Contents

1 Quantization of the Electromagnetic Field **1**
- 1.1 Harmonic Oscillator Revision . . . 1
- 1.2 Heisenberg and Schrödinger pictures Revision . . . 2
- 1.3 The classical electromagnetic field . . . 4
- 1.4 Quantizing the electromagnetic field . . . 7

2 Quantum States of the Electromagnetic Field **10**
- 2.1 Single-Mode Pure States . . . 10
- 2.2 Single Mode Mixed States . . . 13
- 2.3 Multimode Field States . . . 17

3 Quantum Electrodynamics **23**
- 3.1 The minimal coupling Hamiltonian . . . 23
- 3.2 * The Lorentz Force . . . 23
- 3.3 * Radiation Sources . . . 25
- 3.4 Interpretation of the Minimal Coupling Hamiltonian . . . 28

4 Examples of QED processes **30**
- 4.1 Radiative Transitions in Atoms . . . 30
- 4.2 Thomson scattering . . . 37

5 Relativistic Quantum Mechanics **40**
- 5.1 Spin . . . 40
- 5.2 Relativity and Quantum Mechanics . . . 45
- 5.3 The Klein-Gordon Equation . . . 48
- 5.4 The Dirac Equation: Derivation . . . 50
- 5.5 The Dirac Equation: Properties . . . 52
- 5.6 The Dirac Equation: Solutions . . . 54

6 Relativistic QED and the Standard Model **58**
- 6.1 Dirac's Relativistic QED . . . 58
- 6.2 * Fundamental Particles . . . 62
- 6.3 * Other Quanta . . . 64
- 6.4 * Renormalization revisited . . . 65
- 6.5 The spin-statistics theorem . . . 67

1 Quantization of the Electromagnetic Field

1.1 Harmonic Oscillator Revision

Although the photon is commonly referred to as an elementary particle of light, it is quite different from elementary particles such as electrons and quarks. There is no such thing as a photon wavefunction, although there are some text books which describe the electric field in those terms. This is quite wrong. The electric field is a real field which exists in real space, not a probability amplitude which exists in Hilbert space. The relation between the electric field and the photon should become clear later, but for now the following description will have to do: A photon is an elementary excitation of a mode of the quantized electromagnetic field. In that sense it is analogous to an excitation of a quantized particle in a harmonic potential. That is, the discreteness of the electromagnetic field energy (photons) is present for the same reason as the discreteness of the energy levels of a particle in a potential well. For this reason it is useful first to review the quantized harmonic oscillator.

The energy of a classical harmonic oscillator is of course

$$E = \tfrac{1}{2}m\omega^2 q^2 + \tfrac{1}{2}m\dot{q}^2 \tag{1.1.1}$$

$$= \frac{\hbar\omega}{2}\left(X^2 + \omega^{-2}\dot{X}^2\right) \tag{1.1.2}$$

the second line, despite the appearance of $\hbar$, is still classical as I have merely defined a dimensionless position

$$X = \sqrt{\frac{m\omega}{\hbar}}\, q. \tag{1.1.3}$$

Now in the Hamiltonian formulation we have $p = m\dot{q}$ so that

$$H = \frac{m\omega^2 q^2}{2} + \frac{p^2}{2m} \tag{1.1.4}$$

$$= \frac{\hbar\omega}{2}(X^2 + Y^2), \tag{1.1.5}$$

where I have also defined a dimensionless momentum

$$Y = \omega^{-1}\dot{X} = \sqrt{\frac{1}{m\omega\hbar}}\, p\,. \tag{1.1.6}$$

In quantizing the harmonic oscillator we have $p \to \hat{p}$, which in the position basis has the representation $-i\hbar\frac{\partial}{\partial q}$. This means that for an arbitrary quantum state $|\psi\rangle$,

$$\langle q|\hat{p}|\psi\rangle = -i\hbar\frac{\partial}{\partial q}\langle q|\psi\rangle = -i\hbar\psi'(q), \tag{1.1.7}$$

where $\psi(q)$ is called the position wavefunction. Thus $Y \to \hat{Y} \sim -i\frac{\partial}{\partial X}$. These imply the commutation relations

$$[\hat{q}, \hat{p}] = i\hbar \Longrightarrow [\hat{X}, \hat{Y}] = i. \tag{1.1.8}$$

Now we define (non-Hermitian) *lowering* and *raising* operators

$$a = 2^{-1/2}(\hat{X} + i\hat{Y}), a^\dagger = 2^{-1/2}(\hat{X} - i\hat{Y}). \tag{1.1.9}$$

(Note that for convenience we do not use hats for these.) Then we can rewrite the Hamiltonian as

$$\hat{H} = \hbar\omega \left(a^\dagger a + aa^\dagger\right)/2. \tag{1.1.10}$$

Exercise 1.1 *From the commutation relations of $\hat{X}$ and $\hat{Y}$, show that*

$$[a, a^\dagger] = 1, \tag{1.1.11}$$

and hence that $\hat{H} = \hbar\omega \left(a^\dagger a + \frac{1}{2}\right)$.

It is simple to show that the state $|\psi_0\rangle$ with wavefunction

$$\psi_0(X) = \langle X|\psi_0\rangle \propto \exp(-X^2/2) \tag{1.1.12}$$

is an eigenstate of a with eigenvalue 0, by working in the X-basis:

$$\begin{aligned} \langle X|(\hat{X} + i\hat{Y})|\psi_0\rangle &= \left[X + i\left(-i\frac{\partial}{\partial X}\right)\right]\psi_0(X) \\ &= (X - X)\exp(-X^2/2) = 0 = \langle X|0|\psi_0\rangle. \end{aligned} \tag{1.1.13}$$

From this it is easy to show that the eigenvalues of $a^\dagger a$ are the non-negative integers, as follows. From the commutation relations (1.1.11) it follows that (for integer k)

$$[a^\dagger a, (a^\dagger)^k] = a^\dagger [a, (a^\dagger)^k] = k(a^\dagger)^k \tag{1.1.14}$$

Exercise 1.2 *Show this. Hint: Start by showing it for $k = 1$ and $k = 2$ and then find a proof by induction.*

Then, if we define an unnormalized state $|\psi_n\rangle = (a^\dagger)^n|\psi_0\rangle$ we can easily show that

$$\begin{aligned} (a^\dagger a)|\psi_n\rangle &= (a^\dagger a)(a^\dagger)^n|\psi_0\rangle \\ &= \left[(a^\dagger)^n(a^\dagger a) + n(a^\dagger)^n\right]|\psi_0\rangle = n(a^\dagger)^n|\psi_0\rangle = n|\psi_n\rangle \end{aligned} \tag{1.1.15}$$

which establishes the result and identifies the eigenstates.

Thus we have derived the eigenvalues of the harmonic oscillator as $\hbar\omega\left(n + \frac{1}{2}\right)$. The corresponding unnormalized eigenstates are $|\psi_n\rangle$, which we will denote $|n\rangle$ when normalized. These states have an integer number of elementary excitations of the vibration of the particle. They are therefore sometimes called *vibron* number states, that is, states with a definite number of vibrons.

1.2 Heisenberg and Schrödinger pictures Revision

An isolated quantum system will undergo reversible evolution governed by its Hamiltonian (that is, energy) operator $\hat{H}$. There are two basic ways of describing this time evolution, called the Schrödinger picture

(SP) and the Heisenberg picture (HP). In the former, the state of the system changes and operators are unchanging, while in the latter the state is time-independent and the operators time-dependent.

The SP evolution of the state vector is the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle. \tag{1.2.1}$$

The solution of this equation is

$$|\psi(t)\rangle = \hat{U}(t,0)|\psi(0)\rangle \tag{1.2.2}$$

where $\hat{U}$ is a unitary operator ($\hat{U}^\dagger = \hat{U}^{-1}$) given by

$$\hat{U}(t,0) = \exp(-i\hat{H}t/\hbar). \tag{1.2.3}$$

In the HP, the equation of motion for an arbitrary operator $\hat{A}$ is

$$\frac{\partial}{\partial t}\hat{A}(t) = (i\hbar)^{-1}[\hat{A}(t), \hat{H}(t)]. \tag{1.2.4}$$

Note that, because $\hat{H}(t)$ commutes with itself at a particular time t, the Hamiltonian operator is one operator that is always the same in the HP and SP. The solution of the HP equation is

$$\hat{A}(t) = \hat{U}^\dagger(t,0)\hat{A}(0)\hat{U}(t,0). \tag{1.2.5}$$

Exercise 1.3 *Show this using Eq. (1.2.3).*

The two pictures are equivalent because all expectation values are identical:

$$\begin{aligned}\langle\psi|\hat{A}(t)|\psi\rangle &= \langle\psi|\hat{U}^\dagger(t,0)\hat{A}(0)\hat{U}(t,0)|\psi\rangle \\ &= \langle\psi(t)|\hat{A}|\psi(t)\rangle.\end{aligned} \tag{1.2.6}$$

Here the placement of the time-argument t indicates which picture we are in.

We can also allow for an explicitly time-dependent Hamiltonian as would arise from a classical external modulation of the system. In this case $\hat{H}(t)$ is still the same in both pictures, but the expression for the unitary operator is much more complicated:

$$\hat{U}(t,0) = I + \sum_{n=1}^{\infty}(i\hbar)^{-n}\int_0^t ds_n\hat{H}(s_n)\int_0^{s_n} ds_{n-1}\hat{H}(s_{n-1})\cdots\int_0^{s_2} ds_1\hat{H}(s_1) \tag{1.2.7}$$

Exercise 1.4 *Derive the differential equation for $\hat{U}(t,0)$, and show that this is the solution satisfying $\hat{U}(0,0) = I$. Show also that $\hat{U}(t,0)$ is unitary [that is, that $\hat{U}^\dagger\hat{U} = 1$], by showing that $\hat{U}^\dagger(t,0)\hat{U}(t,0)$ is a constant of motion.*

In the case that $\hat{H}$ is time-independent, Eq. (1.2.7) reduces to Eq. (1.2.3)

Exercise 1.5 *Show this.*

1.3 The classical electromagnetic field

1.3.1 Maxwell's Equations

Not surprisingly, we begin with Maxwell's equations. Superficially this might seem similar to considering Schrödinger's equation, but it must be remembered that Schrödinger's equation is an equation for a wavefunction with the usual probability interpretation in quantum mechanics. Maxwell's equations are equations of motion for the *classical* electromagnetic field. They are

$$\begin{aligned} \boldsymbol{\nabla}\cdot\mathbf{E} &= \rho/\epsilon_0\,, & (1.3.1)\\ \boldsymbol{\nabla}\times\mathbf{B} &= \frac{1}{c^2}\left[\mathbf{j}/\epsilon_0 + \frac{\partial\mathbf{E}}{\partial t}\right]\,, & (1.3.2)\\ \boldsymbol{\nabla}\cdot\mathbf{B} &= 0\,, & (1.3.3)\\ \boldsymbol{\nabla}\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial t}\,. & (1.3.4) \end{aligned}$$

Here ϵ_0 is the permittivity of free space, and the permeability μ_0 has been replaced using the identity $\mu_0\epsilon_0 = c^{-2}$.

Eqn(1.3.3) and eqn(1.3.4) are automatically satisfied by introducing the potentials ϕ and $\mathbf{A}$, so that

$$\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}\,,\ \mathbf{E} = -\boldsymbol{\nabla}\phi\ - \frac{\partial\mathbf{A}}{\partial t}\,. \qquad (1.3.5)$$

Eqn(1.3.1) and eqn(1.3.2) then become

$$\begin{aligned} -\nabla^2\phi - \frac{\partial}{\partial t}\boldsymbol{\nabla}\cdot\mathbf{A} &= \rho/\epsilon_0\,, & (1.3.6)\\ \left(\frac{\partial^2}{\partial t^2} - c^2\nabla^2\right)\mathbf{A} + \boldsymbol{\nabla}\left(\frac{\partial\phi}{\partial t} + c^2\boldsymbol{\nabla}\cdot\mathbf{A}\right) &= \mathbf{j}/\epsilon_0\,. & (1.3.7) \end{aligned}$$

Here the identity $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\ = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\ ) - \nabla^2$ has been used.

1.3.2 The Coulomb Gauge

The potentials are not determined uniquely, since the replacements

$$\phi \to \phi' = \phi + \dot{f}\,,\ \mathbf{A} \to \mathbf{A}' = \mathbf{A} - \boldsymbol{\nabla} f\,, \qquad (1.3.8)$$

for arbitrary $f(\mathbf{r},t)$, leave the fields unaltered. (Here the notations $\dot{f}$, $\frac{\partial f}{\partial t}$ and $\partial_t f$ will be used interchangeably.) This is known as a *gauge transformation*. It can be shown that we can choose the gauge such that

$$\boldsymbol{\nabla}\cdot\mathbf{A} = 0. \qquad (1.3.9)$$

This is known as the *Coulomb* or *radiation* gauge. In this gauge, Maxwell's equations become a wave equation for $\mathbf{A}$:

$$\left(\partial_t^2 - c^2\nabla^2\right)\mathbf{A} = \mathbf{j}/\epsilon_0 - \boldsymbol{\nabla}\dot{\phi}\,, \qquad (1.3.10)$$

and Poisson's equation for ϕ:

$$\nabla^2\phi = -\rho/\epsilon_0. \qquad (1.3.11)$$

The latter has the solution

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int \mathrm{d}^3\mathbf{x}\frac{\rho(\mathbf{x})}{|\mathbf{r}-\mathbf{x}|}. \tag{1.3.12}$$

Exercise 1.6 *For a single charge q at position $\mathbf{q}$, verify that this gives the usual Coulomb's law.*

In the Coulomb gauge the vector potential $\mathbf{A}$ is what is known as a *transverse* field, because $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$. It can be shown that an arbitrary vector field can be split into transverse and longitudinal parts. For example, the current density

$$\mathbf{j} = \mathbf{j}^\perp + \mathbf{j}^\parallel \; ; \quad \boldsymbol{\nabla}\cdot\mathbf{j}^\perp = 0 \; , \quad \boldsymbol{\nabla}\times\mathbf{j}^\parallel = \mathbf{0}. \tag{1.3.13}$$

These components are everywhere orthogonal:

$$\mathbf{j}^\perp\cdot\mathbf{j}^\parallel \equiv 0 \tag{1.3.14}$$

Now using $\boldsymbol{\nabla}\times\boldsymbol{\nabla} \equiv 0$ one can split Eq. (1.3.10) into transverse and longitudinal parts. This gives a wave equation with sources for $\mathbf{A}$

$$\left(\partial_t^2 - c^2\nabla^2\right)\mathbf{A} = \mathbf{j}^\perp/\epsilon_0 \tag{1.3.15}$$

and the following identity

$$\mathbf{j}^\parallel/\epsilon_0 = \boldsymbol{\nabla}\dot{\phi}. \tag{1.3.16}$$

This last equation, with Eq. (1.3.11), and the fact that $\boldsymbol{\nabla}\cdot bldj^\parallel = \boldsymbol{\nabla}\cdot\mathbf{j}$ then imply the continuity equation

$$\dot{\rho} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0. \tag{1.3.17}$$

If the continuity equation is taken for granted, the only independent equations are then the wave equation with sources (1.3.15) and Poisson's equation (1.3.11). We will return to these equations in Chapter 3.

1.3.3 Free Fields

Now consider the case of fields with no current sources: $\mathbf{j}^\perp = \mathbf{0}$. Then $\mathbf{A}$ satisfies the pure wave equation

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} = 0, \tag{1.3.18}$$

which has propagating solutions so $\mathbf{A}$ is not necessarily zero. It will be convenient to represent the field by a discrete set of variables rather than a continuum, so we shall assume that the field is confined to a cube of side L with the origin at one corner. Assuming periodic boundary conditions (this is an arbitrary assumption), we consider standing waves, with wavevectors

$$\mathbf{k} = \frac{2\pi}{L}\,(l,m,n)^T\,, \tag{1.3.19}$$

where l, m, n are integers, and two transverse polarization vectors

$$\boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}} \; : \; \boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}} \cdot \mathbf{k} = 0 \; , \;\; \boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}} \cdot \boldsymbol{\varepsilon}^{\lambda'}_{\mathbf{k}} = \delta^{\lambda,\lambda'} . \tag{1.3.20}$$

We can thus write the solution in terms of dimensionless coefficients $X^{\lambda}_{s,\mathbf{k}}$ and $X^{\lambda}_{c,\mathbf{k}}$ as

$$\mathbf{A}(\mathbf{x}) = \sum_{\mathbf{k},\lambda} \boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}} \sqrt{\frac{2\hbar}{\epsilon_0 c |\mathbf{k}| L^3}} \left[X^{\lambda}_{s,\mathbf{k}} \sin(\mathbf{k}\cdot\mathbf{x}) + X^{\lambda}_{c,\mathbf{k}} \cos(\mathbf{k}\cdot\mathbf{x}) \right] \tag{1.3.21}$$

Once again, the presence of $\hbar$ here is just a convenient normalization. Because sin and cos remain the same (up to a sign change) when the sign of their arguments are reversed, the modes with wavenumbers $\mathbf{k}$ and $-\mathbf{k}$ are actually the same mode. Thus to avoid redundancy it is necessary that the sum over all $\mathbf{k}$ actually be only one half of all possible $\mathbf{k}$. We can impose this restriction by demanding for example that in Eq. (1.3.19), $l \geq 0$.

Now the electric field can be split into transverse and longitudinal parts

$$\mathbf{E}^{\perp} = -\dot{\mathbf{A}} \tag{1.3.22}$$

$$\mathbf{E}^{\parallel} = -\boldsymbol{\nabla}\phi \tag{1.3.23}$$

From the above expression (1.3.21) we get

$$\mathbf{E}^{\perp}(\mathbf{x},t) = -\frac{\partial}{\partial t}\mathbf{A} = -\sum_{\mathbf{k},\lambda} \boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}} \sqrt{\frac{2\hbar}{\epsilon_0 c |\mathbf{k}| L^3}} \left[\dot{X}^{\lambda}_{s,\mathbf{k}} \sin(\mathbf{k}\cdot\mathbf{x}) + \dot{X}^{\lambda}_{c,\mathbf{k}} \cos(\mathbf{k}\cdot\mathbf{x}) \right] \tag{1.3.24}$$

$$\mathbf{B}(\mathbf{x},t) = \boldsymbol{\nabla} \times \mathbf{A} = \sum_{\mathbf{k},\lambda} \hat{\mathbf{k}} \times \boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}} \sqrt{\frac{2|\mathbf{k}|\hbar}{\epsilon_0 c L^3}} \left[X^{\lambda}_{s,\mathbf{k}} \cos(\mathbf{k}\cdot\mathbf{x}) - X^{\lambda}_{c,\mathbf{k}} \sin(\mathbf{k}\cdot\mathbf{x}) \right] \tag{1.3.25}$$

where $\hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|$. From Eq. (1.3.12), $\mathbf{E}^{\parallel}$ can be expressed directly in terms of the charge density ρ.

Now since $\mathbf{E}^{\perp} \cdot \mathbf{E}^{\parallel} = 0$, the electromagnetic field energy can be split into the radiation field energy is

$$E = \int_0^L \mathrm{d}x \int_0^L \mathrm{d}y \int_0^L \mathrm{d}z \, \frac{\epsilon_0}{2} \left((\mathbf{E}^{\perp})^2 + c^2 \mathbf{B}^2 \right) . \tag{1.3.26}$$

plus the integral of the Coulomb energy density, $\epsilon_0 (\mathbf{E}^{\parallel})^2/2$. The latter can be expressed in terms of the charge density ρ alone and so need not be considered independently of the matter fields which give rise to these charges. We will return to it in Chapter 3. However, the radiation field can store energy independent of any charges and we find from the above expressions

$$E = \sum_{\mathbf{k},\lambda} \sum_{m=s,c} \frac{\hbar\omega_{\mathbf{k}}}{2} \left[\left(X^{\lambda}_{m,\mathbf{k}} \right)^2 + \omega_{\mathbf{k}}^{-2} \left(\dot{X}^{\lambda}_{m,\mathbf{k}} \right)^2 \right] , \tag{1.3.27}$$

where we have defined $\omega_{\mathbf{k}} = c|\mathbf{k}|$, and m stands for mode (sin or cos). Thus we see that the radiation field consists of an infinite sum of independent harmonic oscillators of frequencies $\omega_{\mathbf{k}}$.

Exercise 1.7 *Verify Eq. (1.3.27). Your expression should initially contain double sums of the form $\sum_{\mathbf{k},\lambda,m}\sum_{\mathbf{k}',\lambda',m'}$, and the integrals will yield Kronecker δ functions which reduce these to the single sums in Eq. (1.3.27).*

1.4 Quantizing the electromagnetic field

We now write $Y^{\lambda}_{m,\mathbf{k}} = \omega_{\mathbf{k}}^{-1}\dot{X}_{m,\mathbf{k}}$ to get the Hamiltonian form of the energy

$$H = \sum_{\mathbf{k},\lambda,m} \frac{\hbar\omega_{\mathbf{k}}}{2}\left[\left(X^{\lambda}_{m,\mathbf{k}}\right)^2 + \left(Y^{\lambda}_{m,\mathbf{k}}\right)^2\right]. \tag{1.4.1}$$

When this is quantized we have, as in Sec. 1.1,

$$[\hat{X}^{\lambda}_{m,\mathbf{k}}, \hat{Y}^{\lambda'}_{m',\mathbf{k}'}] = i\delta_{\mathbf{k},\mathbf{k}'}\delta_{\lambda,\lambda'}\delta_{m,m'}. \tag{1.4.2}$$

Using these relations, and the following identity

$$[\hat{C}^2, \hat{D}] = \hat{C}[\hat{C}, \hat{D}] + [\hat{C}, \hat{D}]\hat{C}, \tag{1.4.3}$$

we get the Heisenberg equations of motion

$$\frac{d}{dt}\hat{X}^{\lambda}_{m,\mathbf{k}} = (i\hbar)^{-1}[\hat{X}^{\lambda}_{m,\mathbf{k}}, \hat{H}] = \omega_{\mathbf{k}}\hat{Y}^{\lambda}_{m,\mathbf{k}} \tag{1.4.4}$$

$$\frac{d}{dt}\hat{Y}^{\lambda}_{m,\mathbf{k}} = (i\hbar)^{-1}[\hat{Y}^{\lambda}_{m,\mathbf{k}}, \hat{H}] = -\omega_{\mathbf{k}}\hat{X}^{\lambda}_{m,\mathbf{k}}. \tag{1.4.5}$$

these are as expected for a Harmonic oscillator and are consistent with the above definition of $\hat{Y}^{\lambda}_{\mathbf{k}}$. They also imply

$$\frac{d^2}{dt^2}\hat{X}^{\lambda}_{m,\mathbf{k}} = -\omega^2_{\mathbf{k}}\hat{X}^{\lambda}_{m,\mathbf{k}} \tag{1.4.6}$$

From this it is easy to show that the original expression for the vector potential (1.3.21) is indeed a solution of Maxwell's wave equation (1.3.18).

Exercise 1.8 *Verify this.*

We have in fact completed the quantization of the free electromagnetic field, without mentioning the word photon. We introduce this concept simply by defining an annihilation operator for each mode of the field

$$a^{\lambda}_{m,\mathbf{k}} = 2^{-1/2}\left(\hat{X}^{\lambda}_{m,\mathbf{k}} + i\hat{Y}^{\lambda}_{m,\mathbf{k}}\right), \tag{1.4.7}$$

exactly analogously to Sec. 1.1. Then the field Hamiltonian can be written

$$\hat{H} = \sum_{\mathbf{k},\lambda,t} \hbar\omega_{\mathbf{k}}\left[\hat{n}^{\lambda}_{m,\mathbf{k}} + \tfrac{1}{2}\right], \tag{1.4.8}$$

where $\hat{n}^\lambda_{m,\mathbf{k}} = (a^\lambda_{m,\mathbf{k}})^\dagger a^\lambda_{m,\mathbf{k}}$. Here the *photon* annihilation and creation operators obey

$$\left[a^\lambda_{m,\mathbf{k}}, (a^{\lambda'}_{m',\mathbf{k}'})^\dagger\right] = \delta_{\mathbf{k},\mathbf{k}'}\delta_{\lambda,\lambda'}\delta_{m,m'}. \tag{1.4.9}$$

We have thus arrived at a description with a similar structure to the field-theory description of a system of noninteracting bose particles. (It describes bosons, not fermions, because there is no limit to the number of photons in a mode.) However, instead of quantizing the particles and then constructing the field, we have quantized the field and found the "particles" (photons). I prefer to maintain a distinction between particles such as electrons, and quanta such as photons. One can be misled if thinking of photons as particles that it should be possible to write down a wavefunction $\psi(x)$ for a single photon with $|\psi(x)|^2$ interpretable as a photon density. This is not possible. If one wished to introduce a wavefunction representation, it would be a function

$$\psi(\{X^\lambda_{m,\mathbf{k}}\}) = \psi(X^1_{s,\mathbf{k}_1}, X^2_{s,\mathbf{k}_1}, X^1_{c,\mathbf{k}_1}, X^2_{c,\mathbf{k}_1}, X^1_{s,\mathbf{k}_2}, X^2_{s,\mathbf{k}_2}, \ldots) \tag{1.4.10}$$

of the infinite set of canonical coordinates $X^\lambda_{m,\mathbf{k}}$. Remember these coordinates are *not* the positions of photon with momentum $\hbar\mathbf{k}$ *et cetera*. They are the amplitudes of the vector potential for the field mode with wavenumber $\mathbf{k}$ *et cetera*.

