11 Canonical quantization of fields

11.1 The canonical quantization machine

You’ve got to have a system
Harry Hill (1964– )

Quantum field theory allows us to consider a Universe in which there exist different, yet indistinguishable, copies of elementary particles. These particles can be created and destroyed by interactions amongst themselves or with external entities. Quantum field theory allows us to describe such phenomena because particles themselves are simply excitations of quantum fields.

To see how particles emerge from fields we need to develop a way to quantize the classical fields we have looked at previously. The good news is that there is a machine available that takes a classical field theory and, after we turn the handle, it spits out a quantum field theory: that is, a theory where quantities are described in terms of the number of quantum particles in the system. The name of the machine is canonical quantization. In developing canonical quantization we’ll see that particles are added or removed from a system using field operators¹ and these are formed from the creation and annihilation operators we found so useful in previous chapters.

1¹Field operators were introduced in Chapter 4.

11.1 The canonical quantization machine

Canonical quantization is the turn-the-handle method of obtaining a quantum field theory from a classical field theory. The method runs like this:

- **Step I**: Write down a classical Lagrangian density in terms of fields. This is the creative part because there are lots of possible Lagrangians. After this step, everything else is automatic.
- **Step II**: Calculate the momentum density and work out the Hamiltonian density in terms of fields.
- **Step III**: Now treat the fields and momentum density as operators. Impose commutation relations on them to make them quantum mechanical.
- **Step IV**: Expand the fields in terms of creation/annihilation operators. This will allow us to use occupation numbers and stay sane.
• **Step V**: That’s it. Congratulations, you are now the proud owner of a working quantum field theory, provided you remember the normal ordering interpretation.

We’ll illustrate the method with one of the simplest field theories: the theory of the massive scalar field.

**Step I**: We write down a Lagrangian density for our theory. For massive scalar field theory this was given in eqn 7.7, which we rewrite as
\[
\mathcal{L} = \frac{1}{2} [\partial_\mu \phi(x)]^2 - \frac{1}{2} m^2 \phi(x)^2. \tag{11.1}
\]
[Recall that the equation of motion for this theory is the Klein–Gordon equation \((\partial^2 + m^2)\phi = 0\) leading to a dispersion \(E_p^2 = p^2 + m^2\).]

**Step II**: Find the momentum density (eqn 10.10) \(\Pi^\mu(x)\) given by
\[
\Pi^\mu(x) = \frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi(x))}. \tag{11.2}
\]
For our Lagrangian in eqn 11.1 this gives \(\Pi^\mu(x) = \partial_\mu \phi(x)\). The time-like component\(^2\) of this tensor is \(\Pi^0(x) = \pi(x) = \partial^0 \phi(x)\). This allows us to define the Hamiltonian density in terms of the momentum density
\[
\mathcal{H} = \Pi^0(x) \partial_0 \phi(x) - L, \tag{11.3}
\]
and using our Lagrangian\(^3\) leads to a Hamiltonian density
\[
\mathcal{H} = \partial^0 \phi(x) \partial_0 \phi(x) - \mathcal{L} = \frac{1}{2} [\partial_\mu \phi(x)]^2 + \frac{1}{2} [\nabla \phi(x)]^2 + \frac{1}{2} m^2 \phi(x)^2. \tag{11.4}
\]
The last line in eqn 11.5 tells us that the energy has contributions from
(i) a kinetic energy term reflecting changes in the configuration in time,
(ii) a ‘shear term’ giving an energy cost for spatial changes in the field and
(iii) a ‘mass’ term reflecting the potential energy cost of there being
a field in space at all. Taken together this has a reassuring look of
\(E = (\text{kinetic energy}) + (\text{potential energy}),\) which is what we expect from
a Hamiltonian.

**Step III**: We turn fields into field operators. That is to say, we make them operator-valued fields: one may insert a point in spacetime into such an object and obtain an operator. We therefore take \(\phi(x) \rightarrow \hat{\phi}(x)\) and \(\Pi^0(x) \rightarrow \hat{\Pi}^0(x)\). To make these field operators quantum mechanical we need to impose commutation relations between them. In single-particle quantum mechanics we have \([\hat{x}, \hat{p}] = i\hbar\). By analogy, we quantize the field theory by defining the equal-time commutator for the field operators
\[
[\hat{\phi}(t, x), \hat{\Pi}^0(t, y)] = i\hbar^{(3)}(x - y). \tag{11.6}
\]
As the name suggests, this applies at equal times only and otherwise the fields commute. We also have that \([\hat{\phi}(x), \hat{\phi}(y)] = [\hat{\Pi}^0(x), \hat{\Pi}^0(y)] = 0\) (and likewise for the daggered versions). Expressed in terms of these fields, the Hamiltonian density \(\mathcal{H}\) is now an operator \(\hat{\mathcal{H}}\) which acts on
state vectors. This is all well and good, except that we don’t know how operators like \( \hat{\phi}(x) \) act on occupation number states like \(|n_1n_2n_3\ldots\rangle\). What we do know is how creation and annihilation operators act on these vectors. If only we could build fields out of these operators!

