto \langle \text{atom}|H_I|\text{atom}^\star\rangle \sqrt{n_k + 1}, \tag{2.100}$$

where we have used

$$\langle n_k - 1|\phi|n_k\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\langle n_k - 1|a_p|n_k\rangle \propto \sqrt{n_k}, \tag{2.101}$$

$$\langle n_k+1|\phi|n_k\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\langle n_k+1|a_p^\dagger|n_k\rangle \propto \sqrt{n_k+1}. \tag{2.102}$$

Thus, $\mathcal{M}_{1\to 2}$ and $\mathcal{M}_{2\to 1}$ agree with what we used in Chapter 1 to reproduce Dirac's calculation of the Einstein coefficients. Note that we only used one photon mode, of momentum k, so this was really just quantum mechanics. Quantum field theory just gave us a δ-function from the d^3p integration.

Problems

2.1 Derive the transformations $x \to \frac{x+vt}{\sqrt{1-v^2}}$ and $t \to \frac{t+vx}{\sqrt{1-v^2}}$ in perturbation theory. Start with the Galilean transformation $x \to x + vt$. Add a transformation $t \to t + \delta t$ and solve for δt assuming it is linear in x and t and preserves $t^2 - x^2$ to $\mathcal{O}\left(v^2\right)$. Repeat for δt and δx to second order in v and show that the result agrees with the second-order expansion of the full transformations.

2.2 Special relativity and colliders.

(a) The Large Hadron Collider was designed to collide protons together at 14 TeV center-of-mass energy. How many kilometers per hour less than the speed of light are the protons moving?

(b) How fast is one proton moving with respect to the other?

2.3 The GZK bound. In 1966 Greisen, Zatsepin and Kuzmin argued that we should not see cosmic rays (high-energy protons hitting the atmosphere from outer space) above a certain energy, due to interactions of these rays with the cosmic microwave background.

(a) The universe is a blackbody at 2.73 K. What is the average energy of the photons in outer space (in electronvolts)?

(b) How much energy would a proton (p^+) need to collide with a photon (γ) in outer space to convert it to a 135 MeV pion (π^0)? That is, what is the energy threshold for $p^+ + \gamma \to p^+ + \pi^0$?

(c) How much energy does the outgoing proton have after this reaction?

This GZK bound was finally confirmed experimentally 40 years after it was conjectured [Abbasi *et al.*, 2008].

2.4 Is the transformation $Y : (t, x, y, z) \to (t, x, -y, z)$ a Lorentz transformation? If so, why is it not considered with P and T as a discrete Lorentz transformation? If not, why not?

2.5 Compton scattering. Suppose we scatter an X-ray off an electron in a crystal, but we cannot measure the electron's momentum, just the reflected X-ray momentum.

(a) Why is it OK to treat the electrons as free?

(b) Calculate the frequency dependence of the reflected X-ray on the scattering angle. Draw a rough plot.

(c) What happens to the distribution as you take the electron mass to zero?

(d) If you did not believe in quantized photon momenta, what kind of distribution might you have expected? [Hint: see [Compton, 1923].]

2.6 Lorentz invariance.

(a) Show that

$$\int_{-\infty}^{\infty} dk^0 \delta(k^2 - m^2)\theta(k^0) = \frac{1}{2\omega_k}, \tag{2.103}$$

where $\theta(x)$ is the unit step function and $\omega_k \equiv \sqrt{\vec{k}^2 + m^2}$.

(b) Show that the integration measure d^4k is Lorentz invariant.

(c) Finally, show that

$$\int \frac{d^3k}{2\omega_k} \tag{2.104}$$

is Lorentz invariant.

2.7 Coherent states of the simple harmonic oscillator.

(a) Calculate $\partial_z(e^{-za^\dagger} a e^{za^\dagger})$ where z is a complex number.

(b) Show that $|z\rangle = e^{za^\dagger}|0\rangle$ is an eigenstate of a. What is its eigenvalue?

(c) Calculate $\langle n|z\rangle$.

(d) Show that these "coherent states" are minimally dispersive: $\Delta p \Delta q = \frac{1}{2}$, where $\Delta q^2 = \langle q^2\rangle - \langle q\rangle^2$ and $\Delta p^2 = \langle p^2\rangle - \langle p\rangle^2$, where $\langle q\rangle = \frac{\langle z|q|z\rangle}{\langle z|z\rangle}$ and $\langle p\rangle = \frac{\langle z|p|z\rangle}{\langle z|z\rangle}$.

(e) Why can you not make an eigenstate of $a^\dagger$?