1.4.1 * Field Commutation Relations

Having quantized the electromagnetic fields we can now write expressions for the fields $\hat{\mathbf{A}}$ and $\hat{\mathbf{E}}^\perp = -\dot{\hat{\mathbf{A}}} = i[\hat{\mathbf{A}}, \hat{H}]/\hbar$ in terms of the annihilation and creation operators. The result is

$$\hat{\mathbf{A}}(\mathbf{x}) = \sum_{\mathbf{k},\lambda} \boldsymbol{\varepsilon}^\lambda_{\mathbf{k}} \sqrt{\frac{\hbar}{\epsilon_0 \omega_{\mathbf{k}} L^3}} \left\{[a^\lambda_{s,\mathbf{k}} + (a^\lambda_{s,\mathbf{k}})^\dagger]\sin(\mathbf{k}\cdot\mathbf{x}) + [a^\lambda_{c,\mathbf{k}} + (a^\lambda_{c,\mathbf{k}})^\dagger]\cos(\mathbf{k}\cdot\mathbf{x})\right\} \tag{1.4.11}$$

$$\hat{\mathbf{E}}^\perp(\mathbf{x}) = \sum_{\mathbf{k},\lambda} \boldsymbol{\varepsilon}^\lambda_{\mathbf{k}} \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{\epsilon_0 L^3}}\, i \left\{[a^\lambda_{s,\mathbf{k}} - (a^\lambda_{s,\mathbf{k}})^\dagger]\sin(\mathbf{k}\cdot\mathbf{x}) + [a^\lambda_{c,\mathbf{k}} - (a^\lambda_{c,\mathbf{k}})^\dagger]\cos(\mathbf{k}\cdot\mathbf{x})\right\} \tag{1.4.12}$$

From these it can be show that

$$[\hat{A}_j(\mathbf{r}), \hat{A}_l(\mathbf{r}')] = 0 \tag{1.4.13}$$

$$[\hat{E}^\perp_j(\mathbf{r}), \hat{E}^\perp_l(\mathbf{r}')] = 0 \tag{1.4.14}$$

$$[\hat{A}_j(\mathbf{r}), \hat{E}^\perp_l(\mathbf{r}')] = \frac{-i\hbar}{\epsilon_0}\frac{2}{L^3}\sum_{\mathbf{k},\lambda}(\boldsymbol{\varepsilon}^\lambda_{\mathbf{k}})_j(\boldsymbol{\varepsilon}^\lambda_{\mathbf{k}})_l \left[\sin\mathbf{k}\cdot\mathbf{r}\,\sin\mathbf{k}\cdot\mathbf{r}' + \cos\mathbf{k}\cdot\mathbf{r}\,\cos\mathbf{k}\cdot\mathbf{r}'\right] \tag{1.4.15}$$

Exercise 1.9 *Show Eq. (1.4.15).*

These equations mean that we could simultaneously measure **A** for all points in space, or **E** for all points in space, but not both.

Using $(\boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}})_l(\boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}})_j + \hat{k}_j\hat{k}_l = \delta_{jl}$ enables us to write

$$[\hat{A}_j(\mathbf{r}), \hat{E}^{\perp}_l(\mathbf{r}')] \quad = \quad (-i\hbar/\epsilon_0)\delta^{\perp}_{jl}(\mathbf{r}-\mathbf{r}'), \tag{1.4.16}$$

where the symbol $\delta^{\perp}_{jl}(\mathbf{r}-\mathbf{r}')$ is known as the *transverse* δ function, and can be defined by

$$\delta^{\perp}_{jl}(\mathbf{x}) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}}\left(\delta_{jl} - \hat{k}_j\hat{k}_l\right). \tag{1.4.17}$$

It is distinct from the usual δ function

$$\delta_{jl}(\mathbf{x}) = \delta_{jl}\frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{1.4.18}$$

which is defined so that for some vector field $\mathbf{C}(\mathbf{x})$,

$$\int d^3\mathbf{r}C_j(\underline{\mathrm{r}})\delta_{jk}(\mathbf{x}-\mathbf{r})\mathbf{e}_k = \mathbf{C}(\mathbf{x}). \tag{1.4.19}$$

To see that (1.4.16) is equivalent to the definition in Eq. (1.4.15) requires that we remember that the fields are defined only in the region $0 \leq x, y, z \leq L$, and taking the limit $L \to \infty$. The result using the transverse δ function is superior because it is independent of the mode functions used to quantize the field. The results (1.4.13)–(1.4.16) will be used later.

2 Quantum States of the Electromagnetic Field

2.1 Single-Mode Pure States

For simplicity, consider a single sin mode of the electromagnetic field with spatial frequency $\mathbf{k}$ and polarization vector $\boldsymbol{\varepsilon}_{\mathbf{k}}^{\lambda}$. This mode has a single co-ordinate, $\hat{X} = \hat{X}_{s,\mathbf{k}}^{\lambda}$, proportional to the amplitude of the vector potential. As described in Sec. 1.4 it is possible to represent the state of the field as a wavefunction $\psi(X)$, in this case of just a single co-ordinate X. This is completely analogous to the wavefunction of a single particle. We can define the eigenstates of the operator $\hat{X} = (a + a^{\dagger})/\sqrt{2}$ as $|X\rangle$ in which case we have

$$\psi(X) = \langle X|\psi\rangle\,. \tag{2.1.1}$$

2.1.1 Number States

The Hamiltonian for this single mode is just

$$\hat{H} = \hbar\omega a^{\dagger}a. \tag{2.1.2}$$

Here we have dropped the $\frac{1}{2}\hbar\omega$ because this is just a constant. In all that follows we will follow this practice, which simply amounts to setting the energy to be zero when there are no photons present.

Because the vector potential amplitude operator X does not commute with the Hamiltonian, it is more convenient to use a different basis, namely the eigenstates $|n\rangle$ of $a^{\dagger}a$. Since $a^{\dagger}a$ can be interpreted as the photon number operator, these are known as photon number eigenstates, or more simply, *number states*.

It was shown in Sec. 1.1 that the number states $|n\rangle$ satisfies

$$|n\rangle \propto (a^{\dagger})^{n}|0\rangle \tag{2.1.3}$$

where $|0\rangle$ is the *vacuum* state, the state with no excitations. That is, $a^{\dagger}$ (the creation operator) raises the number of photons by one. Similarly, a lowers it by one. However, we require the number states to be normalized, so that

$$\langle n|m\rangle = \delta_{nm}. \tag{2.1.4}$$

Now since $|n\rangle$ is an eigenstate of $a^{\dagger}a$ of eigenvalue n,

$$\langle n|\big[a^{\dagger}a|n\rangle\big] = n\,\langle n|n\rangle = n. \tag{2.1.5}$$

However we also have

$$\big[\langle n|a^{\dagger}\big]\big[a|n\rangle\big] = \langle\psi|\psi\rangle\,, \tag{2.1.6}$$

where $|\psi\rangle = a|n\rangle \propto |n-1\rangle$. Therefore the constant of proportionality must be

$$|\psi\rangle = a|n\rangle = e^{i\phi}\sqrt{n}|n-1\rangle \tag{2.1.7}$$

for some phase ϕ. We choose the convention that $\phi = 0$, so that

$$a|n\rangle = \sqrt{n}|n-1\rangle. \tag{2.1.8}$$

With this convention it can be shown similarly that

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{2.1.9}$$

Exercise 2.1 *Show this, in particular showing that once $\phi = 0$ has been chosen in Eq. (2.1.7), Eq. (2.1.9) is the unique solution guaranteeing $[a^\dagger, a] = 1$. Show also that the above two relations are consistent with $|n\rangle$ being an eigenstate of $a^\dagger a$.*

Note that a acting on the vacuum state $|0\rangle$ produces nothing, a null state.

Exercise 2.2 *Show that the normalized number state is $|n\rangle = (n!)^{-1/2}(a^\dagger)^n|0\rangle$.*

2.1.2 Coherent States

The electromagnetic field does not have to be in an eigenstate of the Hamiltonian (in fact, it is very difficult to prepare a mode of the field in a number state, except for the number zero). Also, no matter how large the number state, the system never approaches the classical limit. That is because for a system in a number state the average value of the electric and magnetic fields is always zero.

Exercise 2.3 *Show this.*

For this reason, it is useful to consider a state for which there is a classical limit, the *coherent* state. This state is defined as an eigenstate of the annihilation operator

$$a|\alpha\rangle = \alpha|\alpha\rangle \tag{2.1.10}$$

where α is any complex number (remember a is not an Hermitian operator). There are no such eigenstates of the creation operator $a^\dagger$.

Exercise 2.4 *Show this. Assume that there exists states $|\beta\rangle$ such that $a^\dagger|\beta\rangle = \beta|\beta\rangle$ and consider the inner product $\langle n|(a^\dagger)^{n+1}|\beta\rangle$. Hence show that the inner product of $|\beta\rangle$ with any number state is zero.*

It is easy to find an expression for $|\alpha\rangle$ in terms of the number states as follows. In general we have

$$|\alpha\rangle = \sum_{n=0}^{\infty} c_n|n\rangle. \tag{2.1.11}$$

Since $a|\alpha\rangle = \alpha|\alpha\rangle$ we get

$$\sum_{n=0}^{\infty} \sqrt{n}c_n|n-1\rangle = \sum_{n=0}^{\infty} \alpha c_n|n\rangle. \tag{2.1.12}$$

In the first sum the $n = 0$ term does not contribute because of the $\sqrt{n}$. Therefore we can change the limits of the sum to $n = 1 \cdots \infty$, then change variables from n to $m = n - 1$ to get

$$\sum_{m=0}^{\infty} \sqrt{m+1}c_{m+1}|m\rangle = \sum_{n=0}^{\infty} \alpha c_n|n\rangle \tag{2.1.13}$$

Equating the co-efficients of the number states on both sides we get the recursion relation

$$c_{n+1} = \frac{\alpha}{\sqrt{n+1}} c_n. \tag{2.1.14}$$

Thus $c_1 = \alpha c_0$, $c_2 = \alpha c_1/\sqrt{2} = \alpha^2 c_0/\sqrt{2}$ *et cetera* so that

$$c_n = \frac{\alpha^n}{\sqrt{n!}} c_0. \tag{2.1.15}$$

To normalize the state we require

$$1 = \sum_n |c_n|^2 = |c_0|^2 \sum_n \frac{\left(|\alpha|^2\right)^n}{n!}. \tag{2.1.16}$$

The final sum is simply the series expansion for the exponential $\exp\left(|\alpha|^2\right)$. Thus $|c_0|^2 = \exp\left(-|\alpha|^2\right)$ and without loss of generality we can take $c_0 = \exp\left(-|\alpha|^2/2\right)$. Thus we have

$$|\alpha\rangle = \exp\left(-|\alpha|^2/2\right) \sum_{n=0} \infty \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \tag{2.1.17}$$

The state $|0\rangle$ (where here 0 is the value of α) is the same state as the state $|0\rangle$ (where here 0 is the value of n). In other words, the vacuum state is both a coherent state and a number state.

Exercise 2.5 *Verify this.*

For α finite the coherent state has a non-zero mean photon number:

$$\langle\alpha|a^\dagger a|\alpha\rangle = |\alpha|^2 \langle\alpha|\alpha\rangle = |\alpha|^2. \tag{2.1.18}$$

Here we have used $\langle\alpha|a^\dagger = (a|\alpha\rangle)^\dagger = (\alpha|\alpha\rangle)^\dagger = \langle\alpha|\alpha^*$. The photon number distribution (the probability for measuring a certain photon number) for a coherent state is just $|c_n|^2$, which is a *Poissonian* distribution of mean $|\alpha|^2$:

$$P_n = |\langle n|\alpha\rangle|^2 = e^{-|\alpha|^2} \frac{\left(|\alpha|^2\right)^n}{n!} \tag{2.1.19}$$

This distribution has the property that the variance is equal to the mean. That is,

$$\left\langle (a^\dagger a)^2 \right\rangle - \left\langle a^\dagger a \right\rangle^2 = |\alpha|^2 \tag{2.1.20}$$

Exercise 2.6 *Verify this, either from the distribution (2.1.19) or directly from the coherent state using the commutation relations for a and $a^\dagger$.*

Unlike a number state, a field in a coherent state can have a mean field. Specifically, if all of the other modes are in a number state then

$$\left\langle \hat{\mathbf{E}} \right\rangle (\mathbf{x}) = \mathbf{E}_0(\mathbf{x}) \left\langle -ia + ia^\dagger \right\rangle /2 = \mathbf{E}_0(\mathbf{x}) \langle\alpha|(-ia + ia^\dagger)|\alpha\rangle/2 = \mathbf{E}_0(\mathbf{x})(-i\alpha + i\alpha^*)/2. \tag{2.1.21}$$

Here $\mathbf{E}_0(\mathbf{x})$ is the mode function of the single mode in the coherent state. For the quantization procedure used in Chapter 1 this may for example be equal to

$$\mathbf{E}_0(\mathbf{x}) = \boldsymbol{\varepsilon}_{\mathbf{k}}^{\lambda}\sqrt{\frac{\hbar\omega_{\mathbf{k}}}{\epsilon_0 L^3}}\, 2\sin(\mathbf{k}\cdot\mathbf{x}) \tag{2.1.22}$$

This quantity can roughly be thought of as the electric field per photon, because for a coherent state $\alpha = i$, which has a mean occupation of one photon ($|\alpha|^2 = 1$), $\left\langle \hat{\mathbf{E}}(\mathbf{x}) \right\rangle = \mathbf{E}_0(\mathbf{x})$.

Because a is not an Hermitian operator, the coherent states are not orthogonal. In fact it can be shown that

$$|\left\langle \alpha|\alpha' \right\rangle|^2 = \exp(-|\alpha - \alpha'|^2). \tag{2.1.23}$$

If α and α' are very different (as they would be if they represent two macroscopically distinct fields) then the two coherent states are very nearly orthogonal. Another consequence of their nonorthogonality is that the coherent states form an *overcomplete* basis. Whereas for number states we have

$$\sum_n |n\rangle\langle n| = I, \tag{2.1.24}$$

the identity, for coherent states we have

$$\int \mathrm{d}^2\alpha |\alpha\rangle\langle\alpha| = \pi I. \tag{2.1.25}$$

Exercise 2.7 *Show this using the expansion (2.1.17). The result* $n! = \int_0^\infty dx x^n e^{-x}$ *may be useful.*

Unlike number states, coherent states are not eigenstates of the Hamiltonian. However, they have the nice property that they remain coherent states under the free Hamiltonian $\hat{H} = \hbar\omega a^\dagger a$. The amplitude $|\alpha|$ of the states remain the same; only the phase changes at rate ω (as expected):

$$\exp(-i\hat{H}t/\hbar)|\alpha\rangle = |e^{-i\omega t}\alpha\rangle \tag{2.1.26}$$

Exercise 2.8 *Show this, using Eq. (2.1.17).*

2.2 Single Mode Mixed States

2.2.1 Mixed States

All of the sates considered so far have been pure states. That is, states which can be represented by a state vector

$$|\psi\rangle = \sum_n \psi_n |n\rangle \tag{2.2.1}$$

This state can also be represented by a *state matrix*

$$\rho = \sum_{n,m} \rho_{n,m}|n\rangle\langle m| = |\psi\rangle\langle\psi|, \tag{2.2.2}$$

where

$$\rho_{n,m} = \psi_n \psi_m^*. \tag{2.2.3}$$

The object ρ is sometimes also called a density operator. Using it, the expectation values of an operator $\hat{O}$ can be written

$$\left\langle \hat{O} \right\rangle = \mathrm{Tr}[\hat{O}\rho], \tag{2.2.4}$$

where the trace of an operator is defined as

$$\mathrm{Tr}[\hat{A}] = \sum_n \langle n|\hat{A}|n\rangle \tag{2.2.5}$$

using any complete basis $\{|n\rangle\}$ (such as the number states). This can be seen as follows:

$$\begin{aligned}
\mathrm{Tr}[\hat{O}\rho] &\equiv \sum_n \langle n|\hat{O}\rho|n\rangle && (2.2.6)\\
&= \sum_{n,m} \langle n|\hat{O}|m\rangle\langle m|\rho|n\rangle && (2.2.7)\\
&= \sum_{n,m} \langle n|\hat{O}|m\rangle \rho_{m,n} && (2.2.8)\\
&= \sum_{n,m} \langle n|\hat{O}|m\rangle \psi_m \psi_n^* && (2.2.9)\\
&= \sum_{n,m} \langle \psi|n\rangle \, \langle n|\hat{O}|m\rangle \, \langle m|\psi\rangle && (2.2.10)\\
&= \langle \psi|\hat{O}|\psi\rangle = \left\langle \hat{O} \right\rangle && (2.2.11)
\end{aligned}$$

The big advantage of the state matrix formalism is that we can represent mixed states. These are states which are in a classical probabilistic mixture of different pure states [1]. If the system were in state $|\psi^1\rangle$ with probability p_1 and in state $|\psi^2\rangle$ with probability p_2 then the average value of an operator O would clearly be

$$\left\langle \hat{O} \right\rangle = p_1 \langle \psi^1|\hat{O}|\psi^1\rangle + p_2 \langle \psi^2|\hat{O}|\psi^2\rangle \tag{2.2.12}$$

This result is also obtained from the formula

$$\left\langle \hat{O} \right\rangle = \mathrm{Tr}[\rho\hat{O}] \tag{2.2.13}$$

if we define the state matrix for the mixed state by

$$\rho = p_1 |\psi^1\rangle\langle\psi^1| + p_2 |\psi^2\rangle\langle\psi^2| \tag{2.2.14}$$

Exercise 2.9 *Show this.*

[1] Sometimes superpositions as in Eq. (2.2.1) are called coherent mixtures. Here when I say mixture I mean an incoherent mixture.

In general a state matrix can be written as

$$\rho = \sum_\mu p_\mu |\psi^\mu\rangle\langle\psi^\mu|, \tag{2.2.15}$$

where $p_\mu \geq 0$ and $\sum_\mu p_\mu = 1$. This gives

$$\rho_{nm} = \sum_\mu p_\mu \psi_n^\mu (\psi_m^\mu)^*. \tag{2.2.16}$$

From the Schrödinger equation (1.2.1) for the state vector, it is easy to show that the analogous equation of motion for the state matrix is

$$(i\hbar)\frac{\partial}{\partial t}\rho(t) = [\hat{H}(t), \rho(t)]. \tag{2.2.17}$$

Note that this is different (and has a different interpretation) from the Heisenberg picture equation of motion Eq. (1.2.4) for an operator $\hat{A}(t)$.

As a special example of Eq. (2.2.13), the normalization of a state matrix is

$$1 = \mathrm{Tr}[\rho] = \sum_n \rho_{nn}. \tag{2.2.18}$$

For pure states it can be shown that

$$\mathrm{Tr}[\rho^2] = \mathrm{Tr}[|\psi\rangle\langle\psi|\psi\rangle\langle\psi|] = \mathrm{Tr}[|\psi\rangle\langle\psi|] = \langle\psi|\psi\rangle = 1. \tag{2.2.19}$$

For a general mixed state

$$0 \leq \mathrm{Tr}[\rho^2] \leq 1, \tag{2.2.20}$$

and this quantity is a measure of how mixed the state is.

Exercise 2.10 *Prove Eq. (2.2.20).*

2.2.2 Single-Mode Thermal States

Mixed states arise often in nature because of uncontrollable fluctuations, such as thermal noise which arises when the system is coupled to a heat bath at finite temperature. A single mode of the electromagnetic field coupled to such a bath comes to an equilibrium mixed state called a single mode thermal state. From thermodynamics we know that the probability for the system to occupy a state of energy E is proportional to $\exp(-E/k_BT)$. For a single harmonic oscillator we know the energy states $|n\rangle$ have eigenvalues $\hbar\omega n$. Therefore the single mode thermal state has the state matrix

$$\rho = \sum_{n=0}^{\infty} p_n |n\rangle\langle n|, \tag{2.2.21}$$

where $p_n = Z^{-1}\exp(-\hbar\omega n/k_BT)$, where Z is the partition function (needed for normalization) given by

$$Z = \sum_{n=0}^{\infty} \exp(-\hbar\omega n/k_BT) = \frac{1}{1-\exp(-\hbar\omega/k_BT)}. \tag{2.2.22}$$

Thus the full expression is

$$\rho_{\rm th} = [1 - \exp(-\hbar\omega/k_B T)] \sum_{n=0}^{\infty} \exp(-n\hbar\omega/k_B T)|n\rangle\langle n|. \tag{2.2.23}$$

From Eq. (2.2.23) it is easy to show that the mean photon number of a SMT state is

$$\begin{aligned} \langle a^\dagger a \rangle &= \sum_{n=0}^{\infty} p_n n \;=\; (1-r)\sum_{n=0}^{\infty} r^n n & (2.2.24)\\ &= r(1-r)\sum_{n=0}^{\infty} r^{n-1} n \;=\; r(1-r)\frac{\partial}{\partial r}\sum_{n=0}^{\infty} r^n & (2.2.25)\\ &= r(1-r)\frac{\partial}{\partial r}\frac{1}{1-r} = \frac{r}{1-r} & (2.2.26) \end{aligned}$$

where $r = \exp(-\hbar\omega/k_B T)$. That is,

$$\bar{n}_{\rm th}(\omega) = \frac{1}{\exp(\hbar\omega/k_B T) - 1} \tag{2.2.27}$$

This should be familiar as Planck's formula.

$$\bar{E}_{\rm th}(\omega) = \frac{\hbar\omega}{\exp(\hbar\omega/k_B T) - 1}. \tag{2.2.28}$$

We can then rewrite a SMT state as follows, with $\bar{n} = \bar{n}_{\rm th}(\omega)$,

$$\rho_{\rm th}\left(\bar{n}\right) = \frac{1}{1+\bar{n}} \sum_{n=0}^{\infty} \left(\frac{\bar{n}}{\bar{n}+1}\right)^n |n\rangle\langle n|. \tag{2.2.29}$$

Following a similar procedure to that use to derive the mean, it can be shown that the variance of a SMT state is

$$\left\langle (a^\dagger a)^2 \right\rangle - \left\langle a^\dagger a \right\rangle^2 = \bar{n}^2 + \bar{n}. \tag{2.2.30}$$

Exercise 2.11 *Show this.*

Thus in contrast to a coherent state, where for $|\alpha|^2 \gg 1$ the uncertainty in the photon number is much less than the mean, for a thermal state the uncertainty is always at least as large as the mean. That is, the photon number is not very well defined for a thermal state.

The SMT state can also be expressed in terms of coherent states as

$$\rho = \frac{1}{\pi\bar{n}} \int \mathrm{d}^2\alpha \exp(-|\alpha|^2/\bar{n}) |\alpha\rangle\langle\alpha|. \tag{2.2.31}$$

Exercise 2.12 *Show this. Use* $\int dx x^n e^{-\gamma x} = \gamma^{-(n+1)} n!$

This is an example of how mixed states can always be represented in (infinitely many) different ways in terms of pure states.

2.2.3 Single-Mode Laser States

Another example of a single mode mixed state is the state inside a laser cavity. It is sometimes stated in textbooks that the state of the laser mode is a coherent state $|\alpha\rangle$. While the coherent state does have the same photon statistics as the state of an ideal laser cavity, it is not strictly correct to say that it is in a coherent state. The reason is that whenever a laser is turned on, the *phase* of the field $\arg\alpha$ cannot be predicted beforehand. Rather, it takes on a random value. Furthermore, that value changes randomly (diffuses) over a characteristic time $\tau = 1/\Gamma$. Here Γ is called the laser linewidth. Thus at any time the laser is in a mixture of all possible phases. We can write the state corresponding to this as

$$\rho_{\text{laser}} = \int \frac{d\phi}{2\pi} |\sqrt{\mu}\, e^{i\phi}\rangle\langle\sqrt{\mu}\, e^{i\phi}|, \tag{2.2.32}$$

where here $|\sqrt{\mu}\, e^{i\phi}\rangle$ is a a coherent state. Equivalently, this state is equal to

$$\rho_{\text{laser}}(\mu) = \sum_{n=0}^{\infty} e^{-\mu} \frac{\mu^n}{n!} |n\rangle\langle n|. \tag{2.2.33}$$

Here $\mu = |\alpha|^2$ is the mean photon number of the laser mode.

Exercise 2.13 *Show that equations (2.2.33) and (2.2.32) are equivalent.*

2.3 Multimode Field States

The first step to defining multimode field state is to define a basis. It is most convenient to use the number states for each mode. For example, in a hypothetical 3-mode system a joint number state would be denoted

$$|\{n_\kappa\}\rangle = |n_1 n_2 n_3\rangle = |n_1\rangle_1 \otimes |n_2\rangle_2 \otimes |n_3\rangle_3, \tag{2.3.1}$$

and the infinite basis set would be

$$\begin{aligned} \{ \quad & |000\rangle, \\ & |100\rangle, |010\rangle, |001\rangle, \\ & |200\rangle, |020\rangle, |002\rangle, |110\rangle, \cdots \\ & \cdots \ \} \end{aligned} \tag{2.3.2}$$

For a full description of the electromagnetic field we have to consider an infinite number of modes. For that we have to consider states

$$|\{n_\kappa\}\rangle = \bigotimes_\kappa |n_\kappa\rangle_\kappa = |n_1\rangle_1 \otimes |n_2\rangle_2 \otimes |n_3\rangle_3 \cdots, \tag{2.3.3}$$

where here κ is the mode label which is meant to include $\mathbf{k}$ (the propagation vector), m (whether it is a sin or cos mode) and λ (labelling the polarization direction), and the product is an infinite product. In this case the basis set is doubly infinite in size. The first member of the basis set is the *vacuum state*

$$|\{0_\kappa\}\rangle = |000\cdots\rangle \tag{2.3.4}$$

A general pure state of the multimode field can be written

$$|\psi\rangle = \sum_{\{n_\kappa\}} c_{\{n_\kappa\}} |\{n_\kappa\}\rangle = \sum_{N=0}^{\infty} \sum_{\{n_\kappa : \sum_\kappa n_\kappa = N\}} c_{\{n_\kappa\}} |\{n_\kappa\}\rangle. \tag{2.3.5}$$

That is, we have to sum over all possible total photon numbers, and sum over all the possible ways each total photon number can be distributed between an infinite number of modes. A general mixed state can be written

$$\rho = \sum_{\{n_\kappa\},\{m_\kappa\}} \rho_{\{n_\kappa\},\{m_\kappa\}} |\{n_\kappa\}\rangle\langle\{n_\kappa\}|. \tag{2.3.6}$$

Fortunately the states we will consider are much less general than this. Specifically, they will all have the structure

$$\rho = \bigotimes_\kappa \rho_\kappa = \rho_1 \otimes \rho_2 \otimes \cdots \tag{2.3.7}$$

where

$$\rho_\kappa = \sum_n P_\kappa(n) |n\rangle_\kappa \langle n|_\kappa \tag{2.3.8}$$

That is, there are no correlations between the state of each mode, and each mode is in a mixture of number states.

2.3.1 Thermal Light

Since the total energy of the multimode system is the sum of the energy of the individual modes,

$$\hat{H} = \sum_{\mathbf{k},\lambda,t} \hbar\omega_{\mathbf{k}} \left(a^\lambda_{t,\mathbf{k}}\right)^\dagger a^\lambda_{t,\mathbf{k}} = \sum_\kappa \hbar\omega_\kappa a^\dagger_\kappa a_\kappa. \tag{2.3.9}$$

the eigenstates of $\hat{H}$ are just the number states $|\{n_\kappa\}\rangle$. In thermal equilibrium the probability of being in such a state factorizes:

$$P(\{n_\kappa\}) \propto \exp(-\sum_\kappa \hbar\omega_\kappa n_\kappa / k_B T) = \prod_\kappa \exp(-\hbar\omega_\kappa n_\kappa / k_B T) \tag{2.3.10}$$

Thus the total state matrix for multimode thermal light just factorizes into the single-mode thermal states:

$$\rho = \bigotimes_\kappa \rho^\kappa_{\rm th}\left(\bar{n}_{\rm th}(\omega_\kappa)\right) \tag{2.3.11}$$

where $\bar{n}_{\rm th}(\omega_\kappa)$ is defined in Eq. (2.2.27) and, as before,

$$\rho^\kappa_{\rm th}\left(\bar{n}_{\rm th}(\omega_\kappa)\right) = [1 - \exp(-\hbar\omega_\kappa / k_B T)] \sum_{n=0}^{\infty} \exp(-n\hbar\omega_\kappa / k_B T) |n\rangle_\kappa \langle n|_\kappa. \tag{2.3.12}$$

2.3.2 Collimated Polarized Thermal Light

Because all modes are populated, thermal light is isotropic. That is, it propagates in all directions equally. Also, it is unpolarized: for each $\mathbf{k}$ both polarizations λ are equally populated. One of the aims of this chapter is to compare thermal light with laser light, in order to understand what is special about the light produced by a laser. The first obvious property of laser light which contrasts with thermal light is that it is not isotropic. Rather, it propagates in a single direction, and it is usually polarized. In this section we consider thermal light which has been collimated (so as to make it unidirectional) and polarized.

In order to consider unidirectional light it is convenient to change modes from the standing waves used in Chapter 1 to traveling waves. In these new modes we can drop the t subscript and retain the $\mathbf{k}$ and λ labels. This halves the number of modes, but we double it by allowing $\mathbf{k}$ to take on all possible values, not just those with $k_x \geq 0$. The vector potential and electric fields for a free field are

$$\hat{\mathbf{A}}(\mathbf{x}) = \sum_{\mathbf{k},\lambda} \boldsymbol{\varepsilon}_{\mathbf{k}}^{\lambda} \sqrt{\frac{\hbar}{2\epsilon_0 \omega_{\mathbf{k}} L^3}} \left(a_{\mathbf{k}}^{\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} + (a_{\mathbf{k}}^{\lambda})^{\dagger} e^{-i\mathbf{k}\cdot\mathbf{x}} \right), \quad (2.3.13)$$

$$\hat{\mathbf{E}}^{\perp}(\mathbf{x}) = \sum_{\mathbf{k},\lambda} \boldsymbol{\varepsilon}_{\mathbf{k}}^{\lambda} \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2\epsilon_0 L^3}} \left(i a_{\mathbf{k}}^{\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} - i(a_{\mathbf{k}}^{\lambda})^{\dagger} e^{-i\mathbf{k}\cdot\mathbf{x}} \right), \quad (2.3.14)$$

where $\mathbf{k}$ can point in any direction:

$$\mathbf{k} = \frac{2\pi}{L}(l, m, n) \; : \; l, m, m \in \{-\infty \cdots, -2, -1, 0, 1, 2, \cdots \infty\}. \quad (2.3.15)$$

Exercise 2.14 *Determine the new mode operators* $\{a_{\mathbf{k}}^{\lambda}\}$ *in terms of the old ones* $\{a_{t,\mathbf{k}}^{\lambda}\}$.

Collimating the thermal light emitted by a blackbody source (by passing it through a series of finite apertures) will lead, in the ideal limit, to a field with $\mathbf{k}$ vectors in a single direction. Thus $\mathbf{k}$ can be replaced by $k = (2\pi/L)l$, with l a positive integer. Polarizing it will eliminate the need for the polarization vector and index λ. The energy of the field thus becomes

$$\hat{H} = \sum_{k} \hbar c k a_k^{\dagger} a_k, \quad (2.3.16)$$

and the state of the collimated polarized thermal field is

$$\rho = \bigotimes_{k} \rho_{\mathrm{th}}^{k} \left(\bar{n}_{\mathrm{th}}(ck) \right), \quad (2.3.17)$$

where the single-mode thermal state is as before. Here we have included only those modes propagating in the desired direction. The other modes would ideally be in the vacuum state.

Because the light is collimated, the energy density (which scales as $(k_B T)^4/(\hbar c)^3$ for an isotropic thermal field) is not a sensible quantity to

consider. Rather, we should calculate the energy per unit length in the direction of propagation. This yields

$$\frac{\left\langle \hat{H} \right\rangle}{L} = \frac{1}{L}\sum_k \hbar ck\bar{n}_{\rm th}(ck) \tag{2.3.18}$$

$$= \frac{\hbar c}{2\pi}\sum_k (\Delta k)k\bar{n}_{\rm th}(ck), \tag{2.3.19}$$

where $\Delta k = 2\pi/L$ is the separation of k modes. In the limit $L \to \infty$ this separation becomes infinitesimal and the sum can be converted to an integral, $\sum_k (\Delta k) \to \int_0^\infty dk$. The result is

$$\frac{\left\langle \hat{H} \right\rangle}{L} = \frac{\pi}{12}\frac{(k_B T)^2}{\hbar c}. \tag{2.3.20}$$

Exercise 2.15 *Show this. Hint:* $\int_0^\infty dx\, x/(e^x - 1) = \pi^2/6$.

Since the light is propagating in one direction, it is natural to covert this result to the energy per unit time, or power in the collimated beam:

$$P = \frac{\left\langle \hat{H} \right\rangle c}{L} = \frac{\pi}{12}\frac{(k_B T)^2}{\hbar}. \tag{2.3.21}$$

This can now be easily compared to the power of the output of a laser, which is typically of order one milliwatt.

Exercise 2.16 *Show that the temperature of the blackbody required to produce a collimated, polarized output of this power is about* 45000 *K. How does this compare to the temperature of the surface of the sun?*

2.3.3 Monochromatic Collimated Polarized Thermal Light

As well as being unidirectional and polarized, laser light is close to monochromatic. That is, almost all of the power is at approximately the same frequency. This property can be achieved from collimated polarized thermal light by passing it through a frequency filter (such as a series of Fabry-Perot etalons). To mimic the frequency spread of a typical laser, we can take the filter to change $\bar{n}_{\rm th}(\omega)$ to

$$\bar{n}_{\rm f}(\omega) = \bar{n}_{\rm th}(\omega)\frac{(\Gamma/2)^2}{(\Gamma/2)^2 + (\omega - \omega_0)^2} \simeq \bar{n}_{\rm th}(\omega_0)\frac{(\Gamma/2)^2}{(\Gamma/2)^2 + (\omega - \omega_0)^2} \tag{2.3.22}$$

Here ω_0 is the mean frequency and $\Gamma \ll \omega_0$ is the linewidth. (Note that this Γ is the same as the phase diffusion rate mentioned in Sec 2.2.3). The state of the filtered light is thus

$$\rho = \bigotimes_k \rho_{\rm th}^k\left(\bar{n}_{\rm f}(\omega)\right). \tag{2.3.23}$$

The power in this filtered collimated polarized thermal light is

$$\frac{\left\langle \hat{H} \right\rangle c}{L} = \frac{c}{L}\sum_k \hbar ck\bar{n}_{\rm f}(ck). \tag{2.3.24}$$

In the limit $\Gamma \ll \omega_0$ this can be evaluated as

$$P = \bar{n}_{\rm th}(\omega_0)\hbar\omega_0\Gamma/4 \tag{2.3.25}$$

Exercise 2.17 *Show this. Hint* $\int_{-\infty}^{\infty} dx/(1+x^2) = \pi$. *Justify all approximations made.*

For high temperatures ($k_B T \gg \hbar\omega_0$) this is simply

$$P = k_B T\Gamma/4 \tag{2.3.26}$$

This enables easy comparison with a typical laser which has an output power of order one milliwatt and a linewidth Γ of order 10^4 s^{-1}.

Exercise 2.18 *Show that the temperature of the blackbody required to produce a collimated, polarized, monochromatic output of this power and linewidth is about* 3×10^{16} *K. How does this compare to the temperature of the very early universe? Would it be feasible to filter this radiation with optical devices?*

2.3.4 Laser Light

We have seen that it would be completely impractical to try to produce a beam with the same power and linewidth of a laser beam by collimating, polarizing, and filtering a thermal beam. Even if it could be done, the result would still not be equivalent to a laser beam. The remaining difference is essentially the same difference between a single mode thermal state and a single mode laser state. The former has a very poorly defined photon number, while the latter can have a very well-defined photon number. In the multimode case, the filtered collimated polarized thermal beam would have large, slow (at rate Γ) intensity variations over time, whereas a laser beam has very small intensity variations. In the ideal limit where the state of the laser cavity mode is given by Eq. (2.2.33), the multimode state of the laser output beam is

$$\rho = \bigotimes_k \rho^k_{\rm laser}\left(\mu_{\rm f}(\omega)\right) \tag{2.3.27}$$

where

$$\mu_{\rm f}(\omega) = \mu \frac{(\Gamma/2)^2}{(\Gamma/2)^2 + (\omega-\omega_0)^2}, \tag{2.3.28}$$

where

$$\mu = \frac{P}{\hbar\omega_0}\frac{4}{\Gamma}, \tag{2.3.29}$$

which is the same as $\bar{n}_{\rm th}(\omega_0)$ needed to produce the same output power and linewidth from the filtered collimated polarized thermal beam.

Since P is the power and $\hbar\omega_0$ the energy per photon, $P/\hbar\omega_0$ is simply the *photon flux* of the laser, in photons per unit time. Also, since Γ is the phase diffusion rate, $4/\Gamma$ can be thought of as the time for the phase of the laser to become randomized. This time is also known as the *coherence time* of the laser. Thus μ measures the number of photons which come out of the laser which are coherent (that is, which have the

same phase in some sense). It is the fact that this number is very large in a typical laser which makes it difficult to produce a similar beam from a thermal source.

Exercise 2.19 *Calculate, to the nearest order of magnitude, the number of photons emitted from a laser per coherence time with the typical parameters used above, and for a typical optical frequency.*

3 Quantum Electrodynamics

3.1 The minimal coupling Hamiltonian

So far we have only considered the quantum theory of radiation in the absence of sources. To include classical sources we just have to add a new term to the Hamiltonian of the field. However in reality the sources are charged particles: electrons, protons *etc.*. These particles not only affect the field as sources, they are affected by it through the Lorentz force. If the field is quantized then the particles must be quantized as well if they are to be affected by the field. Of course we already know how to describe particles (at least nonrelativistically), using the free particle Hamiltonian operator. But now we need to add a *coupling* Hamiltonian which will produce both the source terms in the quantized Maxwell's equations, but also the Lorentz force for the particles.

The Hamiltonian which does both of these jobs is called the *minimal coupling* Hamiltonian. First consider the classical Hamiltonian for a *noninteracting* particle and field

$$H = H_0(\mathbf{v}) = \int \mathrm{d}^3\mathbf{x}\frac{\epsilon_0}{2}\left((\mathbf{E}^\perp)^2 + c^2(\boldsymbol{\nabla}\times\mathbf{A})^2\right) + \frac{1}{2m}(m\mathbf{v})^2 + V(\mathbf{r}), \tag{3.1.1}$$

where $\mathbf{v}$ is the velocity operator and $V(\mathbf{r})$ is a nonelectromagnetic potential (such as a gravitational potential). The procedure to introduce the minimal coupling is

$$H = H_0(\mathbf{v}) + q\phi(\mathbf{r}), \tag{3.1.2}$$

$$\mathbf{p} = \mathbf{p}_0(\mathbf{v}) + q\mathbf{A}(\mathbf{r}). \tag{3.1.3}$$

Here $\mathbf{p}_0(\mathbf{v}) = m\mathbf{v}$, and q is the charge of the particle. The concept of adding a scalar potential to the Hamiltonian is familiar, but what is probably less familiar is adding a vector potential to the momentum. In the presence of a vector potential the momentum $\mathbf{p}$ of the particle (which still becomes the operator $-i\hbar\boldsymbol{\nabla}$ when quantized) is not equal to $m\mathbf{v}$. However, the kinetic energy is still equal to $(m/2)\mathbf{v}^2$ as in H_0. Thus, from Eq. (6.1.1) and Eq. (6.1.2), the new Hamiltonian, expressed in terms of the momentum $\mathbf{p}$, is

$$H = \int \mathrm{d}^3\mathbf{x}\frac{\epsilon_0}{2}\left((\mathbf{E}^\perp)^2 + c^2(\boldsymbol{\nabla}\times\mathbf{A})^2\right) + \frac{1}{2m}(\mathbf{p}-q\mathbf{A})^2 + q\phi(\mathbf{r}) + V(\mathbf{r}). \tag{3.1.4}$$

The additional term, the coupling, can be split off as

$$H_\mathrm{I} = q\left[\frac{1}{2m}\left(-\mathbf{p}\cdot\mathbf{A} - \mathbf{A}\cdot\mathbf{p} + q\mathbf{A}^2\right) + \phi\right]. \tag{3.1.5}$$

Obviously only charged particles are coupled to the field.

3.2 * The Lorentz Force

The first task is to show that the coupling term does produce the Lorentz force on the charged particle. To show this the free Hamiltonian of the

field is irrelevant so we can work with

$$H = \frac{1}{2m}\left(\mathbf{p} - q\mathbf{A}(\mathbf{r})\right)^2 + q\phi(\mathbf{r}) + V(\mathbf{r}) \quad (3.2.1)$$

$$= \frac{1}{2m}\left(p_j - qA_j(\mathbf{r})\right)\left(p_j - qA_j(\mathbf{r})\right) + q\Phi(\mathbf{r}) + V(\mathbf{r}) \quad (3.2.2)$$

Here the Einstein summation convention for repeated indices is being used, and Φ is a classical scalar potential. We cannot consider the proper quantum expression for ϕ in terms of the charge distribution (1.3.12) because this would involve other particles and we wish to consider only a single particle for he moment.

First, we treat this as classical Hamiltonian. Then, ignoring $V(\mathbf{r})$, the equations of motion (Hamilton's equations) following from this are

$$\dot{r}_j = \frac{\partial H}{\partial p_j} = \frac{1}{m}\left(p_j - qA_j(\mathbf{r})\right) = v_j \,, \quad (3.2.3)$$

$$\dot{p}_j = -\frac{\partial H}{\partial r_j} = -q\frac{\partial \Phi}{\partial r_j} + \frac{q}{m}\left(p_l - qA_l(\mathbf{r})\right)\frac{\partial A_l}{\partial r_j} \,. \quad (3.2.4)$$

The first of these simply confirms the relation between the velocity and the momentum expressed in Eq. (6.1.2). The second can be rewritten

$$\dot{\mathbf{p}} = -q\boldsymbol{\nabla}\Phi + q\left(\boldsymbol{\nabla}\mathbf{A}\right)dt\mathbf{v}. \quad (3.2.5)$$

where the last term means $\mathbf{e}_k\left[\frac{\partial}{\partial r_k}A_j(\mathbf{r})\right]v_j$.

Now consider the total rate of change of the mechanical momentum

$$m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \dot{\mathbf{p}} - q\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}. \quad (3.2.6)$$

Here $\mathrm{d}f/\mathrm{d}t$ is the *total* derivative of a field f, as opposed to the partial derivative $\dot{f}$. For example, if f were a function $f(t,s)$ then we would have $\mathrm{d}f/\mathrm{d}t = \partial_t f + (\partial_t s)(\partial_s f)$. In the current example $\mathbf{A}$ is a function $\mathbf{A}(t,\mathbf{r})$ and $\partial_t r_j = v_j$. Using this, and Eq. (3.2.5), gives

$$m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \left[-q\boldsymbol{\nabla}\Phi + q\boldsymbol{\nabla}(\mathbf{A}\cdot\mathbf{v})\right] - q\left[\frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{A}\right]$$

$$= -q\left(\boldsymbol{\nabla}\Phi + \frac{\partial}{\partial t}\mathbf{A}\right) - q\left[(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{A} - \boldsymbol{\nabla}(\mathbf{v}\cdot\mathbf{A})\right], \quad (3.2.7)$$

Using the identity $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{b})\mathbf{c} - (\mathbf{a}\cdot\mathbf{c})\mathbf{b}$ and the relations

$$\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A} \,,\; \mathbf{E} = -\boldsymbol{\nabla}\Phi \; - \frac{\partial \mathbf{A}}{\partial t} \,, \quad (3.2.8)$$

we get finally

$$m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = q[\mathbf{E} + \mathbf{v}\times\mathbf{B}] \,. \quad (3.2.9)$$

This is the Lorentz force, which we know to be correct for a classical particle in an electromagnetic field.

Exactly the same result would be obtained if the Hamiltonian in Eq. (3.2.1) were treated as a quantum operator, and, instead of Hamilton's equations, Heisenberg's equations

$$\dot{\hat{\mathbf{r}}} = (i/\hbar)[\hat{H}, \hat{\mathbf{r}}] \tag{3.2.10}$$
$$\dot{\hat{\mathbf{p}}} = (i/\hbar)[\hat{H}, \hat{\mathbf{p}}] \tag{3.2.11}$$

were used. This completes the demonstration that the minimal coupling Hamiltonian correctly describes the effect of the field on a particle.

3.3 * Radiation Sources

The next step is to show that the minimal coupling Hamiltonian gives the correct source terms for the radiation field. For this we have to consider many particles indexed by a greek superscript, so the total Hamiltonian is

$$\begin{aligned}\hat{H} &= \int \mathrm{d}^3\mathbf{x}\frac{\epsilon_0}{2}\left((\hat{\mathbf{E}}^\perp)^2 + c^2(\boldsymbol{\nabla}\times\hat{\mathbf{A}})^2\right) \\ &\quad + \sum_\mu \left[\frac{1}{2m^\mu}\left(\hat{\mathbf{p}}^\mu - q^\mu\mathbf{A}(\hat{\mathbf{r}}^\mu)\right)^2 + \frac{1}{2}q^\mu\hat{\phi}(\hat{\mathbf{r}}^\mu) + V(\hat{\mathbf{r}}^\mu)\right]\end{aligned} \tag{3.3.1}$$

The position and momentum operators for the many particles obey

$$[\hat{p}^\mu_j, \hat{r}^\nu_l] = -i\hbar\delta_{jl}\delta^{\mu\nu} \tag{3.3.2}$$

Here $\hat{\phi}$ is expressed in terms of the particle positions:

$$\hat{\phi}(\mathbf{x}) = \frac{1}{4\pi\epsilon_0}\int \mathrm{d}^3\mathbf{r}\frac{\hat{\rho}(\mathbf{r})}{|\mathbf{r}-\mathbf{x}|} = \frac{1}{4\pi\epsilon_0}\sum_\nu \frac{q^\nu}{|\hat{\mathbf{r}}^\nu - \mathbf{x}|} \tag{3.3.3}$$

where the following

$$\hat{\rho}(\mathbf{r}) = \sum_\nu q^\nu\delta(\mathbf{r}-\hat{\mathbf{r}}^\nu) \tag{3.3.4}$$

has been used. Any classical (i.e. with a source external and irrelevant to the system) scalar potential can be included in $V(\mathbf{r})$.

3.3.1 Coulomb Energy

Before dealing with the radiation sources, we will show that the Coulomb energy

$$\hat{H}_{\text{Coulomb}} = \frac{1}{2}\sum_\mu q^\mu\hat{\phi}(\hat{\mathbf{r}}^\mu) \tag{3.3.5}$$

can be expressed in terms of the longitudinal electric field. There is a factor of 1/2 in this expression not included in the original minimal coupling Hamiltonian in order to avoid counting the Coulomb energy between particles twice. Substituting the above expression (3.3.3) gives

$$\hat{H}_{\text{Coulomb}} = \frac{1}{2}\sum_\mu q^\mu\frac{1}{4\pi\epsilon_0}\sum_\nu\frac{q^\nu}{|\hat{\mathbf{r}}^\nu - \hat{\mathbf{r}}^\mu|} \tag{3.3.6}$$
$$= \frac{1}{4\pi\epsilon_0}\sum_{\mu>\nu}\frac{q^\mu q^\nu}{|\hat{\mathbf{r}}^\nu - \hat{\mathbf{r}}^\mu|} \tag{3.3.7}$$

Strictly the infinite self-Coulomb energy when $\mu = \nu$, [which is in Eq. (3.3.6) but which has been omitted in Eq. (3.3.7)] needs to be included to calculate the exact relativistic quantum dynamics of the system, but it will not be important in our treatment.

Now the energy (3.3.5) can be re-expressed as

$$\hat{H}_{\rm Coulomb} = \frac{1}{2}\int {\rm d}^3\mathbf{x}\hat{\rho}(\mathbf{x})\hat{\phi}(\mathbf{x}) \tag{3.3.8}$$

From Eq. (1.3.11) this can be written as

$$\hat{H}_{\rm Coulomb} = -\frac{\epsilon_0}{2}\int {\rm d}^3\mathbf{x}\hat{\phi}(\mathbf{x})\nabla^2\hat{\phi}(\mathbf{x}) = -\frac{\epsilon_0}{2}\int {\rm d}^3\mathbf{x}\hat{\phi}(\mathbf{x})\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\hat{\phi}(\mathbf{x}) \tag{3.3.9}$$

Using integration by parts and assuming that $\hat{\phi}(\mathbf{x})$ vanishes at infinity as it will if our system is finite, we get

$$\hat{H}_{\rm Coulomb} = \frac{\epsilon_0}{2}\int {\rm d}^3\mathbf{x}\boldsymbol{\nabla}\hat{\phi}(\mathbf{x})\cdot\boldsymbol{\nabla}\hat{\phi}(\mathbf{x}) = \frac{\epsilon_0}{2}\int {\rm d}^3\mathbf{x}\left(\hat{\mathbf{E}}^{\parallel}(\mathbf{x})\right)^2 \tag{3.3.10}$$

Thus the Coulomb energy is exactly the non-radiative part of the electromagnetic field energy.

3.3.2 Radiation Sources

Ignoring the Coulomb energy and free Hamiltonian of the particles, the radiation field Hamiltonian is

$$\hat{H} = \hat{H}_{\rm free} + \hat{H}_{\rm source} \tag{3.3.11}$$

where

$$\hat{H}_{\rm free} = \int {\rm d}^3\mathbf{x}\frac{\epsilon_0}{2}\left((\hat{\mathbf{E}}^{\perp})^2 + c^2(\boldsymbol{\nabla}\times\hat{\mathbf{A}})^2\right) \tag{3.3.12}$$

$$\hat{H}_{\rm source} = \sum_\mu \frac{1}{2m^\mu}\left[-2q^\mu\hat{\mathbf{p}}^\mu\cdot\hat{\mathbf{A}}(\hat{\mathbf{r}}^\mu) + \left(q^\mu\hat{\mathbf{A}}(\hat{\mathbf{r}}^\mu)\right)^2\right] \tag{3.3.13}$$

Here the ordering of $\hat{\mathbf{p}}$ and $\hat{\mathbf{A}}$ does not matter, even in the quantum case where $\hat{\mathbf{p}}^\mu \sim -i\hbar\boldsymbol{\nabla}^\mu$, because $\boldsymbol{\nabla}\cdot\hat{\mathbf{A}} = 0$. Consider a single particle:

$$\langle\mathbf{r}|\hat{\mathbf{p}}\cdot\hat{\mathbf{A}}(\hat{\mathbf{r}})|\psi\rangle = \hat{p}_j\hat{A}_j(\hat{\mathbf{r}})|\psi\rangle = -i\hbar\frac{\partial}{\partial r_j}\left[\hat{A}_j(\mathbf{r})\langle\mathbf{r}|\psi\rangle\right] \tag{3.3.14}$$

$$= 0 - i\hbar\hat{A}_j(\mathbf{r})\frac{\partial}{\partial r_j}\langle\mathbf{r}|\psi\rangle = \langle\mathbf{r}|\hat{\mathbf{A}}_{\mathbf{j}}(\hat{\mathbf{r}})\hat{\mathbf{p}}_{\mathbf{j}}|\psi\rangle \tag{3.3.15}$$

$$= \langle\mathbf{r}|\hat{\mathbf{A}}(\hat{\mathbf{r}})\cdot\mathbf{p}|\psi\rangle. \tag{3.3.16}$$

We have shown in Sec. 1.4 that $\hat{H}_{\rm free}$ implies

$$\partial_t\partial_t\hat{\mathbf{A}} = c^2\nabla^2\hat{\mathbf{A}}, \tag{3.3.17}$$

which is the source-free Maxwell equation. To find the effect of the source we need to go back a step and find

$$\partial_t\hat{\mathbf{A}} = (i/\hbar)[\hat{H}_{\rm free} + \hat{H}_{\rm source}, \hat{\mathbf{A}}] \tag{3.3.18}$$

Now since $\hat{H}_{\rm source}$ is a function of $\hat{\mathbf{A}}$, which commutes with itself at different positions [see Eq. (1.4.13)], we have

$$\begin{aligned}
\partial_t \hat{\mathbf{A}}(\mathbf{r}) &= (i/\hbar)[\hat{H}_{\rm free}, \hat{\mathbf{A}}(\mathbf{r})] \\
&= (i\epsilon_0/2\hbar)\int \mathrm{d}^3\mathbf{x}[\left(\hat{\mathbf{E}}^\perp(\mathbf{x})\right)^2 + c^2\left(\boldsymbol{\nabla}\times\hat{\mathbf{A}}(\mathbf{x})\right)^2, \hat{\mathbf{A}}(\mathbf{r})] \\
&= (i\epsilon_0/2\hbar)\int \mathrm{d}^3\mathbf{x}[\left(\hat{\mathbf{E}}^\perp(\mathbf{x})\right)^2, \hat{\mathbf{A}}(\mathbf{r})] \\
&= (i\epsilon_0/2\hbar)\int \mathrm{d}^3\mathbf{x}\, 2\hat{E}_j^\perp(\mathbf{x})(i\hbar/\epsilon_0)\delta_{jl}^\perp(\mathbf{x}-\mathbf{r})\mathbf{e}_l, \qquad (3.3.19)
\end{aligned}$$

where both Eq. (1.4.13) and Eq. (1.4.16) have been used, and where $\mathbf{e}_l$ is the unit vector in the l direction. We have also used the handy operator identity (1.4.3).

Now because $\hat{\mathbf{E}}^\perp$ is a transverse vector field, integrating over all space with the transverse δ function is equivalent to integrating with a normal δ function (see Eq. (1.4.19)). The result is thus

$$\partial_t \hat{\mathbf{A}}(\mathbf{r}) = -\hat{\mathbf{E}}^\perp(\mathbf{r}), \qquad (3.3.20)$$

which is as it should be since this was the original definition of $\hat{\mathbf{E}}^\perp$! Now, we find

$$\begin{aligned}
\partial_t\partial_t\hat{\mathbf{A}} &= (i/\hbar)[\hat{H}_{\rm free}+\hat{H}_{\rm source}, \partial_t\mathbf{A}] \qquad (3.3.21)\\
&= (i/\hbar)[\hat{H}_{\rm free}+\hat{H}_{\rm source}, -\hat{\mathbf{E}}^\perp] \qquad (3.3.22)
\end{aligned}$$

Recall that $\mathbf{a}\times\mathbf{b} \equiv \epsilon_{ijk}\mathbf{e}_i a_j b_k$, where

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1 \qquad (3.3.23)$$
$$\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1 \qquad (3.3.24)$$
$$\text{all other } \epsilon_{ijk} = 0 \qquad (3.3.25)$$

Thus

$$(i/\hbar)[\hat{H}_{\rm free}, -\hat{\mathbf{E}}^\perp(\mathbf{r})] = (i\epsilon_0/2\hbar)\int \mathrm{d}^3\mathbf{x}[c^2\epsilon_{ijk}\epsilon_{lmk}\partial_i\hat{A}_j(\mathbf{x})\partial_l\hat{A}_m(\mathbf{x}), -\hat{E}_n^\perp(\mathbf{r})\mathbf{e}_n] \qquad (3.3.26)$$

Using the relation $\epsilon_{ijk}\epsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{lj}$ and the commutation relations, this can be evaluated as

$$(i\epsilon_0/2\hbar)2(i\hbar/\epsilon_0)c^2\int \mathrm{d}^3\mathbf{x}[\partial_i\delta_{jn}^\perp(\mathbf{x}-\mathbf{r})\mathbf{e}_n][\partial_i\hat{A}_j(\mathbf{x})] - [\partial_i\delta_{jn}^\perp(\mathbf{x}-\mathbf{r})\mathbf{e}_n][\partial_j\hat{A}_i(\mathbf{x})] \qquad (3.3.27)$$

Now integrating by parts and using the fact that $\hat{\mathbf{A}}$ is a transverse vector field gives

$$\begin{aligned}
(i/\hbar)^2[\hat{H}_{\rm free}, [\hat{H}_{\rm free}, \hat{\mathbf{A}}(\mathbf{r})]] &= c^2\left[\partial_i\partial_i\hat{A}_n(\mathbf{r})\mathbf{e}_n - \partial_i\partial_n\hat{A}_i(\mathbf{r})\mathbf{e}_n\right] \\
&= c^2\nabla^2\hat{\mathbf{A}}(\mathbf{r}), \qquad (3.3.28)
\end{aligned}$$

where $\boldsymbol{\nabla}\cdot\hat{\mathbf{A}} = 0$ has been used. This result is as expected, giving the source free wave equation (3.3.17) as derived earlier.

The final step is to calculate

$$\begin{aligned}[\hat{H}_{\rm source}, -\hat{\mathbf{E}}^{\perp}(\mathbf{x})] &= \sum_{\mu} \frac{1}{2m^{\mu}} \left[\{-2q^{\mu}\hat{\mathbf{p}}^{\mu}\cdot\hat{\mathbf{A}}(\hat{\mathbf{r}}^{\mu}) + \left(q^{\mu}\hat{\mathbf{A}}(\hat{\mathbf{r}}^{\mu})\right)^2\}, -\hat{\mathbf{E}}^{\perp}(\mathbf{x}) \right] \\ &= (i\hbar/\epsilon_0) \sum_{\mu} \frac{1}{2m^{\mu}} \left[-2q^{\mu}\hat{\mathbf{p}}^{\mu}\cdot\mathbf{e}_i\delta^{\perp}_{ij}(\mathbf{x}-\hat{\mathbf{r}}^{\mu})\mathbf{e}_j \right. \\ &\qquad \left. + 2(q^{\mu})^2\hat{\mathbf{A}}(\hat{\mathbf{r}}^{\mu})\cdot\mathbf{e}_i\delta^{\perp}_{ij}(\mathbf{x}-\hat{\mathbf{r}}^{\mu}) \right]. \end{aligned} \tag{3.3.29}$$

Now using

$$\hat{\mathbf{v}}^{\mu} = \left(\hat{\mathbf{p}}^{\mu} - q^{\mu}\hat{\mathbf{A}}(\hat{\mathbf{r}}^{\mu})\right)/m^{\mu} \tag{3.3.30}$$

we can rewrite this as

$$(i/\hbar)[\hat{H}_{\rm source}, -\hat{\mathbf{E}}^{\perp}] = \frac{1}{\epsilon_0}\sum_{\mu} \hat{\mathbf{v}}^{\mu}\cdot\mathbf{e}_i\delta^{\perp}_{ij}(\mathbf{x}-\hat{\mathbf{r}}^{\mu})\mathbf{e}_j q^{\mu} \tag{3.3.31}$$

Thus, combining the evolution from the free and source Hamiltonians we obtain

$$\partial_t\partial_t\hat{\mathbf{A}} = c^2\nabla^2\hat{\mathbf{A}} + \hat{\mathbf{j}}^{\perp}/\epsilon_0, \tag{3.3.32}$$

which is the wave equation with sources, as desired. In this equation,

$$\hat{\mathbf{j}}^{\perp}(\mathbf{x}) = \sum_{\mu} \hat{\mathbf{v}}^{\mu}\cdot\mathbf{e}_i\delta^{\perp}_{ij}(\mathbf{x}-\hat{\mathbf{r}}^{\mu})\mathbf{e}_j q^{\mu} \tag{3.3.33}$$

is the transverse part of the current. The operator ordering here is unimportant because $\mathbf{e}_i\delta^{\perp}_{ij}(\mathbf{x}-\hat{\mathbf{r}}^{\mu})\mathbf{e}_j$ is transverse and so commutes with $\hat{p}^{\mu} \sim -i\hbar\boldsymbol{\nabla}$. If we were to ignore the operator nature of $\hat{\mathbf{v}}$ we could write the total current as

$$\mathbf{j}(\mathbf{x}) = \sum_{\mu} \mathbf{v}^{\mu}\cdot\mathbf{e}_i\delta_{ij}(\mathbf{x}-\mathbf{r}^{\mu})\mathbf{e}_j q^{\mu} = \sum_{\mu} \mathbf{v}^{\mu}\delta(\mathbf{x}-\mathbf{r}^{\mu})q^{\mu}. \tag{3.3.34}$$

from which Eq. (3.3.33) is easy to understand as the transverse part.

3.4 Interpretation of the Minimal Coupling Hamiltonian

We restate the coupling term which arises from the minimal coupling method, once again for a single particle for simplicity.

$$\hat{H}_I = \frac{-q\hat{\mathbf{p}}\cdot\hat{\mathbf{A}}(\hat{\mathbf{r}})}{m} + \frac{q^2\mathbf{A}^2(\hat{\mathbf{r}})}{2m} + q\hat{\phi}(\hat{\mathbf{r}}). \tag{3.4.1}$$

Recall the expression for $\hat{\mathbf{A}}$ in terms of annihilation and creation operators (1.4.11) — it is of the form

$$\hat{\mathbf{A}}(\hat{\mathbf{r}}) = \sum_{\kappa} \mathbf{u}_{\kappa}(\hat{\mathbf{r}})(a_{\kappa} + a^{\dagger}_{\kappa}). \tag{3.4.2}$$

Thus the first term (linear in the charge) involves either the annihilation or the creation of a photon. Of course the charged particle (call it an

electron) is not destroyed, nor a new electron created, but its state may change as a result of the coupling. The second term (bilinear in the charge) involves the annihilation of two photons, or the creation of two photons, or the annihilation of one photon and the creation of another. Again the state of the electron may be changed, because of dependence of $\mathbf{A}$ on the position $\mathbf{r}$ of the particle. We will see in the next section that this term is a result of the non-relativistic approximation. The third term (linear in the charge) does not involve any photons, but represents the action of other charged particles on the electron under consideration through the Coulomb force.

4 Examples of QED processes

4.1 Radiative Transitions in Atoms

Atomic energy levels are not infinitely sharp, but have a finite width, called the *natural linewidth*. More precisely, the energy of the photon emitted by a spontaneous transition between two levels need not be exactly the energy difference of the levels. This can be thought of as a consequence of the energy-time uncertainty relation

$$\Delta E \Delta t \approx \tfrac{1}{2}\hbar \ : \tag{4.1.1}$$

the shorter-lived the atomic state, the less well-defined its energy and the greater the natural linewidth. If there was no uncertainty in the energy difference between the levels, the atom would not decay at all.

There is an uncertainty in the energy levels of an atom because the atomic Hamiltonian is not the full Hamiltonian of the system. Rather, as we have seen, there is a coupling between any charged particle and the electromagnetic field. Thus the explanation for radiative atomic transitions is to be sought in the minimal coupling Hamiltonian introduced in the preceding chapter. In this chapter we will see how the spontaneous emission rate between two levels of an atom can be calculated from first principles, by making lots of approximations.

4.1.1 The Non-Relativistic Approximation

Let us consider the situation in which the radiation field interacts with just electrons in a bound states in an atom. This applies to the case of a one-electron atom (or an effectively one-electron atom) in which $\hat{\phi}(\hat{\mathbf{r}})$ can be treated as a $c-$number $V(\hat{\mathbf{r}})$ due to the nucleus (or nucleus and inner core of electrons). As in all the theory we have considered so far, we use a *non-relativistic Hamiltonian*. This Hamiltonian, for the two systems of electron and radiation field, is

$$\hat{H} = \hat{H}_{\rm el} + \hat{H}_{\rm rad} + \hat{H}_{\rm I}. \tag{4.1.2}$$

The first term, $\hat{H}_{\rm el}$, is the non-relativistic Hamiltonian for the electron bound to the atom:

$$\hat{H}_{\rm el} = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{r}}) = \sum_j E_j |j\rangle\langle j|. \tag{4.1.3}$$

Here $E_j, |j\rangle$ are the eigenvalues and eigenstates of $\hat{H}_{\rm el}$, found by solving Schrödinger's equation.

The second term is the energy of the radiation field,

$$\hat{H}_{\rm rad} = \int {\rm d}^3\mathbf{x} \frac{\epsilon_0}{2}\left((\hat{\mathbf{E}}^\perp)^2 + c^2\hat{\mathbf{B}}^2\right) = \sum_{\mathbf{k},\lambda} \hbar\omega_{\mathbf{k}} \left(a^\lambda_{\mathbf{k}}\right)^\dagger a^\lambda_{\mathbf{k}}. \tag{4.1.4}$$

The third term is the interaction between these two systems, which we split into two parts:

$$\hat{H}_{\rm I} = \hat{H}_{\rm I,1} + \hat{H}_{\rm I,2} \tag{4.1.5}$$

where

$$\hat{H}_{\mathrm{I},1} = \left(\frac{e}{m}\right)\hat{\mathbf{A}}(\mathbf{r})\cdot\hat{\mathbf{p}}, \tag{4.1.6}$$

$$\hat{H}_{\mathrm{I},2} = \left(\frac{e^2}{2m}\right)\hat{A}^2(\hat{\mathbf{r}}). \tag{4.1.7}$$

These describe one and two-photon processes respectively, as noted in Sec. 3.4.

4.1.2 Single-photon Transition Approximation

Radiative transitions in an atom are dominated by $\hat{H}_{\mathrm{I},1}$, unless it is a forbidden transition with zero dipole moment. Quantitatively, the effects of $\hat{H}_{\mathrm{I},2}$ are typically smaller than those of $\hat{H}_{\mathrm{I},1}$ by a factor of

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \approx \frac{1}{137} \quad (\text{the } \textit{fine structure constant}). \tag{4.1.8}$$

4.1.3 Electric Dipole Approximation

Recall that the vector potential $\hat{\mathbf{A}}$ is defined in terms of annihilation and creation operators for photons of definite momenta in Eq. (2.3.13). We make the further approximation that

$$e^{i\mathbf{k}\cdot\hat{\mathbf{r}}} = e^{i\mathbf{k}\cdot\mathbf{r}_0}\,, \tag{4.1.9}$$

where $\mathbf{r}_0$ is the position of the centre of the atom. This is justified by the observation that in the optical region (which is the relevant region of the spectrum for most atomic transitions), the wavelength of the light is much greater than the linear dimensions of the atom. Typically, $r_{\mathrm{atom}} \approx 1\text{Å}$ while $\lambda \approx 10^3\text{Å}$). Under this *dipole approximation*, we can replace $\hat{\mathbf{A}}$ in Eq. (4.1.6) by

$$\hat{\mathbf{A}} = \sum_{\mathbf{k},\lambda} \boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}L^3}}\,\left(a^{\lambda}_{\mathbf{k}} + (a^{\lambda}_{\mathbf{k}})^{\dagger}\right), \tag{4.1.10}$$

where we have taken $\mathbf{r}_0 = \mathbf{0}$ without loss of generality.

Exercise 4.1 *Show, using only first year physics, that* $2\pi r_{\mathrm{Bohr}}/\lambda = O(\alpha)$*, where* r_{Bohr} *is the characteristic length for an atom.*

The electron operator $\hat{\mathbf{p}}$ can be written in the atomic state basis as

$$\hat{\mathbf{p}} = \sum_{jl} |j\rangle\langle j|\hat{\mathbf{p}}|l\rangle\langle l|. \tag{4.1.11}$$

Thus the interaction Hamiltonian can be taken to be

$$\hat{H}_{\mathrm{I},1} = \frac{e}{m}\sum_{\mathbf{k}\lambda jl}\left(\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}L^3}\right)^{\frac{1}{2}} \boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}}\cdot\langle j|\hat{\mathbf{p}}|l\rangle[\left(a^{\lambda}_{\mathbf{k}}\right) + (a^{\lambda}_{\mathbf{k}})^{\dagger}]|j\rangle\langle l|. \tag{4.1.12}$$

To continue we note that $\hat{\mathbf{p}} = im[\hat{H}_{\rm el}, \hat{\mathbf{r}}]/\hbar$, so that $\langle j|\hat{\mathbf{p}}|l\rangle$ can be written in terms of the dipole matrix element:

$$\langle j|\hat{\mathbf{p}}|l\rangle = \frac{im}{\hbar}\langle j|[\hat{H}_{\rm el}, \hat{\mathbf{r}}]|l\rangle = imE_{jl}\langle j|\hat{\mathbf{r}}|l\rangle/\hbar\,, \tag{4.1.13}$$

where $E_{jl} = E_j - E_l$. We then have

$$\hat{H}_{\rm I,1} = ie\sum_{\mathbf{k}\lambda jl}\left(\hbar 2\epsilon_0\omega_{\mathbf{k}}L^3\right)^{-1/2}\boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}}\cdot\langle j|\hat{\mathbf{r}}|l\rangle E_{jl}[\left(a^{\lambda}_{\mathbf{k}}\right) + (a^{\lambda}_{\mathbf{k}})^{\dagger}]|j\rangle\langle l|. \tag{4.1.14}$$

4.1.4 Rotating-Wave Approximation

We may simplify this further by noting that the energy non-conserving terms in eqn(4.1.14) will not contribute significantly to one-photon processes (although they may be important for multiphoton ones). Dropping these terms (the *rotating wave approximation* (RWA) or *secular approximation*) gives

$$\hat{H}_{\rm I,1} \approx \sum_{\mathbf{k},\lambda,E_j<E_l}\left(\hbar 2\epsilon_0\omega_{\mathbf{k}}L^3\right)^{-1/2}\boldsymbol{\varepsilon}^{\lambda}_{\mathbf{k}}\cdot\mathbf{d}_{jl}E_{jl}[i\left(a^{\lambda}_{\mathbf{k}}\right)^{\dagger}|j\rangle\langle l| - i(a^{\lambda}_{\mathbf{k}})|l\rangle\langle j|]\,, \tag{4.1.15}$$

where $\mathbf{d}_{jl} = e\langle j|\mathbf{r}|l\rangle$ is the *electric dipole moment* for the $l \to j$ transition. In wavefunction terms, this is

$$\mathbf{d}_{jl} = e\int d^3\mathbf{r}\phi^*_j(\mathbf{r})\mathbf{r}\phi_k(\mathbf{r}). \tag{4.1.16}$$

The RWA Hamiltonian means that either the atom can make a transition to a lower energy level by emitting a photon, or to a higher energy level by absorbing a photon.

4.1.5 Two-level Assumption

We assume that initially the field has no photons in it (the vaccum field), and that the atom is initially in an excited state $|e\rangle$. We also assume that there is only a single lower state, the ground state $|g\rangle$, to which it can make a transition by emitting a photon. We define the energy of the ground state to be 0 and the excited state energy to be $\hbar\omega_0$. The initially excited atom in a vacuum field will spontaneously return to the ground state. That is, there is a transition, due to the interaction Hamiltonian Eq. (4.1.15), from an initial state $|i\rangle = |e, \{0\}\rangle$ to a final state $|f^{\lambda}_{\mathbf{k}}\rangle = |g, 1^{\lambda}_{\mathbf{k}}\rangle$ in which a photon of momentum $\mathbf{k}$ and polarization λ has been emitted. The ket $|1^{\lambda}_{\mathbf{k}}\rangle$ is short for the ket $|00\cdots00100\cdots\rangle$ in the Fock basis, where the 1 appears as the photon number of the mode $\mathbf{k}, \lambda$.

Because the interaction Hamiltonian describes a one-photon process, the states $|i\rangle$ and $|f^{\lambda}_{\mathbf{k}}\rangle$ are the only states we need consider. Thus the state of the atom-field system at time t can be expanded as

$$|\psi(t)\rangle = c_0(t)\exp(-i\omega_0 t)|i\rangle + \sum_{\mathbf{k},\lambda}c^{\lambda}_{\mathbf{k}}(t)\exp(-i\omega_{\mathbf{k}}t)|f^{\lambda}_{\mathbf{k}}\rangle\,. \tag{4.1.17}$$

Now the combined state of the system obeys Schrödinger's equation which is, ignoring $\hat{H}_{\mathrm{I},2}$,

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \left(\hat{H}_{\mathrm{el}} + \hat{H}_{\mathrm{rad}} + \hat{H}_{\mathrm{I},1}\right)|\psi(t)\rangle \tag{4.1.18}$$

Subsitituting Eq. (4.1.17) into this equation, the left-hand-side evaluates to

$$i\hbar\left[(\dot{c}_0 - i\omega_0 c_0)\exp(-i\omega_0 t)|i\rangle + \sum_{\mathbf{k},\lambda}\left(\dot{c}^\lambda_{\mathbf{k}} - i\omega_{\mathbf{k}}c^\lambda_{\mathbf{k}}\right)\exp(-i\omega_{\mathbf{k}}t)|f^\lambda_{\mathbf{k}}\rangle\right], \tag{4.1.19}$$

while the three Hamiltonians on the right-hand-side give the three terms

$$\hbar\omega_0 c_0\exp(-i\omega_0 t)|i\rangle + \sum_{\mathbf{k},\lambda}\hbar\omega_{\mathbf{k}}c^\lambda_{\mathbf{k}}\exp(-i\omega_{\mathbf{k}}t)|f^\lambda_{\mathbf{k}}\rangle + \hat{H}_{\mathrm{I},1}|\psi(t)\rangle. \tag{4.1.20}$$

Cancelling common terms on both sides, and acting upon the remaining terms with $\langle f^\lambda_{\mathbf{k}}|$ and $\langle i|$ respectively gives

$$\exp(-i\omega_{\mathbf{k}}t)i\hbar\frac{\mathrm{d}c^\lambda_{\mathbf{k}}}{\mathrm{d}t} = \langle f^\lambda_{\mathbf{k}}|\hat{H}_{\mathrm{I},1}|\psi\rangle \tag{4.1.21}$$

$$\exp(-i\omega_0 t)i\hbar\frac{\mathrm{d}c_0}{\mathrm{d}t} = \langle i|\hat{H}_{\mathrm{I},1}|\psi\rangle \tag{4.1.22}$$

Noting that $\langle i|\hat{H}_{\mathrm{I},1}|i\rangle = \langle f^\lambda_{\mathbf{k}}|\hat{H}_{\mathrm{I},1}|f^{\lambda'}_{\mathbf{k}'}\rangle = 0$, and denoting the matrix element $\langle f^\lambda_{\mathbf{k}}|\hat{H}_{\mathrm{I},1}|i\rangle$ by $\kappa^\lambda_{\mathbf{k}}$ we thus obtain

$$\frac{\mathrm{d}c^\lambda_{\mathbf{k}}}{\mathrm{d}t} = (-i/\hbar)\kappa^\lambda_{\mathbf{k}}c_0(t)\exp(-i(\omega_0 - \omega_{\mathbf{k}})t) \tag{4.1.23}$$

$$\frac{\mathrm{d}c_0}{\mathrm{d}t} = (-i/\hbar)\sum_{\mathbf{k},\lambda}\kappa^{\lambda *}_{\mathbf{k}}c^\lambda_{\mathbf{k}}(t)\exp(i(\omega_0 - \omega_{\mathbf{k}})t)\,. \tag{4.1.24}$$

These equations are exact given the approximate Hamiltonian Eq. (4.1.15). If the RWA had not been made there would be terms oscillating as $\exp\left(i(\omega_0 + \omega_{\mathbf{k}})\right)$. These average to zero over any significant time scale because they are oscillating at twice the optical frequency. Hence the name (and justification) of the rotating-wave approximation.

Eq. 4.1.23 can be formally integrated to give

$$c^\lambda_{\mathbf{k}}(t) = \frac{-i}{\hbar}\kappa_{\mathbf{k}}\int_0^t c_0(t')\exp(-i(\omega_0 - \omega_{\mathbf{k}})t')\mathrm{d}t'\,. \tag{4.1.25}$$

Substituting this into Eq. (4.1.24) gives

$$\frac{\mathrm{d}c_0}{\mathrm{d}t} = -\frac{1}{\hbar^2}\sum_{\mathbf{k},\lambda}|\kappa^\lambda_{\mathbf{k}}|^2\int_0^t c_0(t')\exp(-i(\omega_0 - \omega_{\mathbf{k}})(t'-t))\mathrm{d}t'\,. \tag{4.1.26}$$

4.1.6 Continuum limit

To proceed further, note that as the normalization volume L^3 becomes infinite, the final states $|f_{\mathbf{k}}^{\lambda}\rangle$ become infinitely closely spaced in energy. Also note that the integral in Eq. (4.1.26) depends only on the frequency of the emitted photon, not its direction or polarization. Thus we can replace the sum over $\mathbf{k}$ vectors by an integral over frequencies, with a mode density $\rho(\omega_f)\mathrm{d}\omega_f$,

$$\frac{\mathrm{d}c_0}{\mathrm{d}t} = -\frac{2}{\hbar^2}\int_0^\infty |\kappa_f|^2 \rho(\omega_f) \int_0^t c_0(t') e^{-i(\omega_0-\omega_f)(t'-t)} \mathrm{d}t' \mathrm{d}\omega_f \,. \qquad (4.1.27)$$

4.1.7 Markovian Approximation

A systematic solution to this can be developed using Laplace transform techniques. However, a simple approximation allows us to get a direct estimate of the spontaneous emission rate and hence the linewidth. We note that the exponential in the integrand of

$$f(t,\omega_f) \equiv \int_0^t c_0(t') e^{-i(\omega_0-\omega_f)(t'-t)} \mathrm{d}t' \qquad (4.1.28)$$

is rapidly oscillating (as a function of t') except for $\omega_f \approx \omega_0$. Thus $f(t,\omega_f)$ will average to zero except in the region $\omega_f \approx \omega_0$. This is physically reasonable too: it says that the frequency of the emitted photon is close to the transition frequency of the atom. Thus we can replace the coupling constant and mode density by their values at the resonant frequency, and extend the lower limit of the integral to $-\infty$ without greatly changing the value of the integral. This is valid provded that $\rho(\omega_f)$ is a slowly varying function of ω_f in the region of ω_0. This gives

$$\frac{\mathrm{d}c_0}{\mathrm{d}t} \approx -\frac{2}{\hbar^2}|\kappa_0|^2 \rho(\omega_0) \int_{-\infty}^\infty \mathrm{d}\omega_f \int_0^t \mathrm{d}t'\, c_0(t') e^{-i(\omega_0-\omega_f)(t'-t)} \,. \qquad (4.1.29)$$

Now we can change the order of the integrals to get

$$\frac{\mathrm{d}c_0}{\mathrm{d}t} \approx -\frac{2}{\hbar^2}|\kappa_0|^2 \rho(\omega_0) \int_0^t \mathrm{d}t'\, c_0(t') \int_{-\infty}^\infty \mathrm{d}\omega_f e^{-i(\omega_0-\omega_f)(t'-t)} \,. \qquad (4.1.30)$$

But the second integral, over ω_f, is the standard complex Fourier representation of the Dirac delta function:

$$2\pi\delta(t) = \int_{-\infty}^\infty e^{-i\omega t} \mathrm{d}\omega \,, \qquad (4.1.31)$$

making the integral over t' trivial. The result is

$$\frac{\mathrm{d}c_0}{\mathrm{d}t} \approx -\frac{\gamma}{2} c_0(t) \,, \qquad (4.1.32)$$

where

$$\gamma = \frac{2\pi|\kappa_0|^2\rho(\omega_0)}{\hbar^2} \qquad (4.1.33)$$

is the spontaneous decay rate.

4.1.8 Evaluating the spontaneous decay rate

To fully evaluate this, we need the mode density $\rho(\omega_0)$ and the coupling coefficient κ_0. These can be found as follows.

$$\rho(\omega_0) = \frac{dN(\omega)}{d\omega}|_{\omega=\omega_0}, \tag{4.1.34}$$

where $N(\omega)$ is the number of modes with frequency less than ω. Recall that with a normalization volume of L^3, $\mathbf{k} = (2\pi/L)(l,m,n) = (2\pi/L)(\nu_1,\nu_2,\nu_3)$. The number of modes is thus

$$\begin{aligned} N(\omega) &= \sum_{(\nu_1,\nu_2,\nu_3) \text{ such that } (2\pi c/L)^2(\nu_1^2+\nu_2^2+\nu_3^2)<\omega^2} \sum_{\lambda=1,2} 1 \\ &\approx 2\int_0^{L\omega/2\pi c} 4\pi\nu^2 d\nu = \frac{L^3\omega^3}{3\pi^2c^3} \end{aligned} \tag{4.1.35}$$

Thus

$$\rho(\omega_0) = \frac{L^3\omega_0^2}{\pi^2c^3}. \tag{4.1.36}$$

Next,

$$\kappa_0 = \langle f_{\mathbf{k}}^{\lambda}|\hat{H}_{\mathrm{I}}|i\rangle = i\left(\hbar 2\epsilon_0\omega_0 L^3\right)^{-1/2} E_0 \boldsymbol{\varepsilon}_{\mathbf{k}}^{\lambda}\cdot\mathbf{d}. \tag{4.1.37}$$

Here $E_0 = \hbar\omega_0 = E_{+-}$, and $\mathbf{d} = \mathbf{d}_{+-}$. Thus we have

$$|\kappa_0|^2 = \hbar\omega_0(2\epsilon_0 L^3)^{-1} d^2\cos^2\theta, \tag{4.1.38}$$

where θ is the angle between the polarization vector of the mode and the dipole vector $\mathbf{d} = d\hat{\mathbf{d}}$ of the atom. Since we don't care about the polarization of the output, we want to average over all possible polarization directions to find $\overline{\cos^2\theta}$. This can be evaluated using the Euler angles

$$\overline{\cos^2\theta} = \frac{\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta d\theta \cos^2\theta}{\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta d\theta} = \tfrac{1}{2}\int_{-1}^{1} dx x^2 = \frac{1}{3}. \tag{4.1.39}$$

The final expression for the total spontaneous emission rate is thus

$$\gamma = \frac{\omega_0^3 d^2}{3\pi\epsilon_0 c^3\hbar}. \tag{4.1.40}$$

Note that the normalization volume L^3 has disappeared from this expression, which is as required for physical consistency.

4.1.9 Fermi's Golden Rule

For a typical atom, γ is of the order of $10^8 s^{-1}$ or smaller, which is much less than the transition frequency which is of order $10^{15}s^{-1}$. This difference in time scales is what justifies the rotating wave approximation. The ratio can be expressed as follows:

$$\frac{\gamma}{\omega_0} = \frac{4}{3}\alpha\left(\frac{|\langle +|\mathbf{r}|-\rangle|}{\lambda/2\pi}\right)^2. \tag{4.1.41}$$

Now the fine structure constant α (see Eq. (4.1.8)) is small, and was used to ignore $\hat{H}_{\mathrm{I},2}$. Also, we have used the fact that the wavelength λ is much greater than the displacement of the electron $|\langle +|\mathbf{r}|-\rangle|$ in making the dipole approximation. In fact, using the exercise in Sec. 4.1.3, we have $\gamma/\omega_0 \sim \alpha^2$. Thus all of the approximations are justified by the smallness of α. Indeed, this is why QED is such a good theory for calculating with (compared to other quantum field theories we will meet in Chapter 6).

This result for spontaneous emission is one application of a general theory of irreversible processes in QM. The irreversibility arises from the fact that there are many final states into which the system may evolve, so the chances of it finding its way back into its initial state is negligible. By taking the continuum limit (by using a density of states $\rho(\omega)$), true irreversibility is obtained. It applies to any scattering problem in which the particles (or quanta as in this case) are scattered into an infinite region of free space. The general result, known as Fermi's golden rule, is that the rate of the irreversible process is

$$\gamma_{i\to f} = 2\pi |V_{fi}|^2 \rho(E_f)/\hbar, \tag{4.1.42}$$

where V_{fi} is the amplitude of the interaction Hamiltonian matrix element taking the initial state i to final state f, and $\rho(E_f)$ is the density of final states as a function of of energy.

4.1.10 The Lamb shift and Renormalization

A more careful calculation of the radiative process in atoms agrees with this result for the emission rate, but shows that there is in addition a frequency shift $\Delta\omega$, the *Lamb shift*, due to emission and reabsorption of virtual photons, so that

$$\frac{\mathrm{d}c_0}{\mathrm{d}t} = -\left(\frac{\gamma}{2} + i\Delta\omega\right) c_0(t)\,, \tag{4.1.43}$$

and hence

$$c_0(t) = \exp\left[-\left(\frac{\gamma}{2} + i\Delta\omega\right) t\right]\,. \tag{4.1.44}$$

A naive calculation of $\Delta\omega$ returns the value infinity; to obtain a finite answer requires relativistic QED and the theory of *renormalization*.

Basically, the naive calculation gives the shift in the energy of an atomic level due to coupling to the infinite number of electromagnetic modes in the vacuum state. But these infinite number of modes in fact give an infinite energy shift to a *free* electron — it can be thought of as giving it an infinite mass (since $E_0 = mc^2$). Since we know electrons don't have an infinite mass, we assume the electron actually has an infinite *negative* mass to begin with, such that this infinite positive mass contribution from the electromagnetic field is cancelled, leaving over just the usual finite mass of the electron. (There are an awful lot of mathematical tricks in this as you might imagine). This is called *mass-renormalization*. Now this infinite negative mass also cancels the infinity in the naive calculation of the energy of the electron in the atom, but what is left over is not only the usual electron mass term, but also an

additional small energy shift, which is the Lamb shift. The calculation of the Lamb shift by Bethe in 1947 was the first triumph of renormalization theory.

4.1.11 The Spectrum

To find the linewidth, we just substitute the solution eqn(4.1.44) back into eqn(4.1.25), to get

$$c_{\mathbf{k}}^{\lambda}(t) = \frac{\kappa_{\mathbf{k}}^{\lambda}/\hbar}{\omega_0 - \omega_{\mathbf{k}} + \Delta\omega - i\frac{\gamma}{2}} \left[\exp(-i(\omega_0 - \omega_{\mathbf{k}} + \Delta\omega)t - \frac{\gamma}{2}t) - 1\right] . \tag{4.1.45}$$

The emission spectrum $S(\mathbf{k})$ can be defined as the probability for the photon to be emitted with propagation vector $\mathbf{k}$. This includes both direction and frequency information. It is given by

$$\begin{aligned} S(\mathbf{k},\lambda) &= \lim_{t\to\infty} |c_{\mathbf{k}}^{\lambda}(t)|^2 = \frac{|\kappa_0|^2/\hbar^2}{(\omega_0 - \omega_{\mathbf{k}} + \Delta\omega)^2 + (\gamma/2)^2} \\ &= \frac{\omega_0 (2\epsilon_0 L^3 \hbar)^{-1} d^2 \cos^2\theta}{(\omega_0 - \omega_{\mathbf{k}} + \Delta\omega)^2 + (\gamma/2)^2} \end{aligned} \tag{4.1.46}$$

That is, the spectrum is a Lorentzian, peaked at $\omega_0 + \Delta\omega$, with linewidth (FWHMH) γ. This is exactly the linewidth expected from the simple uncertainty principle argument outlined at the beginning of this section, taking Δt to be the lifetime $1/\gamma$. The fact that $\gamma \ll \omega_0$ (as found above) then justifies the use of the RWA and dipole approximations, since the spectrum derived here indicates that only photons with frequencies within a few linewidths of ω_0 are populated.

Note that the dipole radiation distribution, proportional to $\cos^2\theta$, is incorporated, following from Eq. (4.1.38). Recall that θ is the angle between the atomic dipole and the photon polarization vector. If the dipole moment is defined as being in the z-direction then a polarization vector in the θ, ϕ direction implies that the photon propagation vector $\mathbf{k}$ is in the θ', ϕ direction, where $\theta' = \theta + \pi/2$. Thus the dipole radiation pattern can be rewritten in the more familiar form of $\sin^2\theta'$.

It might be thought strange that the normalization volume L^3 appears on the RHS of Eq. (4.1.46), but that is because we have assumed that the photon propagation vectors $\mathbf{k}$ can be differentiated no matter how small the difference between them. More realistically we can consider the probability for the photon to be emitted in a small solid angle $\sin\theta' \mathrm{d}\theta' \mathrm{d}\phi$ and small frequency range $\mathrm{d}\omega$, which equals

$$S(\omega,\theta',\phi)\rho(\omega_0)\mathrm{d}\omega \frac{\sin\theta'\mathrm{d}\theta'\mathrm{d}\phi}{4\pi} = \frac{\gamma}{2\pi} \frac{3\sin^2\theta'}{(\omega_0 - \omega + \Delta\omega)^2 + (\gamma/2)^2} \mathrm{d}\omega \frac{\sin\theta'\mathrm{d}\theta'\mathrm{d}\phi}{4\pi} \tag{4.1.47}$$

Exercise 4.2 *Show this and verify that it integrates to unity.*

4.2 Thomson scattering

In Thomson scattering a photon with propagation vector $\mathbf{k}_i$ and polarization $\boldsymbol{\varepsilon}_{\mathbf{k}_i}^{\lambda_i}$ is scattered by a free electron with initial momentum $\mathbf{p}_i$.

After the scattering process the electron has momentum $\mathbf{p}_f$ while the scattered photon has propagation vector $\mathbf{k}_f$ and polarization $\boldsymbol{\varepsilon}_{\mathbf{k}_f}^{\lambda_f}$. For convenience we can always transform to a frame where the initial momentum of the electron is $\mathbf{p}_i = \mathbf{0}$. In that frame, the nonrelativistic QED Hamiltonian can be used as long as $\hbar\omega_i \ll mc^2$, where m is the rest mass of the electron. We will assume this to be the case.

Exercise 4.3 *Show classically from conservation of free energy* ($H_{\rm el} + H_{\rm rad} = (\mathbf{p})^2/2m + \hbar c|\mathbf{k}|$) *and momentum* ($\mathbf{p} + \hbar\mathbf{k}$) *that if* $\mathbf{p}_i = \mathbf{0}$ *then*

$$\frac{\omega_i - \omega_f}{\omega_i} < \frac{2\hbar\omega_i}{mc^2} \ll 1. \tag{4.2.1}$$

This result shows that the change in the energy of the photon is negligible in this scattering process. From this it follows that we can assume $\omega_i \approx \omega_f$.

Exercise 4.4 *Show that* $\hat{H}_{\rm I,1}$ *does not contribute to the scattering process according to Fermi's golden rule (which is really a first-order perturbation result).*

Exercise 4.5 *Show that from* $\hat{H}_{\rm I,2}$ *the amplitude* V_{fi} *is given by*

$$V_{fi} = \langle \mathbf{p}_f, 1_{\mathbf{k}_f}^{\lambda_f} | \hat{H}_{\rm I,2} | \mathbf{p}_i, 1_{\mathbf{k}_i}^{\lambda_i} \rangle \tag{4.2.2}$$

$$= \frac{e^2\hbar}{2m\epsilon_0 L^3} (\boldsymbol{\varepsilon}_{\mathbf{k}_i}^{\lambda_i} \cdot \boldsymbol{\varepsilon}_{\mathbf{k}_f}^{\lambda_f}) \langle \mathbf{p}_f + \hbar\mathbf{k}_f | \mathbf{p}_i + \hbar\mathbf{k}_i \rangle (\omega_i\omega_f)^{-1/2}, \tag{4.2.3}$$

where (for $\alpha = i, f$*)* $\omega_\alpha = c|\mathbf{k}_\alpha|$ *and the states* $|\mathbf{p}_\alpha + \hbar\mathbf{k}_\alpha\rangle$ *are momentum eigenstates of the electron. Hint: show first that* $e^{i\mathbf{k}\cdot\hat{\mathbf{r}}}|\mathbf{p}\rangle = |\mathbf{p} + \hbar\mathbf{k}\rangle$ *by considering the momentum eigenstate in the position representation.*

Exercise 4.6 *Show using Fermi's golden rule that the probability of scattering into the final state per unit time is*

$$\gamma_{i\to f} = 2\pi r_0^2 \delta_{\mathbf{p}_f + \hbar\mathbf{k}_f, \mathbf{p}_i + \hbar\mathbf{k}_i} \frac{c}{L^3} 4 (\boldsymbol{\varepsilon}_{\mathbf{k}_i}^{\lambda_i} \cdot \boldsymbol{\varepsilon}_{\mathbf{k}_f}^{\lambda_f})^2, \tag{4.2.4}$$

where

$$r_0 = \frac{e^2}{4\pi\epsilon_0 mc^2} \tag{4.2.5}$$

is known as the "classical electron radius". Hint: $\langle \mathbf{p}|\mathbf{p}'\rangle = \delta_{\mathbf{p},\mathbf{p}'}$.

In Eq. (4.2.4), the rate $\gamma_{i\to f}$ still contains the normalization volume L^3, and in the limit $L^3 \to \infty$ the rate goes to zero. That is because our initial state contained one photon and in the limit $L^3 \to \infty$ that photon (which is spread over all space) has a zero probability of hitting the electron. To obtain a quantity with a physical interpretation it is necessary to divide $\gamma_{i\to f}$ by the beam intensity which this single photon represents. Since the speed of the photon is c, and the length of the volume L, the rate of photons in the beam is c/L photons per second. Since the cross section of the beam is L^2, the beam flux is c/L^3 photons per second per square meter.

Exercise 4.7 *Show that the rate of scattering, normalized relative to the beam intensity, and summed over all possible final electron momenta* $\mathbf{p}_f$ *and photon polarizations* $\boldsymbol{\varepsilon}_{\mathbf{k}_f}^{\lambda_f}$ *is*

$$\sigma = \frac{8\pi}{3} r_0^2. \tag{4.2.6}$$

Evaluate this quantity. How can it be interpreted, and why?

5 Relativistic Quantum Mechanics

5.1 Spin

5.1.1 Rotations

Spin was a surprising feature of quantum mechanics, with no classical analogue. However, it is intimately related to rotations, so we begin by reviewing rotations. These form a *group* of transformations denoted $SO(3)$, whose elements are 3×3 matrices R which are special ($\det(R) = 1$) and orthogonal ($RR^T = 1$). In index notation, the latter equation is

$$R^i{}_j R_k{}^j = \delta^i_k \tag{5.1.1}$$

Here we are using the Einstein summation convention — if an an index appears twice, once "upstairs" and once "downstairs", it is summed over. Under a rotation of the frame of reference, a vector transforms as

$$r^i \to r^{i'} = R^{i'}_j \, r^j. \tag{5.1.2}$$

Rotations preserve the scalar product between two vectors $\mathbf{q}$ and $\mathbf{r}$:

$$\mathbf{q}\cdot\mathbf{r} = \delta_{ij} q^i r^j. \tag{5.1.3}$$

Exercise 5.1 *Using Eq. (5.1.1), show that the scalar product (5.1.3) is invariant under a rotation of axes.*

This invariance is what defines the structure of Euclidean space, and the matrix δ_{ij} is known as the metric tensor for Euclidean space.

An arbitrary 3×3 matrix has 9 elements. The restriction of Eq. (5.1.1) gives six equations these must satisfy (only six because RR^T is automatically a symmetric matrix). Moreover, these restrictions imply that $\det(R) = \pm 1$, so that the requirement that $\det(R) = 1$ is not an independent restriction (it simply rules out spatial reflection). Thus R can be specified by three real numbers, which can be represented by a vector $\mathbf{w}$ known as the **twist vector**, where the rotation is by an angle $\theta = |\mathbf{w}|$ around the axis $\check{\mathbf{w}}$. (Here I use a $\check{}$ to represent a unit vector, rather than the usual $\hat{}$, to avoid confusion with the notation for an operator.) For example, if $\mathbf{w} = \theta \mathbf{e}_3$ (a rotation around the z-axis), then

$$R = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.1.4}$$

5.1.2 Spinors

The most important case of spin is $s = \frac{1}{2}$, for two reasons. First, particles of all other spins (including zero spin) can be made from a compound of spin half particles. Second, as we will see, all fundamental *particles* are spin half. (Here I am maintaining the distinction between particles such as electrons and quanta such as photons.)

A spin-half particle is described by a two-dimensional Hilbert space, and so spin is represented by 2×2 matrices. Because they are so important, they are given a special name, the Pauli matrices. Specifically,

$$\mathbf{S} = \tfrac{1}{2}\hbar\boldsymbol{\sigma}, \tag{5.1.5}$$

where the three Pauli matrices are[2]

$$\boldsymbol{\sigma} = \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right)^T . \tag{5.1.6}$$

The Pauli matrices obey the following relations

$$\sigma_i \sigma_j = i\epsilon_{ijk}\sigma_k + \delta_{ij} \; ; \tag{5.1.7}$$

From this it may be shown that, for any vectors $\mathbf{A}$ and $\mathbf{B}$,

$$(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B}) \, , \tag{5.1.8}$$

$$\boldsymbol{\sigma}(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{B} - i(\boldsymbol{\sigma}\times\mathbf{B}) \, . \tag{5.1.9}$$

Exercise 5.2 *Check (5.1.7) and from this prove the above properties of the Pauli matrices.*

A spin half particle has only two orthogonal spin states. Ignoring the spatial dependence of the wavefunction for the present, we can write the system state in terms of the eigenstates of σ_3:

$$|\xi\rangle = \xi_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \xi_{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_{-1} \end{pmatrix} . \tag{5.1.10}$$

In the relativistic theory there is not as simple a relation between the components of the wavefunction and the value of the spin. For this reason, it is simpler now to label the states instead as

$$|\xi\rangle = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} . \tag{5.1.11}$$

If we wish to describe the system relative to a different set of coordinates, rotated relative to the original axes by twist vector $\mathbf{w}$, it can be shown that the state transforms to

$$|\xi'\rangle = S_{\mathbf{w}}|\xi\rangle \equiv \exp(-i\tfrac{1}{2}\mathbf{w}\cdot\boldsymbol{\sigma})|\xi\rangle. \tag{5.1.12}$$

Writing $\mathbf{w} = \theta\check{\mathbf{w}}$ we find

$$S_{\mathbf{w}} = \cos\frac{\theta}{2} - i\check{\mathbf{w}}\cdot\boldsymbol{\sigma}\sin\frac{\theta}{2}. \tag{5.1.13}$$

[2]Note that I will generally not use the $\hat{}$ symbol on Pauli operators and certain related operators that we will see in the relativistic setting. This is for three reasons. First, they are usually represented as matrices, so it should be easy to remember that they do not commute. Second, as well as representing quantum quantities, these matrices also have other applications, as we will see. Third, the quantum quantities they represent have no close classical analogues, so no confusion is likely.

Exercise 5.3 *Verify this and show that this rotation preserves the norm of the state (that is, preserves $|\xi_1|^2 + |\xi_2|^2$).*
Hint: *First show that $\sigma^2 = 1$, where $\sigma = \check{\mathbf{w}}\cdot\boldsymbol{\sigma}$, then expand the exponential in an infinite Taylor series.*

The group of transformations of the form (5.1.13) is called the group $SU(2)$. Here 2 indicates that the elements operate in a two dimensional linear space, U that they are unitary ($S_{\mathbf{w}}^\dagger S_{\mathbf{w}} = I$), and S that they are special ($\det(S_{\mathbf{w}}) = 1$)[3] The two-component state on which they act is called a *spinor*. A spinor is an object distinct from scalars, vectors, tensors *etc.*, because of its manner of transformation (5.1.12). Spinors are fundamental to the study of elementary particles. It is obvious from Eq. (5.1.13) that a rotation of $\theta = 2\pi$ yields a phase factor of -1. More generally, a rotation of $\theta + 2\pi$ is identical with a rotation of θ and a sign change. That is, there is a two-to-one mapping from $SU(2)$ matrices to rotations in three dimensions.

5.1.3 Spinors and Vectors

The mapping between the $SU(2)$ group and the group of rotations can be made explicit. A given spinor $|\xi\rangle$ has a corresponding Hermitian 2×2 matrix of zero trace

$$H = 2|\xi\rangle\langle\xi| - \langle\xi|\xi\rangle = \begin{pmatrix} |\xi_1|^2 - |\xi_2|^2 & 2\xi_1\xi_2^* \\ 2\xi_2\xi_1^* & |\xi_2|^2 - |\xi_1|^2 \end{pmatrix}. \tag{5.1.14}$$

A given vector $\mathbf{h}$ can also be represented as a Hermitian 2×2 matrix of zero trace

$$\underline{\mathbf{h}} = \mathbf{h}\cdot\boldsymbol{\sigma} = \begin{pmatrix} h_3 & h_1 - ih_2 \\ h_1 + ih_2 & -h_3 \end{pmatrix}. \tag{5.1.15}$$

The normalized spinor $|\xi\rangle$ may thus be associated with a vector $\mathbf{h}$ of unit norm by setting

$$\underline{\mathbf{h}} = H \tag{5.1.16}$$

Such a vector $\mathbf{h}$ is a useful representation for any two-level quantum system, and is sometimes called the Bloch vector.

Now the identification (5.1.16) by itself is of no great significance. The crucial point is that the matrices H and $\underline{\mathbf{h}}$ transform identically under rotations. First note that $|\xi'\rangle = S_{\mathbf{w}}|\xi\rangle$ implies that under the rotation by the twist vector $\mathbf{w}$,

$$H' = S_{\mathbf{w}} H S_{\mathbf{w}}^\dagger. \tag{5.1.17}$$

Now it can be shown, by working out the corresponding rotation matrix $R_{\mathbf{w}}$, that H undergoes the same transformation:

$$\underline{\mathbf{h}}' = \mathbf{h}'\cdot\boldsymbol{\sigma} = (R_{\mathbf{w}}\mathbf{h})\cdot\boldsymbol{\sigma} = S_{\mathbf{w}}\underline{\mathbf{h}}S_{\mathbf{w}}^\dagger. \tag{5.1.18}$$

From these equations, it is evident that spinors are something like the square root of a vector.

[3]This last fact follows from the fact that the generators $\boldsymbol{\sigma}$ are traceless.

Exercise 5.4 *Verify Eq. (5.1.18) for a rotation around the* $\mathbf{e}_3$ *axis.*

The representation of 3−vectors as 2×2 Hermitian matrices is in one sense older than their representation as an ordered triplet of real numbers. Before Gibbs invented the modern notation for vectors, they were represented by Hamilton (of Hamiltonian fame) in terms of non-commuting elements $\mathbf{i}, \mathbf{j}, \mathbf{k}$ which satisfied $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$ [4]. These three noncommuting numbers can be identified with i (the unit imaginary) multiplied by the three Pauli matrices. This method allows vectors and scalars to be treated on the same level. For example, if $\mathbf{a}$ and $\mathbf{b}$ are vectors, then

$$\underline{\mathbf{a}}\,\underline{\mathbf{b}} = \mathbf{a}\cdot\mathbf{b} I + i\underline{\mathbf{a}\times\mathbf{b}}, \tag{5.1.19}$$

where $\underline{\mathbf{h}} = \mathbf{h}\cdot\boldsymbol{\sigma}$ as before, and I is the 2×2 identity matrix. This formula you have already met as Eq. (5.1.8). Simply using 1 for I, another example is that Maxwell's equations (yes, all four) can be written as

$$\left[\tfrac{1}{c}\tfrac{\partial}{\partial t} - \underline{\boldsymbol{\nabla}}\right]\left[\underline{\mathbf{E}} - i\underline{\mathbf{B}}\right] = \left[\rho + \tfrac{1}{c}\underline{\mathbf{j}}\right]/\epsilon_0. \tag{5.1.20}$$

Fundamentally, this representation exists because the Pauli matrices satisfy a *Clifford algebra*

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij} \tag{5.1.21}$$

appropriate for the Euclidean metric tensor δ_{ij} in three dimensions.

Exercise 5.5 *Verify this result from Eq. (5.1.7).*

5.1.4 * Other Spins

It was stated above that particles of all spins can be formed by combining spin half particles. This is now demonstrated for the case of spin 0 and spin 1, which can be formed from two spin half particles. Let the two basis spinors of one each particle be denoted $|u\rangle$ for the $+1$ eigenstate of σ_3, and $|d\rangle$ for the -1 eigenstate, and let a and b denote the particle. Then one possibility for the four basis states for the compound system is

$$|u_a u_b\rangle, |u_a d_b\rangle, |d_a u_b\rangle, |d_a d_b\rangle. \tag{5.1.22}$$

Here $|u_a u_b\rangle = |u_a\rangle|u_b\rangle$ *etc.* This defines the tensor product of the two Hilbert spaces, as the larger Hilbert space spanned by the basis vectors (5.1.22). We can thus expand an arbitrary state of the joint system as

$$|\psi\rangle = c_{uu}|u_a u_b\rangle + c_{ud}|u_a d_b\rangle + c_{du}|d_a u_b\rangle + c_{dd}|d_a d_b\rangle. \tag{5.1.23}$$

The two particles are said to be not entangled if $|\psi\rangle = |\xi^a\rangle|\xi^b\rangle$. In this case

$$c_{uu} = \xi_1^a\xi_1^b\ ,\ c_{ud} = \xi_1^a\xi_2^b\ ,\ c_{du} = \xi_2^a\xi_1^b\ ,\ c_{dd} = \xi_2^a\xi_2^b, \tag{5.1.24}$$

[4] a formula which Hamilton carved into a stone bridge when it occurred to him on a walk with his wife.

where ξ_1 and ξ_2 are as defined by Eq. (5.1.10). In general, this is not the case, and the two particles are said to be entangled.

Now the spin operator for the compound system is

$$\mathbf{S}_c = \mathbf{S}_a + \mathbf{S}_b. \tag{5.1.25}$$

Although these two matrices are added, it must be remembered that they act in different Hilbert spaces. A more more complete notation for the addition of $\mathbf{S}_a$ and $\mathbf{S}_b$ is

$$\mathbf{S}_c = \mathbf{S}_a \otimes I_b + I_a \otimes \mathbf{S}_b, \tag{5.1.26}$$

where I is the identity operator as above and $\otimes$ is known as the direct product or tensor product. In the tensor product, each element of the first matrix is replaced by itself multiplied by the entire second matrix. Thus the dimension of the tensor product of two matrices is equal to the product of the dimensions of the two matrices. Thus $\mathbf{S}_c$ is a 3-vector of 4×4 matrices. Explicitly,

$$\mathbf{S}_c = \frac{\hbar}{2}\left(\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i & -i & 0 \\ i & 0 & 0 & -i \\ i & 0 & 0 & -i \\ 0 & i & i & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}\right)^T \tag{5.1.27}$$

where the state vector on which this would act is $(c_{uu}, c_{ud}, c_{du}, c_{dd})^T$. The spin-squared operator for the system is

$$\mathbf{S}_c^2 = \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}. \tag{5.1.28}$$

From the result for the $z-$component of $\mathbf{S}_c/\hbar$ it is evident that the spin in that direction can either take the values $-1, 0, 1$, as expected from adding values of $\frac{1}{2}$ or $-\frac{1}{2}$. This eigenvalue we will denote in the usual way as m. The eigenvalues of $\mathbf{S}_c^2/\hbar^2$ are 2 and 0. Since they are defined to be $s(s+1)$ where s is the total spin, we have $s = 0$ or 1, also as expected. Therefore we can find a new basis $|s, m\rangle$ of the joint system in terms of joint eigenvalues of $\mathbf{S}_c^2$ and the $\mathbf{e}_3 \cdot \mathbf{S}_c$. First the single $s = 0$ state is

$$\sqrt{2}|0, 0\rangle = |ud\rangle - |du\rangle. \tag{5.1.29}$$

The three $s = 1$ states are

$$|1, 1\rangle = |uu\rangle \ , \ \sqrt{2}|1, 0\rangle = |ud\rangle + |du\rangle \ , \ |1, -1\rangle = |dd\rangle. \tag{5.1.30}$$

Evidently states with $m = 0$ are entangled states with respect to the original particles. In fact, they are precisely those sorts of states used by Bell to prove the nonlocality inherent in quantum mechanics.

Exercise 5.6 *Verify Eqs. (5.1.27–5.1.30).*

We can write the two-particle wavefunction (5.1.23) using the s, m basis as

$$|\psi\rangle = c_s|0,0\rangle + c_x(|1,1\rangle - |1,-1\rangle)/\sqrt{2} + c_y i(|1,1\rangle + |1,-1\rangle)/\sqrt{2} + c_z(-|1,0\rangle) \tag{5.1.31}$$

From the results of the preceding exercise, it can be shown that, under rotations, c_s transforms as a scalar and (c_x, c_y, c_z) as a vector. It is thus clear why spin zero particles are sometimes called scalar particles, and spin one particles vector particles. A scalar is a rank-zero tensor, and a vector is a rank-one tensor. A spin two particle would be a rank-two tensor particle because its components could be arranged into a matrix which transformed as a rank-two tensor. A spin $\frac{3}{2}$ particle would transform as a vector whose 3 components each were spinors. The relation between the nature of the object that is the wavefunction (spinorial, vectorial *etc.*), and the value of the spin s, reinforces the idea that a spinor is like a tensor of rank $\frac{1}{2}$ (that is, the square root of a vector in some sense).

If we accept that all fundamental particles are fermions of spin half (which will be justified in the next chapter), the spin-statistics theorem for particles also follows from the analysis here. Any particle with integer spin n must be composed of an even number $2n$ of fermions. Swapping two such compound particles swaps $2n$ fermions, so that the total effect is to multiply the wavefunction by $(-1)^{2n} = 1$. That is to say, an integer spin particle must be a boson. Conversely, a half integer spin particle made up of an odd number of fermions must be a fermion.

5.2 Relativity and Quantum Mechanics

5.2.1 Summary of relativistic notation

A coordinate 4-vector is denoted by $x = (x^0, x^1, x^2, x^3) = (ct, \mathbf{x})$. (We use greek-letter indices for 4-vectors and latin-letter indices for 3-vectors.) Under a Lorentz transformation from frame S to frame S'

$$x^\mu \to x^{\mu'} = \Lambda^{\mu'}_{\ \nu}\, x^\nu \,, \tag{5.2.1}$$

where the Einstein summation convention is again being used. The Lorentz transformations includes frames that are rotated relative to one another, and frames that are moving at a constant velocity $\mathbf{v}$ relative to one another (called boosts). For example, for a boost by speed $\mathbf{v} = v\mathbf{e}_1$ (in the x direction),

$$\Lambda = \gamma \begin{pmatrix} 1 & 0 & 0 & -v/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -v/c & 0 & 0 & 1 \end{pmatrix}, \tag{5.2.2}$$

where $\gamma = (1 - v^2/c^2)^{-\frac{1}{2}}$.

In a relativistic theory, physical quantities should transform under Lorentz transformations in well-defined ways. The simplest case is **scalar** quantities, which do not change at all, such as the rest mass of a particle. The next simplest case is a **vector** such as the coordinate vector x

itself. The quantities q^μ are components of a 4-vector (technically, a *contravariant* 4-vector). if they transform in the same way as a coordinate vector, that is, if

$$q^\mu \to q^{\mu'} = \Lambda^{\mu'}{}_\nu \, q^\nu \, . \tag{5.2.3}$$

Writing $q = (q^0, \mathbf{q})$, we define the length of a 4-vector as

$$q^2 = g_{\mu\nu} q^\mu q^\nu = (q^0)^2 - \mathbf{q}\cdot\mathbf{q} \, ; \tag{5.2.4}$$

that is, $g_{00} = 1$, $g_{0j} = 0$, $g_{ij} = -\delta_{ij}$. (N.B. Greek indices run from 0 to 3, Latin from 1 to 3.) Lorentz transformations preserve vector length:

$$q^2 = q'^2 \, . \tag{5.2.5}$$

That is, the length of a 4-vector is a scalar, just as in Euclidean geometry the length of a 3-vector is a scalar.

It is convenient to define the *covariant* version of a contravariant 4-vector by

$$q_\mu = g_{\mu\nu} q^\nu . \tag{5.2.6}$$

This transforms as

$$q_\mu \to q_{\mu'} = \Lambda_{\mu'}{}^\nu \, q_\nu \, . \tag{5.2.7}$$

Note that the derivative with respect to the components of a contravariant vector automatically gives a covariant vector:

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \leftrightarrow \left(\tfrac{1}{c}\partial_t, \boldsymbol{\nabla}\right) . \tag{5.2.8}$$

and vice-versa.

The covariant–contravariant notation allows us to write the analogue of a scalar product for 3-vectors by

$$q.r \equiv q_\mu r^\mu = q^0 r^0 - \mathbf{q}\cdot\mathbf{r} \tag{5.2.9}$$

Just as rotations leave the scalar product of Euclidean geometry unchanged, Lorentz transformations leave the scalar product of 4-vectors unchanged. This requirement implies that

$$\Lambda^{\mu'}{}_\nu \Lambda_{\mu'}{}^\lambda = \delta^\lambda_\nu . \tag{5.2.10}$$

(Compare with Eq. (5.1.1).) This is the only restriction that need be placed on the matrix $\Lambda^{\mu'}{}_\nu$ for it to represent a Lorentz transformation, but we will follow the common convention of also assuming that $\det(\Lambda) = 1$, just as for rotation matrices, and also that $\Lambda^0{}_0 > 0$ (this rules out reversing the arrow of time). Equation (5.2.10) gives 10 equations restricting the elements of Λ, so that the number of free elements is 6. These can be determined from a twist vector $\mathbf{w}$ and a boost vector $\mathbf{v}$.

Exercise 5.7 *Verify that the example (5.2.2) satisfies Eq. (5.2.10).*

Proper time τ is time as measured by a clock at rest in a moving frame. By definition it is a scalar (unchanged by changing frames) and is given by

$$c^2(d\tau)^2 = dx_\mu dx^\mu = c^2dt^2 - d\mathbf{x}^2. \tag{5.2.11}$$

(This follows because if the clock is at rest relative to the moving frame, $d\mathbf{x} = \mathbf{0}$ relative to that frame, so $d\tau = dt$.) That is,

$$d\tau = dt\sqrt{1 - (d\mathbf{x}/dt)^2/c^2} = dt/\gamma, \tag{5.2.12}$$

where γ as defined above is the usual time-dilation factor.

The 4-velocity u is then defined by

$$u^\mu = \frac{dx^\mu}{d\tau} = (c\gamma, \mathbf{v}\gamma)\,, \tag{5.2.13}$$

It follows that

$$\begin{aligned} u^2 &= u_\mu u^\mu && (5.2.14)\\ &= \frac{dx_\mu}{d\tau}\frac{dx^\mu}{d\tau} && (5.2.15)\\ &= \frac{c^2(d\tau)^2}{(d\tau)^2} = c^2. && (5.2.16) \end{aligned}$$

The 4-momentum is then defined as

$$p^\mu = mu^\mu = (E/c, \mathbf{p})^T, \tag{5.2.17}$$

where m is *rest* mass. From the definition of u^μ it follows

$$\mathbf{p} = m\mathbf{v}\gamma. \tag{5.2.18}$$

For low velocities $\gamma \approx 1$, so that $\mathbf{p}$ is approximately the Newtonian momentum, and we refer to it as the 3-momentum. Note also that from Eq. (5.2.16), $p^2 = m^2c^2$. But $p^2 = \mathbf{p}^2 - E^2/c^2$, so

$$E = \sqrt{m^2c^4 + c^2\mathbf{p}^2} \tag{5.2.19}$$

Expanding this by Taylor series, we have

$$E = mc^2 + \frac{\mathbf{p}^2}{2m} + \cdots, \tag{5.2.20}$$

so we identify this as the energy (rest energy plus kinetic energy).

Exercise 5.8 *Verify Eq. (5.2.20) and find the next term in the expansion.*

Note that it is possible to also express the energy as

$$E = \mathbf{p}\cdot\mathbf{v} + \gamma^{-1}mc^2, \tag{5.2.21}$$

Exercise 5.9 *Show this.*

5.2.2 Quantizing time?

The striking new feature of relativity is that it replaces Galileian space-time, in which space and time are distinct entities, by Minkowskian space-time, in which they are inextricably linked. In putting quantum mechanics on a relativistic footing it might therefore be thought that we should quantize time, so that $t \to \hat{t}$ in the same way that $\mathbf{x} \to \hat{\mathbf{x}}$. Some further thought shows that this is untenable.

In non-relativistic quantum mechanics, we are not actually quantizing space when we put $\mathbf{x} \to \hat{\mathbf{x}}$. Rather, we are quantizing a property of the system which happens to equal the position of the particle in space. If there are many particles, we must quantize the position of each one individually: $\mathbf{x}_i \to \hat{\mathbf{x}}_i$, but there is still only one time. In relativistic mechanics, whatever frame we happen to be working in, the system state in configuration space is still defined by the position of each of the particles $\mathbf{x}_i$ at a particular time t. It makes no sense to have an operator $\hat{t}_i$ for each particle; what is the average time for a particle? The usual structure of non-relativistic QM is based on a time evolution equation

$$i\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle, \tag{5.2.22}$$

for a single time t. The effect of the Minkowskian space-time in relativistic QM is that t depends on one's frame of reference, and hence that the wavefunction will transform differently from in non-relativistic QM.

5.3 The Klein-Gordon Equation

The obvious guess for how to make Schrödinger's equation relativistically invariant is as follows. The Schrödinger equation for a free particle with wave function ψ,

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi\,, \tag{5.3.1}$$

can be obtained from the Newtonian equation for kinetic energy,

$$E = \frac{\mathbf{p^2}}{2m}\,, \tag{5.3.2}$$

by the identification

$$E \to i\hbar\frac{\partial}{\partial t}\,,\ \mathbf{p} \to -i\hbar\boldsymbol{\nabla}\,. \tag{5.3.3}$$

We may make the same identification in the relativistic energy equation

$$E^2 = m^2c^4 + c^2\mathbf{p}^2\,, \tag{5.3.4}$$

to give the *Klein-Gordon equation* for a wave function ϕ:

$$(i\hbar\frac{\partial}{\partial t})^2\phi = (i\hbar\nabla)^2c^2\phi + m^2c^4\phi\,. \tag{5.3.5}$$

This can be rearranged as

$$(\partial_\mu\partial^\mu + \lambda_{\rm C}^{-2})\phi = 0\,, \tag{5.3.6}$$

where $\lambda_{\rm C} = \hbar/mc$ is the *Compton wavelength*.

Exercise 5.10 *Show this.*

This equation was suggested independently by no fewer than 6 authors in 1926 (the same year as Schrödinger published his non-relativistic equation), including Schrödinger himself. However, for no particularly good reason, it has come to be known under the name of two of those authors, Klein and Gordon.

5.3.1 Problems with the Klein-Gordon equation

For the Schrödinger equation we have the probability density

$$\rho(\mathbf{x},\mathbf{t}) = \psi^*(\mathbf{x},\mathbf{t})\psi(\mathbf{x},\mathbf{t}) \tag{5.3.7}$$

and probability current

$$\mathbf{j}(\mathbf{x},t) = -\frac{i\hbar}{2m}(\psi^*(\mathbf{x},\mathbf{t})\boldsymbol{\nabla}\psi(\mathbf{x},\mathbf{t}) - \psi(\mathbf{x},\mathbf{t})\boldsymbol{\nabla}\psi^*(\mathbf{x},\mathbf{t}))\,. \tag{5.3.8}$$

In hindsight, the first of these is a fundamental part of quantum mechanics (the probability density to find, at time t, the particle at position $\mathbf{x}$). In 1926 this was not obvious at all. The probability current is such that

$$\frac{\partial}{\partial t}\rho + \boldsymbol{\nabla}\cdot\mathbf{j} = 0. \tag{5.3.9}$$

Exercise 5.11 *Show that this guarantees that the total probability* $\int d^3\mathbf{x}\rho$ *is conserved.*
Hint: *Remember Gauss' theorem.*

It is possible to obtain the equivalent relativistic conservation equation from the Klein-Gordon equation

$$\partial_\mu j^\mu = 0\,, \tag{5.3.10}$$

only by choosing

$$j^\mu = \frac{i\hbar}{2mc}(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*)\,. \tag{5.3.11}$$

However this has

$$\rho = j^0 = \frac{i\hbar}{2mc^2}(\phi^*\frac{\partial\phi}{\partial t} - \phi\frac{\partial\phi^*}{\partial t})\,, \tag{5.3.12}$$

which cannot be interpreted as a probability density because it is not positive definite.

This problem with the KG equation arise from two (related) fundamental problems The first is that it is second-order in time, and hence is not of the form (5.2.22). The second is that it is an equation for a scalar field $\phi(x)$, that is, $\phi'(x') = \phi(x)$ under all Lorentz transformations. This means that $|\phi(x)|^2$ is also a scalar. But relativistically, the density $|\psi(x)|^2$ of a wavefunction should be the time component of a 4-vector. That is needed so that $\int d^3\mathbf{x}|\psi(x)|^2$ is a scalar (equal to one if the wavefunction is normalized). In a moving frame $d^3\mathbf{x}$ shrinks by a

factor of γ^{-1} due to Lorentz contraction, so that $|\psi(x)|^2$ must increase by a factor γ, as indeed it would if it were the 0-component of a 4-vector.

Of course, these criticisms do not imply that there is anything wrong with the Klein-Gordon equation *per se*. It is simply that it cannot be interpreted as an evolution equation for the wavefunction of a quantum particle. It is possible in fact to reinterpret $\phi(x)$ as a classical field (which can be quantized just like the electromagnetic field), as we will discuss in Chap. 6. In this case, ρ and $\mathbf{j}$ appear as *charge* density and current, so that negative density is not a difficulty.

5.4 The Dirac Equation: Derivation

5.4.1 Introduction

Since we require $|\psi(x)|^2$ to be the 0-component of a 4-vector, it is evident that $\psi(x)$ itself must be something like the square root of a 4-vector. In Sec. 5.1.3 we saw that a spinor is something like the square root of a 3-vector, which gives us hope that a relativistic theory based on spinors will work. Before proceeding, it may be good for us to have a peek at the end result:

$$(-i\gamma^\nu \partial_\nu + \lambda_{\mathrm{C}}^{-1})\psi = 0, \tag{5.4.1}$$

where $\lambda_{\mathrm{C}} = \hbar/mc$ as above, and the γ^ν will be explained below. This equation was derived by Dirac in 1928. Unlike the Klein-Gordon equation, Dirac's equation is a work of genius, and nobody else was even close to deriving it at the time that he did.

5.4.2 First Derivation

Inspired by Eq. (5.1.19), we can rewrite Eq. (5.3.4) as

$$E^2/c^2 - (\boldsymbol{\sigma}\cdot\mathbf{p})^2 = m^2c^2, \tag{5.4.2}$$

which can be factorized as

$$(E/c - \boldsymbol{\sigma}\cdot\mathbf{p})\,(E/c + \boldsymbol{\sigma}\cdot\mathbf{p}) = m^2c^2. \tag{5.4.3}$$

If ξ is the spinor wavefunction for a relativistic electron, it evidently obeys

$$(i\hbar\partial_0 + i\hbar\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\,(i\hbar\partial_0 - i\hbar\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\,\xi = m^2c^2\xi, \tag{5.4.4}$$

which is simply the KG equation for each component.

Now define a new spinor η as

$$mc\eta = (i\hbar\partial_0 - i\hbar\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\,\xi \tag{5.4.5}$$

Then we have

$$mc\xi = (i\hbar\partial_0 + i\hbar\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\,\eta. \tag{5.4.6}$$

We can make these two coupled equations one equation by defining a 4-component wavefunction

$$\psi = \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \eta_1 \\ \eta_2 \end{pmatrix}. \tag{5.4.7}$$

This wavefunction is known as a *bispinor* or *Dirac spinor*. Note that it is *not* a 4-vector. Now if we define a 4-vector of 4×4 matrices

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \boldsymbol{\gamma} = \begin{pmatrix} \mathbf{0} & -\boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{pmatrix}. \tag{5.4.8}$$

then Eqs. (5.4.5) and (5.4.6) are equivalent to the single equation,

$$(-i\hbar\gamma^\mu\partial_\mu + mc)\psi(x) = 0. \tag{5.4.9}$$

which is Dirac's equation as already stated in Eq. (5.4.1).

Exercise 5.12 *Show this equivalence.*

5.4.3 * Second Derivation (Minkowskian Clifford Algebra)

A quicker, but perhaps less enlightening, path to Dirac's equation is via the Minkowskian Clifford algebra. Recall that non-relativistically the Pauli matrices could be used to define a 3-dimensional Euclidean Clifford Algebra with

$$\tfrac{1}{2}\{\sigma_i, \sigma_j\} = \delta_{ij}, \tag{5.4.10}$$

where here the curly brackets indicate an anticommutator. The relativistic analogue is a set of four basis matrices γ_μ satisfying

$$\tfrac{1}{2}\{\gamma_\mu, \gamma_\nu\} = g_{\mu\nu}. \tag{5.4.11}$$

In an analogy to the non-relativistic $\underline{\mathbf{h}} = \boldsymbol{\sigma}\cdot\mathbf{h}$, here we define a "slash" notation for a 4-vector

$$\not{A} = \gamma_\mu A^\mu. \tag{5.4.12}$$

Then it follows that for arbitrary vectors

$$A.B \equiv A_\mu B^\mu = \tfrac{1}{2}\{\not{A}, \not{B}\}, \tag{5.4.13}$$

which is the analogue of the non-relativistic $\mathbf{a}\cdot\mathbf{b} = \frac{1}{2}\{\underline{\mathbf{a}}, \underline{\mathbf{b}}\}$.

In particular, the relativistic relation between energy and momentum

$$p_\mu p^\mu = m^2c^2 \tag{5.4.14}$$

can be factorized as

$$\not{p} = mc. \tag{5.4.15}$$

Introducing the usual quantum-mechanical relation $p_\mu \to -i\hbar\partial_\mu$ yields

$$(-i\not{\partial} + \lambda_{\mathrm{C}}^{-1})\psi = 0 \tag{5.4.16}$$

which is Dirac's equation in in its most succinct form.

Exercise 5.13 *Verify that the γ^μ as defined in Eq. (5.4.8) do obey the anticommutation relations Eq. (5.4.11).*

In this derivation, the γ matrices need not be 4×4. However it can be shown that this is the minimum size of any set of matrices which satisfy the Minkowskian Clifford algebra (5.4.11). In the same way, the minimum size of the Pauli matrices (2×2) is set by the Euclidean Clifford algebra. Even if we choose 4 dimensions, there is no reason that the γ matrices must be the same as those used above (5.4.8). In fact, Dirac's original choice was

$$\gamma^0 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \boldsymbol{\gamma} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & \mathbf{0} \end{pmatrix}, \tag{5.4.17}$$

This is known as the Dirac-Pauli representation, and will be used in all that follows unless otherwise stated. The previous representation is known as the chiral representation, for reasons we will not explore.

Exercise 5.14 *Verify that the γ^μ as defined in Eq. (5.4.17) also obey the anticommutation relations Eq. (5.4.11).*

5.5 The Dirac Equation: Properties

5.5.1 Hamiltonian form

We can rewrite the Dirac equation as

$$i\hbar \frac{\partial}{\partial t}\psi = \hat{H}\psi \,, \tag{5.5.1}$$

$$\hat{H} = c\hat{\boldsymbol{\alpha}}\cdot\hat{\mathbf{p}} + \hat{\beta}mc^2 \,, \tag{5.5.2}$$

where

$$\hat{\boldsymbol{\alpha}} = \gamma^0\boldsymbol{\gamma} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{pmatrix}, \tag{5.5.3a}$$

$$\hat{\beta} = \gamma^0 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}. \tag{5.5.3b}$$

Exercise 5.15 *Show that this is equivalent to Eq. (5.4.16).*

Equation (5.4.16) was in fact Dirac's original version (using the Dirac-Pauli representation of the γ matrices). Compare this to that other expression (5.2.21) for the relativistic energy

$$E = \mathbf{v}\cdot\mathbf{p} + \gamma(\mathbf{v})^{-1}mc^2, \tag{5.5.4}$$

where $\gamma(\mathbf{v})^{-1} = \sqrt{1 - \mathbf{v}^2/c^2}$ as usual. This suggests that we may expect $c\hat{\boldsymbol{\alpha}}$ to play a role analogous to the velocity, and $\hat{\beta}$ to γ^{-1}. Although the former analogue works well, $\hat{\beta}$ just as often takes on the role of γ, since $\hat{\beta}^2 = 1$.

5.5.2 Probability density

If we have succeeded in deriving a true relativistic version of Schrödinger's equation, the following

$$\rho(x) = \psi^\dagger\psi = |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2 \,. \tag{5.5.5}$$

should be a valid probability density for the Dirac particle. Now we check that.

$$i\hbar\frac{\partial\rho}{\partial t} = i\hbar\left(\frac{\partial\psi^\dagger}{\partial t}\psi + \psi^\dagger\frac{\partial\psi}{\partial t}\right) \tag{5.5.6}$$

$$= -i\hbar c\left[(\boldsymbol{\nabla}\psi)^\dagger\cdot\hat{\boldsymbol{\alpha}}^\dagger\psi + \psi^\dagger\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{\nabla}\psi\right] . \tag{5.5.7}$$

Since $\hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\alpha}}^\dagger$, we can write this as a continuity equation

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0 , \tag{5.5.8}$$

by defining the probability current to be

$$\mathbf{j} = c\psi^\dagger\hat{\boldsymbol{\alpha}}\psi , \tag{5.5.9}$$

as expected if $c\hat{\boldsymbol{\alpha}}$ is the velocity. Integrating over all space confirms the conservation of probability:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^3\mathbf{x}\rho = 0 . \tag{5.5.10}$$

So we *can* take Eq. (5.5.5) as defining a probability density. The probability current 4-vector $j = (c\rho, \mathbf{j})$ is then

$$j^\mu = \psi^\dagger(c, c\hat{\boldsymbol{\alpha}})^\mu\psi = c\psi^\dagger\gamma^0\gamma^\mu\psi . \tag{5.5.11}$$

5.5.3 Motion of a free Dirac particle

Working now in the Heisenberg picture, and using the Dirac Hamiltonian Eq. (5.5.2), the velocity of the electron is

$$\hat{\mathbf{v}} = \frac{d\hat{\mathbf{r}}}{dt} = \frac{-i}{\hbar}[\hat{\mathbf{r}}, \hat{H}] = c\hat{\boldsymbol{\alpha}}, \tag{5.5.12}$$

as expected from previous results. Note that compenents of $\hat{\boldsymbol{\alpha}} = \{\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3\}$ do not commute with eachother, and each has eigenvalues ± 1.

Exercise 5.16 *Show this.*

This means that the velocity can be perfectly measured only in one direction at a time, and that the result of such a perfect measurement would be the speed of light! In practice one cannot measure velocity instantaneously, and it is usually inferred from a time-of-flight measurement, involving a (necessarily imperfect) initial position measurement and a final position measurement, as

$$\hat{\mathbf{v}}_{\text{inferred}}(t_\text{i}) = \frac{\hat{\mathbf{r}}_\text{f} - \hat{\mathbf{r}}_i}{t_\text{f} - t_\text{i}}. \tag{5.5.13}$$

It can be shown that for a sufficiently imperfect initial position measurement,

$$\lim_{t_\text{f}\to\infty}\hat{\mathbf{v}}_{\text{inferred}}(t_\text{i}) = \frac{c^2\hat{\mathbf{p}}(t_i)}{\hat{H}(t_i)}. \tag{5.5.14}$$

This is as expected classically, and can take values from 0 to c in any direction.

Exercise 5.17 *Show that this is indeed what is expected classically. Also explain why the first position measurement must be imperfect.*

5.5.4 * Lorentz Transformations

Recall that under a rotation an ordinary (two-component) spinor translates according to

$$\xi' = \exp(-i\tfrac{1}{2}\mathbf{w}\cdot\boldsymbol{\sigma})\xi, \tag{5.5.15}$$

which is actually a completely general unitary transformation on a spinor (apart from a missing global phase factor). Now a Lorentz transformation has 6 parameters ($\mathbf{v}$ and $\mathbf{w}$), which is another way of seeing why we need two spinors in the relativistic case. The Lorentz transformation of the bispinor is most easily expressed using the chiral representation of Sec. 5.4.2. It is:

$$\begin{pmatrix} \xi' \\ \eta' \end{pmatrix} = \begin{pmatrix} \exp[-i\frac{1}{2}(\mathbf{w}+i\mathbf{v})\cdot\boldsymbol{\sigma}] & 0 \\ 0 & \exp[-i\frac{1}{2}(\mathbf{w}-i\mathbf{v})\cdot\boldsymbol{\sigma}] \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}. \tag{5.5.16}$$

One can actually derive this transformation equation from basic properties of the Lorentz group, and moreover one can then use this transformation equation to derive Dirac's equation itself. But we will not go into those details.

5.6 The Dirac Equation: Solutions

5.6.1 Momentum eigenstates

We start from the Hamiltonian form of the Dirac equation:

$$i\hbar\frac{\partial\psi}{\partial t} = [c\hat{\boldsymbol{\alpha}}\cdot(-i\hbar\boldsymbol{\nabla}) + \hat{\beta}mc^2]\psi\,. \tag{5.6.1}$$

Dirac originally suggested this equation for a free electron, so we expect to be able to find plane wave solutions, corresponding to a well-defined momentum $\mathbf{p}$:

$$\psi(x) = \exp(\frac{-ip.x}{\hbar})u(\mathbf{p})\,, \tag{5.6.2}$$

where $p.x = Et - \mathbf{p}\cdot\mathbf{x}$ and $u(\mathbf{p})$ is a bispinor that is independent of space and time. If we substitute this in as a trial solution, we get

$$Eu(\mathbf{p}) = (c\hat{\boldsymbol{\alpha}}\cdot\mathbf{p} + \hat{\beta}mc^2)u(\mathbf{p})\,. \tag{5.6.3}$$

It is now an algebraic problem to find possible forms of $u(\mathbf{p})$. We can exploit the block form of the Dirac-Pauli representation of $\hat{\boldsymbol{\alpha}}$ and $\hat{\beta}$ by introducing the two-component (non-relativistic) spinors χ_{A} and χ_{B}:

$$u(\mathbf{p}) = \begin{pmatrix} \chi_{\mathrm{A}} \\ \chi_{\mathrm{B}} \end{pmatrix}\,. \tag{5.6.4}$$

Now the Dirac equation in terms of $\chi_{\mathrm{A}}, \chi_{\mathrm{B}}$ is

$$\left\{E - c\begin{pmatrix} 0 & \boldsymbol{\sigma}\cdot\mathbf{p} \\ \boldsymbol{\sigma}\cdot\mathbf{p} & 0 \end{pmatrix} - \begin{pmatrix} mc^2 & 0 \\ 0 & -mc^2 \end{pmatrix}\right\}\begin{pmatrix} \chi_{\mathrm{A}} \\ \chi_{\mathrm{B}} \end{pmatrix} = 0\,, \tag{5.6.5}$$

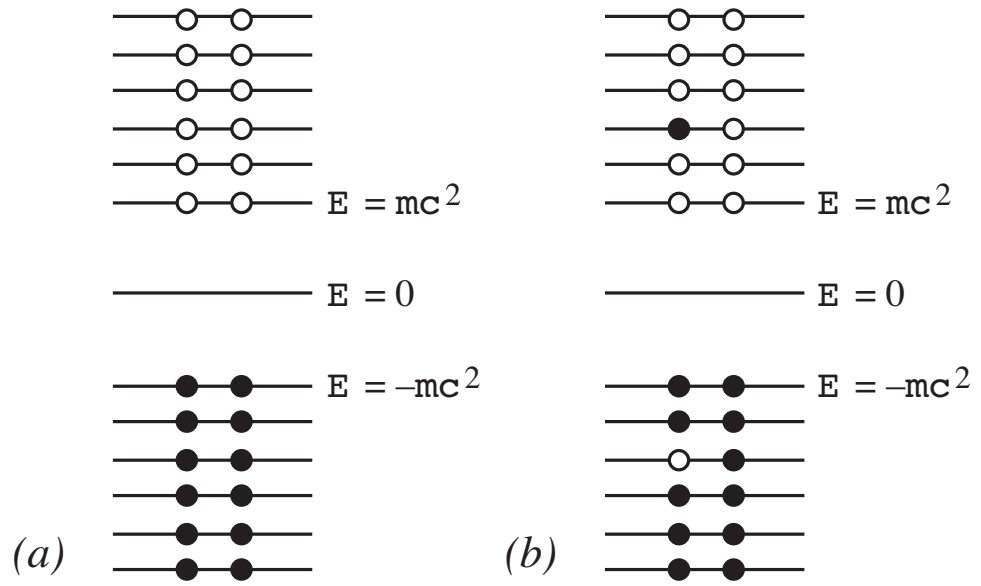

Figure 5.1: The Dirac sea of electrons *(a)* in the ground state and *(b)* after the creation of one electron-positron pair.

or, rearranging,

$$(E - mc^2)\chi_{\text{A}} = c\boldsymbol{\sigma}\cdot\mathbf{p}\chi_{\text{B}} \quad (5.6.6a)$$
$$(E + mc^2)\chi_{\text{B}} = c\boldsymbol{\sigma}\cdot\mathbf{p}\chi_{\text{A}}\,. \quad (5.6.6b)$$

Exercise 5.18 *Show that for these two equations to be consistent, the usual relativistic energy-momentum relation* $E^2 = c^2|\mathbf{p}|^2 + m^2c^4$ *must hold.*

Note that for each value of $\mathbf{p}$, negative energies as well as positive energies are still allowed: $E = \pm\sqrt{c^2|\mathbf{p}|^2 + m^2c^4}$. In fact, these occur twice each (4 eigenvalues for a 4×4 matrix). Negative energies could also exist classically, but do not pose a problem as classical energy must vary continuously, so an electron starting with $E > mc^2$ could never reach a negative energy state. Quantum mechanics, however, allows particles to jump from one state to another, so an electron in a positive energy state may make a transition to a negative energy one, despite the intervening band of impossible energies ($-mc^2 < E < +mc^2$).

5.6.2 Particles and Holes

It might be thought that the existence of negative energy states is a problem with the Dirac theory. From Eq. (5.5.14), a negative energy (or *anomalous*) electron would move in the opposite direction to its momentum, which has never been observed. Also, since there is no lower limit to the negative energies, the system would be unstable as electrons could keep losing energy, while going faster and faster in the "wrong" direction.

In fact, the anomalous states are one of the triumphs of the Dirac theory. Dirac realized that these problems would be solved if one assumed

that the universe were full of electrons in the anomalous states. For a finite normalization volume L^3 the number of anomalous states is infinite but countable, so it is possible to imagine them all being filled. The infinite negative charge and negative energy resulting from this "Dirac sea" would not be observable because it is equally spread everyhwere in space. Of course creating this sea would not be possible if electrons were bosons (as bosons can never "fill" an energy level), and we can also view this as a proof of why electrons (and other fundamental particles) must be fermions. Any extra electrons would be forced into the usual positive energy states and would appear as usual electrons. Any *holes* (i.e. missing electrons) in the anomalous states would appear as missing negative energy, missing negative momentum (that is, in the opposite direction to the velocity), and missing negative charge. That is, it would appear as a electron with positive energy, positive momentum (in the same direction as the velocity) and *positive* charge. In other words, it would be an *anti-electron* or *positron*. See Fig. 5.1 and the table below.

Table 1: Antiparticles are holes in the Dirac sea.

Particle	Velocity	Momentum	Energy	Charge
electron (e)	$\longrightarrow$	$m\gamma\times\longrightarrow$	$> mc^2$	$-e$
anomalous electron (æ)	$\longrightarrow$	$m\gamma\times\longleftarrow$	$< -mc^2$	$-e$
positron (missing æ)	$\longrightarrow$	$m\gamma\times\longrightarrow$	$> mc^2$	$+e$

If an anomalous electron is given enough energy to turn it into an ordinary electron, this creates a hole (a positron) and an ordinary electron. This is the phenomenon of pair creation. The amount of energy required is at least $2mc^2$ (the energy gap). This is exactly as expected for creating two particles each with rest mass m.

The prediction of antiparticles by Dirac in 1930 was one of the many triumphs of his equation, and positrons were first observed in 1932. This explanation of antiparticles relies on the exclusion principle, and so implies that electrons must be fermions. More generally, it requires that all fundamental particles (being spin half) must be fermions. This is part of the proof of the spin-statistics theorem, which will be discussed in greater detail in the next chapter.

5.6.3 Non-relativistic behaviour

The occupation of the negative energy states does not affect low-energy electron dynamics, again because of the exclusion principle — the electron cannot make a transition to a full state. For $v \ll c$ we can take

$$\frac{c|\mathbf{p}|}{E+mc^2} \ll 1 \tag{5.6.7}$$

$$E \approx mc^2\,. \tag{5.6.8}$$

Then from Eqs. (5.6.6a) and (5.6.6b), χ_{B} will be much smaller than χ_{A}. Thus in this limit, we can ignore χ_{B} and use only χ_{A}. This simple non-relativistic limit is the principal attraction of the standard Dirac-Pauli representation. It is obviously consistent with the standard non-

relativistic description of electrons using 2-spinors. We will derive the non-relativistic equation for χ_{A} after we have included the coupling to the electromagnetic field, in the next chapter.

6 Relativistic QED and the Standard Model

6.1 Dirac's Relativistic QED

6.1.1 The minimal coupling

Recall that the minimal coupling Hamiltonian arises from the replacement

$$H = H_0(\mathbf{v}) + q\phi(\mathbf{r}), \tag{6.1.1}$$
$$\mathbf{p} = \mathbf{p}_0(\mathbf{v}) + q\mathbf{A}(\mathbf{r}). \tag{6.1.2}$$

In the relativistic case, these are combined in the 4-vector transformation

$$p^\mu = p_0^\mu + qA^\mu, \tag{6.1.3}$$

where $A^\mu = (\phi/c, \mathbf{A})$ is the 4-vector potential. The Dirac equation for an electron (with charge $q = -e$) coupled to a field is thus

$$(-i\partial\!\!\!/ + eA\!\!\!//\hbar + \lambda^{-1})\psi = 0 \tag{6.1.4}$$

In the Hamiltonian formulation, the Dirac Hamiltonian changes into

$$\hat{H} = \hat{H}_{\rm el} + \hat{H}_{\rm I} \tag{6.1.5}$$

where $\hat{H}_{\rm el} = c\hat{\boldsymbol{\alpha}}\cdot\hat{\mathbf{p}} + \hat{\beta}mc^2$ and

$$\hat{H}_{\rm I} = ec\hat{\boldsymbol{\alpha}}\cdot\hat{\mathbf{A}}(\hat{\mathbf{r}}) - e\hat{\phi}(\hat{\mathbf{r}}). \tag{6.1.6}$$

Here we have included a quantized field. In contrast to the nonrelativistic expression in Eq. (3.4.1), there are just two terms in $\hat{H}_{\rm I}$. The first involves the annihilation or creation of a single photon, and the second does not involve photons. Once again the electron is not created or destroyed, but may change state. However, in Dirac theory an electron can change state from positive to negative energy states (or vice versa) via the electromagnetic field. This is observable as the annihilation (or creation) of electron-positron pairs.

6.1.2 * The Dyson Series

Consider a system of charged particles (labelled by l) interacting with the electromagnetic field. The total Hamiltonian is

$$\hat{H} = \hat{H}_0 + \hat{V}, \tag{6.1.7}$$

where $\hat{H}_0$ is the total "free Hamiltonian":

$$\hat{H}_0 = \sum_l \left[c\hat{\boldsymbol{\alpha}}_l\cdot\hat{\mathbf{p}}_l + \hat{\beta}_l m_l c^2\right] + \int d^3\boldsymbol{x}\frac{\epsilon_0}{2}\left((\mathbf{E}^\perp)^2(\boldsymbol{x}) + c^2\mathbf{B}^2(\boldsymbol{x})\right) \tag{6.1.8}$$

and $\hat{V}$ is the total "interaction Hamiltonian",

$$\hat{V} = \frac{1}{4\pi\epsilon_0}\sum_{l\neq l'}\frac{q_l q_{l'}}{|\hat{\mathbf{r}}_l - \hat{\mathbf{r}}_{l'}|} + \sum_l q_l c\hat{\boldsymbol{\alpha}}_l\cdot\hat{\mathbf{A}}(\hat{\mathbf{r}}_l). \tag{6.1.9}$$

The free Hamiltonian is easy to deal with — it describes particles that propagate with constant momentum (we showed this in Sec. 5.6.1) and a collection of Harmonic oscillators for the field (we showed this in Sec. 1.3.3). That is to say, the unitary evolution

$$\hat{U}_0(t,0) \equiv \exp(-i\hat{H}_0 t/\hbar) \tag{6.1.10}$$

is easy to apply to any initial state. By contrast the interaction Hamiltonian is very difficult to deal with, as it describes electrons interacting with eachother, or interacting with photons in the field. For this reason it is very difficult to work out the action of the full unitary evolution

$$\hat{U}(t,0) \equiv \exp[-i(\hat{H}_0 + \hat{V})t/\hbar]. \tag{6.1.11}$$

It is however possible to develop an approximation to $\hat{U}(t,0)$ that can be applied. First note that for *any* operators $\hat{H}_0$ and $\hat{V}$,

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\hat{U}(t,0) &= [\hat{H}_0 + \hat{V}]\hat{U}(t,0) && (6.1.12)\\ i\hbar\frac{\partial}{\partial t}\hat{U}_0(t,0) &= \hat{H}_0\hat{U}_0(t,0), && (6.1.13) \end{aligned}$$

and (taking the Hermitian adjoint of the last equation)

$$-i\hbar\frac{\partial}{\partial t}\hat{U}_0^\dagger(t,0) = \hat{U}_0^\dagger(t,0)\hat{H}_0 \tag{6.1.14}$$

Now consider the following unitary operator:

$$\hat{u}(t,0) \equiv \hat{U}_0^{-1}(t,0)\hat{U}(t,0) = \hat{U}_0^\dagger(t,0)\hat{U}(t,0), \tag{6.1.15}$$

so that $\hat{U}(t,0) = \hat{U}_0(t,0)\hat{u}(t,0)$. This obeys the following differential equation:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\hat{u}(t,0) &= \left(i\hbar\frac{\partial}{\partial t}\hat{U}_0^\dagger(t,0)\right)\hat{U}(t,0) + \hat{U}_0^\dagger(t,0)\left(i\hbar\frac{\partial}{\partial t}\hat{U}(t,0)\right)\\ &= -\hat{U}_0^\dagger(t,0)\hat{H}_0\hat{U}(t,0) + \hat{U}_0^\dagger(t,0)(\hat{H}_0 + \hat{V})\hat{U}(t,0)\\ &= \hat{U}_0^\dagger(t,0)\hat{V}\hat{U}_0(t,0)\hat{u}(t,0). && (6.1.16) \end{aligned}$$

That is to say, $\hat{u}(t,0)$ is a unitary evolution operator that would result from a *time-dependent* Hamiltonian

$$\hat{v}(t) = \hat{U}_0^\dagger(t,0)\hat{V}\hat{U}_0(t,0). \tag{6.1.17}$$

Now we know from Sec. 1.2 that the unitary operator resulting from a time-dependent Hamiltonian is

$$\hat{u}(t,0) = I + \sum_{n=1}^{\infty}(i\hbar)^{-n}\int_0^t ds_n \hat{v}(s_n)\int_0^{s_n} ds_{n-1}\hat{v}(s_{n-1})\cdots\int_0^{s_2} ds_1\hat{v}(s_1). \tag{6.1.18}$$

Expanding this out and substituting in Eq. (6.1.17), we have the full evolution

$$\begin{aligned}\hat{U}(t,0) &= \hat{U}_0(t,0) \\ &+ (i\hbar)^{-1}\int_0^t ds_1 \hat{U}_0(t,s_1)\hat{V}\hat{U}_0(s_1,0) \\ &+ (i\hbar)^{-2}\frac{1}{2}\int_0^t ds_2 \int_0^{s_2} ds_1 \hat{U}_0(t,s_2)\hat{V}\hat{U}_0(s_2,s_1)\hat{V}\hat{U}_0(s_1,0) \\ &+ (i\hbar)^{-2}\frac{1}{3!}\int_0^t ds_3 \int_0^{s_3} ds_2 \int_0^{s_2} ds_1 \hat{U}_0(t,s_3)\hat{V}\hat{U}_0(s_3,s_2)\hat{V}\hat{U}_0(s_2,s_1)\hat{V}\hat{U}_0(s_1,0) \\ &+ \cdots . \end{aligned} \tag{6.1.19}$$

Exercise 6.1 *Show this.*

This is known as a *Dyson expansion* after the work of British physicist Freeman Dyson in the 1940s. The zeroth-order term is where the electrons and photons don't interact. The first order term is where they freely evolve up to time s_1, then interact, then freely evolve from time s_1 to t, and we have to integrate over all possible interaction times s_1. You can tell the story for the higher-order terms.

6.1.3 Feynman Diagrams

There is one-to-one relation between terms in the Dyson exansion, and Feynman diagrams, named after the simultaneous work of US physicist Richard Feynman. In these diagrams, free evolution corresponds to lines (wiggly for photons and straight for electrons), while interactions correspond to vertices. The minimal coupling Hamiltonian gives "three-line vertices", two straight and one wiggly.[5] Thus as well as giving a picture of quantum interactions, Feynman diagrams are actually a tool for figuring out what the relevant terms in the Dyson expansion are. A Feynman diagram with n vertices is a contribution to the nth order term in the Dyson expansion. Fermi's golden rule comes out from using just the first order term in the Dyson expansion.

6.1.4 The nonrelativistic limit

For simplicity, we will now treat the field classically (which just means dropping the hats on the field). Then the two-component equations for fixed energy E, Eq. (5.6.6), become

$$(E + e\phi - mc^2)\chi_{\mathrm{A}} = c\boldsymbol{\sigma}\cdot(\hat{\mathbf{p}} + e\mathbf{A})\chi_{\mathrm{B}} \tag{6.1.20a}$$

$$(E + e\phi + mc^2)\chi_{\mathrm{B}} = c\boldsymbol{\sigma}\cdot(\hat{\mathbf{p}} + e\mathbf{A})\chi_{\mathrm{A}}\,. \tag{6.1.20b}$$

In the non-relativistic regime $\chi_{\mathrm{A}} \gg \chi_{\mathrm{B}}$, so we consider the equation for χ_{A} only.

$$(E + e\phi + mc^2)(E + e\phi - mc^2)\chi_{\mathrm{A}}$$

[5] Because of the gauge we are working in there are direct electron-electron "four-line" vertices from the Coulomb Hamiltonian. In other gauges these sorts of vertices don't occur, but there are other complications.

$$
\begin{aligned}
&= (E + e\phi + mc^2) c\boldsymbol{\sigma}\cdot(\hat{\mathbf{p}} + e\mathbf{A})\chi_{\mathrm{B}} \\
&= c\left(\boldsymbol{\sigma}\cdot(\hat{\mathbf{p}} + e\mathbf{A})(E + e\phi + mc^2) + [e\phi(\hat{\mathbf{r}}), c\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}]\right)\chi_{\mathrm{B}} \\
&\approx c^2\left(\boldsymbol{\sigma}\cdot(\hat{\mathbf{p}} + e\mathbf{A})\right)^2\chi_{\mathrm{A}}\,, \qquad (6.1.21)
\end{aligned}
$$

where the last line follows if we neglect the commutator (see discussion below). From the properties of the Pauli matrices (5.1.8),

$$
\begin{aligned}
&\boldsymbol{\sigma}\cdot(\hat{\mathbf{p}} + e\mathbf{A})\boldsymbol{\sigma}\cdot(\hat{\mathbf{p}} + e\mathbf{A}) \\
&\quad = (\hat{\mathbf{p}} + e\mathbf{A})^2 + i\boldsymbol{\sigma}\cdot[(\hat{\mathbf{p}} + e\mathbf{A}) \times (\hat{\mathbf{p}} + e\mathbf{A})] \\
&\quad = (\hat{\mathbf{p}} + e\mathbf{A})^2 + i\boldsymbol{\sigma}\cdot[(\hat{\mathbf{p}} \times e\mathbf{A}(\hat{\mathbf{r}}) + (e\mathbf{A}(\hat{\mathbf{r}}) \times \hat{\mathbf{p}})] \\
&\quad = (\hat{\mathbf{p}} + e\mathbf{A})^2 + \hbar\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla} \times e\mathbf{A}(\hat{\mathbf{r}})) \\
&\quad = (\hat{\mathbf{p}} + e\mathbf{A})^2 + \hbar e\boldsymbol{\sigma}\cdot\mathbf{B}\,. \qquad (6.1.22)
\end{aligned}
$$

We can write $E = mc^2 - e\phi + E_{\mathrm{K}}$, and for the non-relativistic regime in which $E_{\mathrm{K}} \ll mc^2$ we then have

$$
(E + e\phi + mc^2)(E + e\phi - mc^2) \approx 2E_{\mathrm{K}}mc^2\,, \qquad (6.1.23)
$$

and hence

$$
E_{\mathrm{K}}\chi_{\mathrm{A}} \approx \left(\frac{1}{2m}(\hat{\mathbf{p}} + e\mathbf{A})^2 + \frac{\hbar e}{2m}\boldsymbol{\sigma}\cdot\mathbf{B}\right)\chi_{\mathrm{A}}\,. \qquad (6.1.24)
$$

Finally, introducing the total non-relativistic energy $\Delta E = E - mc^2$,

$$
\Delta E\chi_{\mathrm{A}} \approx \left(\frac{1}{2m}(\hat{\mathbf{p}} + e\mathbf{A})^2 - e\phi + \frac{\hbar e}{2m}\boldsymbol{\sigma}\cdot\mathbf{B}\right)\chi_{\mathrm{A}}\,. \qquad (6.1.25)
$$

The term we neglected to obtain Eq. (6.1.21) is of order $e\phi'\lambda_{\mathrm{C}}$, the change in the potential energy of the electron across the Compton wavelength of the electron.

Exercise 6.2 *Evaluate λ_{C}. Also write it in terms of the "classical electron radius" r_0 and α.*

Since λ_{C} is so small, this energy is typically much smaller than the other energies here. If it is not, then this whole approach fails; physically, such strong fields lead to pair creation.

Equation (6.1.25) is just the time-independent Schrödinger equation for a non-relativistic charged particle in an electromagnetic field, plus the $-\boldsymbol{\mu}\cdot\mathbf{B}$ term for a magnetic moment, with

$$
\boldsymbol{\mu} = \frac{-e\hbar}{2m}\boldsymbol{\sigma}\,. \qquad (6.1.26)
$$

In terms of the electron spin,

$$
\boldsymbol{\mu} = \frac{-e}{m}\mathbf{S}\,. \qquad (6.1.27)
$$

For a classical electron in a closed orbit of angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, the effective magnetic moment would be

$$
\boldsymbol{\mu}_{\mathrm{classical}} = \frac{-e}{2m}\mathbf{L} \qquad (6.1.28)
$$

Exercise 6.3 *(Challenging.) Show that for a constant magnetic field one can write* $\mathbf{A} = -\frac{1}{2}\mathbf{r}\times\mathbf{B}$. *From the non-relativistic classical minimal coupling Hamiltonian, show that a perturbatively small* $\mathbf{B}$ *produces an energy shift of* $-\mathbf{B}\cdot\boldsymbol{\mu}_{\text{classical}}$.

So quantum spin angular momentum gives a contribution to the magnetic moment twice as large as the same amount of orbital angular momentum. This is another triumph of Dirac's equation, as the so-called anomalous g factor of the electron was a mystery at the time (1928). The extra factor of 2 is sometimes called the Landé g-factor. (The full theory of QED predicts small corrections to this; the corrections have been experimentally confirmed to more than 13 decimal places, making the g-factor of the electron one of the most precisely predicted of all physical quantities.)

6.2 * Fundamental Particles

We have argued above that the only good relativistic wavefunction is the Dirac bispinor. This means that all fundamental particles should, like electrons,

1. be spin-half particles
2. be Fermions
3. have antiparticles of opposite charge
4. have a magnetic moment of (approximately) $2\times q/4m$, where q is the charge and m the mass.

Electrons and neutrinos (which are neutral) certainly obey these rules. What about the heavier constituents of matter? Protons and neutrons both obey all of the above rules, except for the last. The magnetic moment of a proton is about $2.8\times e/4m_pc$, while that of a neutron (which should be zero since it is neutral) is about $-1.9\times e/4m_nc$. This suggests that protons and neutrons are not fundamental particles.

From the accumulated evidence of many experiments it became generally accepted in the early 1970s that nucleons (protons and neutrons) are compound particles made up of quarks. For example, deep (very high energy) inelastic scattering of electrons off protons are consistent with the presence of point-like particles inside the proton. As far as anyone knows, quarks (named by Murray Gell-Mann) are fundamental particles like electrons and neutrinos. Two sorts of quarks are needed to explain protons and neutrons, up quarks u (with charge $+\frac{2}{3}e$) and down quarks d (with charge $-\frac{1}{3}e$). A proton is the compound particle uud and the neutron the compound particle ddu. The orbital motion of quarks within a nucleon explains why the magnetic moment of nucleons is not simply related to their charge and spin. One of the recent triumphs of the quark model is the calculation of these magnetic moments to good agreement with experiment.

Like electrons and neutrinos, quarks have antiquarks represented as $\bar{\text{u}}, \bar{\text{d}}$. In addition to electrical charge, quarks have another sort of charge,

called *colour charge*. This charge has three values, called r, g, b for red, green, blue. These are so-called because a mixture of all three gives white w, representing a lack of colour charge. The three quarks in a nucleon always have colours such that the nucleon is colour-neutral. Antiquarks have the anti-colour values c, m, y for cyan, magenta, yellow (these are the colours opposite red, green, blue on a colour wheel). In a *meson* (which is composed of a quark and an anti-quark), a colour is always paired with its anti-colour, so that mesons are also colour neutral. The fundamental particles which make up the matter of ordinary experience are summarized in table 2

Table 2: The first family of fundamental particles

Particle	E. Charge	Colour Charge	Lepton #	Baryon #
e	-e	w	1	0
$\bar{\text{u}}$	$-(2/3)e$	c, m, y	0	$-1/3$
d	$-(1/3)e$	r, g, b	0	$1/3$
$\nu, \bar{\nu}$	0	w	± 1	0
d	$(1/3)e$	c, m, y	0	$-1/3$
u	$(2/3)e$	r, g, b	0	$1/3$
$\bar{\text{e}}$	e	w	-1	0

Leptons, meaning light ones, comprise electrons and neutrinos. Baryons, meaning heavy ones, comprise quarks and all of the particles quarks make up. Lepton number and baryon number are important quantum numbers because they are conserved. That is, although it turns out that the number of electrons, for example, is not conserved in subatomic interactions, the total lepton number of the universe is conserved. In some sense this means that a neutrino and an electron are really the same particle, just in different states. Similarly, it means that an up quark and a down quark are in some sense just different states of the same particle. But leptons and baryons are distinct (as far as we know). A typical subatomic interaction which conserves lepton and baryon number is β decay:

$$\text{n} \rightarrow \text{p} + \text{e} + \bar{\nu}.$$

Since p=uud and n=ddu, what is really going on is

$$\text{d} \rightarrow \text{u} + \text{e} + \bar{\nu}.$$

Note that this also conserves electric charge and colour charge (the two quarks must be the same colour).

Table 3: The three families of fundamental particles

Family #1	Family # 2	Family # 3
electron e	muon μ	tauon τ
up quark u	strange quark s	top quark t
down quark d	charmed quark c	bottom quark b
e-neutrino ν_{e}	μ-neutrino ν_μ	τ-neutrino ν_τ

Counting all of the particles in the above table, including their possible colours, gives a total of 16. This already seems like quite a lot of elementary particles. However, nature is not satisfied with this. Instead, she has replicated the above structure two more times. The three groups of this structure are called families. The first family given above accounts for all of ordinary matter. The second family has higher mass and the third family higher still. Counting all three families gives a grand total of 48 elementary particles. These particles, with their often whimsical names, are summarized in table 3.

6.3 * Other Quanta

The list of elementary particles given above does not include photons for the reasons explained in Chapters 1 and 3: photons are quanta arising from quantizing the electromagnetic field, not particles that have to be quantized. The electromagnetic field is of course not the only force field in nature. Three others are known: the weak force, the strong force, and gravity. Gravity was of course the first force to be mathematically described. It is also the only one which has not been successfully quantized. Gravity is intimately connected with space and time, as Einstein showed. All other fields are quantized by *assuming* a certain space-time metric.

The strong and weak force fields can both be quantized in a similar manner to the electromagnetic force fields, and each gives rise to quanta which are bosons. However, they are much more complicated. Firstly, each field has a number of different quanta (three for the the weak force, eight for the strong force) compared to just one sort of electromagnetic quanta (photons). Secondly, the force-carrying quanta are massive, so that influences do not propagate at a constant speed. This is the reason that the strong and the weak forces are said to be finite range forces. The range of the force is roughly equal to the Compton wavelength $\lambda_C = \hbar/mc$ of the force-carrying "particle". (This can be shown by considering a radial, time-independent solution $\phi(r)$ of the Klein-Gordon equation.) Thirdly, the strong and weak fields couple to themselves. That is, there are terms in the Hamiltonians of these field theories which are nonlinear in the fields. The weak field also couples to the electromagnetic field because it carries charge.

Because of these complications we will not consider the details of the strong or weak forces. Table 4 summarizes the quanta of these forces (and electromagnetism), and how they couple to matter.

Table 4: The three known quantized forces of nature.

Force	Quanta	Symbols	Action
Electromag.	Photons	γ	couples to electric charge
Strong	Gluons	$g^1 \ldots g^8$	couples to colour charge
Weak	IVBs	W^+,W^-,Z^0	transmutes particles

Note that the action of the weak force, transmuting particles, is what causes interactions such as β decay. Such interactions take place rela-

tively rarely, which is why this is known as the weak force. The name of the force carrier, IVB, stands for intermediate vector boson.

The strong force is what holds the quarks together in nucleons. It also holds the nucleons together in a nucleus. Since nucleons are white (colour neutral), the bonding of nucleons is really just a residual strong force. That is, the strong nuclear force is analogous to the van-der-Waals forces of electromagnetic origin which hold together neutral atoms or molecules in condensed matter. The residual strong force in the nucleus of an atom is many orders of magnitude stronger than the electromagnetic force between the nucleus and the electrons. This gives an idea as to how strong the strong force really is.

In addition to these force-fields, there is believed to exist another fundamental field to complete the so-called *standard model* of particle physics. This is the Higgs field (or perhaps fields), named after British physicist Peter Higgs. This is quantized in the usual way and gives rise to a Higgs boson. The Higgs field is believed to couple to all particles and fields (except the electromagnetic field), giving all the particles and quanta their masses (see below). The Higgs boson has not yet been experimentally observed, as it is itself very massive, but the LHC (Large Hadron Collider) is hoping to see it.

6.4 * Renormalization revisited

Recall from Sec. 4.1.10 that if one goes beyond the first-order calculation (Fermi's golden rule), one finds an infinite energy shift for the atom. This calculation (which physicists first tried to perform in the 1930s) corresponds to a second-order term in Dyson's expansion. As explained in that earlier section, performing this calculation for an unbound electron also gives an infinite energy shift even to a free electron. it can be shown that this is equivalent to adding an infinite contribution Δm to the electron's mass, coming from the interaction with the continuum of electromagnetic vacuum modes. We deal with this by *mass renormalization* — we give the electron an infinite negative mass m_{bare}, called the *bare mass* to start with. This is chosen to cancel the infinite contribution to the energy of the unbound electron, leaving only the *physical* mass $m_{\text{physical}} = m_{\text{bare}} + \Delta m \approx 10^{-30}\text{kg}$.

Of course if $\Delta m \to \infty$ one cannot simply say $m_{\text{bare}} = -\infty + m_{\text{physical}}$ — that is a meaningless equation. Rather, one must deal with finite quantities the whole way, and then take the infinite limit at the end. This is a tricky business, and is known as *regularization*. For instance, one type of regularization is *dimensional regularization*, in which we assumes that space is $3 - \delta$ dimensional. For $\delta > 0$, the integrals in Dyson's expansion actually converge! We then take the limit $\delta \to 0$ at the end.

You might think that these infinities are an artefact of truncating the Dyson expansion, and that if we kept all orders that the infinites would cancel out and the result would be finite. In fact this is not the case. Each higher-order term in the Dyson expansion generates its own infinite terms, and they add up rather than cancel. All of these infinites are actually due to *loops* in the corresponding Feynman diagrams, which

are explained in terms of *virtual particles*, as opposed to the the "real" particles (photons included here) that enter or exit the interaction region. Thankfully, all of these loops generate the same sort of infinities, so they can all be cancelled by using the trick of mass renormalization.

There is one final comment that should be made about mass renormalization. It turns out that the standard model is only renormalizable if the bare masses of all the particles is zero. This sounds crazy, since I said above that renormalization with regularization sets the bare mass to a very large negative value. The trick is to get an *effective* bare mass which is nonzero (and in fact, very large and negative) by coupling the zero-mass particles to another massless field. This is the Higgs field (or fields). The Higgs field is special in that it interacts with itself in such a way that its lowest energy state is not the vacuum state $|0\rangle$ with no quanta, but rather something like a coherent state $|\alpha\rangle$, with a large number of quanta. This gives rise to an effective mass both for the Higgs field itself, and to all other fields or particles it interacts with. If the Higgs field did not exist, then the standard model could not explain why elementary particles have mass[6]. This is why the Higgs field is predicted to exist.

As it turns out, mass renormalization is not the whole story. Not only the mass of the electron (or whatever particle we are considering) must be renormalized, but so too must ϵ_0, which sets the scale of the electromagnetic field energy — see Eq. (6.1.8). And finally (for QED), so too must the coupling constant q (the charge) between the particle and the field. Like mass, the infinite contributions to these quantities come from Feynman loops, and every term in the Dyson expansion can be dealt with by using a regularization procedure with suitable values of $m_{\rm bare}$, $(\epsilon_0)_{\rm bare}$, and $q_{\rm bare}$.

The other interactions in the standard model can also be dealt with in a similar way — by renormalizing a finite number of parameters in the theory. This is what we mean by saying the theory is renormalizable. Other field theories are not renormalizable — each term in the Dyson expansion throws up a *different sort* of infinity, so one must introduce an infinite number of parameters in the theory in order to calculate anything.

These days, most physicists think of renormalization in the following way. It does not mean that the electron really has an infinite negative bare mass, for instance. Rather it means that on a small enough length scale (or high enough energy scale), quantum field theory breaks down. We may never know what theory applies at very small length scales — it may be a theory of strings of length $\sim 10^{-43}$m, or it may be something not yet dreamt of. But at the "everyday" length scales of high energy physics ($\sim 10^{-12}$m or greater) the details of the "theory of everything" will be invisible. What will *emerge* is an *effective field theory*, which, like the standard model, will necessarily be renormalizable[7]

[6]If an electron's bare mass was exactly zero, the corrections to its mass through couplings to the EM vacuum field is also zero, so its physical mass would remain zero.

[7]Gravity seems to be an exception. Einstein's field equations, when quantized, are non-renormalizable. On the other hand, gravity is extraordinarily weak compared to the other forces, about 10^{-42} times weaker. It may be an anomalous non-

6.5 The spin-statistics theorem

The spin-statistics theorem has been stated a number of times already, but now that we have completed our look at elementary particles and fields it is worth revisiting. That is because explaining the spin-statistics connection is one of the triumphs of relativistic quantum mechanics and quantum field theory. Another triumph is the prediction of antiparticles, and that is part of the spin-statistics theorem.

To reiterate, the spin-statistics connection is that

> All particles having half-integer spins are fermions, while all particles and quanta having integer spin are bosons.

This is an observed fact of nature. Assuming relativity and quantum mechanics to be correct it can be proven (and thus becomes a theorem). The rigorous proof was first constructed by Pauli, but we can understand it in three steps as follows.

1. Elementary particles are spin-half fermions. (See Chap. 5)
   (a) The only relativistically invariant object which can be interpreted as a particle wavefunction is the Dirac bispinor.
   (b) The particles described by the Dirac bispinor are spin-half.
   (c) The Dirac equation has anomalous (negative energy) solutions which would make all dynamics unstable unless particles were prevented from occupying these states.
   (d) The only way particles can be prevented from occupying negative energy states is if they are (almost) full already.
   (e) The anomalous states can only be full if the particles are fermions (obeying the Pauli exclusion principle).
2. Compound particles composed of elementary spin-half fermions obey the theorem (see Sec. 5.1.4)
   (a) A compound particle composed of an odd number of fermions is a fermion; one composed of an even number is a boson.
   (b) A compound particle composed of an odd number of fermions has half-integer spin; one composed of an even number has integer spin.
3. Elementary quanta for force-fields behave like integer spin bosons
   (a) Quantizing a classical field results in a description of quanta analogous to that of bose particles. (see Chap. 1).
   (b) The fundamental objects in relativity are integer-rank tensors (scalars, 4-vectors, 4×4-tensors etc.) and half-integer-rank tensors (spinors, spinors of vectors etc.)

renormalizable low-energy "relic" of the true "theory of everything", which we can see the effects of only because there is only one sort of gravitational "charge". That is, masses add up to give perceptible gravitational fields in the vicinity of stars and planets, unlike electrical and other charges which tend to cancel on macroscopic scales.

(c) A classical field must be an integer-rank tensor, because a classical spinor field would have an energy density which is not bounded below (see Sec. 5.6.2).

(d) A tensor field of rank r transforms in the same way as a particle wavefunction of spin r (see Sec. 5.1.4).