**Step IV:** How do we get any further? The machinery of creation and annihilation operators is so attractive that we’d like to define everything in terms of them. In particular, we found a neat analogy between particles in momentum eigenstates and quanta in oscillators. What we do, therefore, is to expand the field operators in terms of the creation and annihilation operators. We’ve already seen this in the coupled oscillator problem (eqn 2.68) where we obtained a time-independent position operator of the form

\[
\hat{x}_j = \left(\frac{\hbar}{m}\right)^{\frac{1}{2}} \sum_k \frac{1}{(2\omega_k N)^{\frac{1}{2}}} (\hat{a}_k e^{i j k a} + \hat{a}_k^\dagger e^{-i j k a}).
\]

(11.7)

By analogy we write down a time-independent field operator for the continuous case which will look like

\[
\hat{\phi}(x) = \int \frac{d^3 p}{(2\pi)^{\frac{3}{2}}} \frac{1}{(2E_p)^{\frac{1}{2}}} (\hat{a}_p e^{ip \cdot x} + \hat{a}_p^\dagger e^{-ip \cdot x}),
\]

(11.8)

with \( E_p = + (p^2 + m^2)^{\frac{1}{2}} \) (that is, we consider positive roots only, a feature discussed below) and where, as before, our creation and annihilation operators have a commutation relation \([\hat{a}_p, \hat{a}_q^\dagger] = \delta^{(3)}(p - q)\).

The field operators that we’ve written down are intended to work in the Heisenberg picture. To obtain their time dependence, we hit them with time-evolution operators. We use the Heisenberg prescription for making an operator time dependent:

\[
\hat{\phi}(x) = \hat{\phi}(t, x) = \hat{U}^\dagger(t, 0)\hat{\phi}(x)\hat{U}(t, 0) = e^{iHt}\hat{\phi}(x)e^{-iHt}.
\]

(11.9)

The only part that the \( \hat{U}^\dagger(t, 0) = e^{-iHt} \) operators affect are the creation and annihilation operators, and we have

\[
\hat{U}^\dagger(t, 0)\hat{a}_p\hat{U}(t, 0) = e^{-iE_pt}\hat{a}_p,
\]

(11.10)

which shows that the \( \hat{a}_p \) picks up a \( e^{-iE_pt} \). Similarly, the \( \hat{a}_p^\dagger \) part will pick up a \( e^{iE_pt} \).

---

**Example 11.1**

To see this unfolding we’ll consider a simple example of one component of the field acting on an example state. The component we’ll consider is \( \hat{a}_q e^{iq \cdot x} \) and the example state is \( |n_p n_q n_r\rangle \). Taking it one step at a time:

\[
\begin{align*}
\hat{a}_q e^{iq \cdot x} e^{-iHt}\hat{a}_q^\dagger |n_p n_q n_r\rangle &= e^{iHt}\hat{a}_q |n_p n_q(n_q - 1)\rangle e^{iq \cdot x} e^{-i(n_p E_p + n_q E_q + n_r E_r)t} \\
&= \sqrt{n_q / n_p} e^{iHt} |(n_p n_q - 1)\rangle e^{iq \cdot x} e^{-i(n_p E_p + n_q E_q + n_r E_r)t} \\
&= \sqrt{n_q / n_p} e^{iHt} n_q |(n_q - 1)\rangle e^{iq \cdot x} e^{-i(n_p E_p + n_q E_q + n_r E_r)t} \\
&= \sqrt{n_q / n_p} n_q |(n_q - 1)\rangle e^{-iE_q t} e^{iq \cdot x}.
\end{align*}
\]

(11.11)
Part of what we’re left with, namely $\sqrt{n_{p}}(n_{q} - 1)n_{p}$, is exactly the same as if we’d just acted on the original state with $\hat{a}_{p}$. The effect of dynamicizing this operator has just been to multiply by a factor $e^{-ik_{a}q}$, so we conclude that the operator we seek is $\hat{a}_{p}e^{-i(\mathbf{k}a - \mathbf{q} \cdot \mathbf{x})} = \hat{a}_{p}e^{-i\mathbf{k} \cdot \mathbf{x}}$.

In summary, what we call the mode expansion of the scalar field is given by

$$\hat{\phi}(x) = \int \frac{d^{3}p}{(2\pi)^{3/2}} \frac{1}{(2E_{p})^{1/2}} \left( \hat{a}_{p}e^{-ip \cdot x} + \hat{a}_{p}^{\dagger}e^{ip \cdot x} \right),$$

with $E_{p} = +(\mathbf{p}^{2} + m^{2})^{1/2}$.

Note that by expanding out the position field we’ll get the momentum expansion for free, since for our scalar field example $\Pi^{\mu}(x) = \partial^{\mu}\phi(x)$. Also note that because the field in our classical Lagrangian is a real quantity, the field operator $\phi(x)$ should be Hermitian. By inspection $\phi^{\dagger}(x) = \phi(x)$, so this is indeed the case.

### 11.2 Normalizing factors

Before proceeding, we will justify the normalization factors in eqn 11.12. In evaluating integrals over momentum states we have the problem that $d^{3}p$ is not a Lorentz-invariant quantity. We can use $d^{4}p$ where $p = (p^{0}, \mathbf{p})$ is the four-momentum, but for a particle of mass $m$ then only values of the four-momentum which satisfy $\sum p^{2} = m^{2}$ need to be considered. This is known as the mass shell condition (see Fig. 11.1). Consequently we can write our integration measure

$$d^{4}p \delta(p^{2} - m^{2}) \theta(p^{0}).$$

We have included a Heaviside step function $\theta(p^{0})$ to select only positive mass particles.

---

**Example 11.2**

Show that $\delta(p^{2} - m^{2}) \theta(p^{0}) = \frac{1}{2E_{p}}[\delta(p^{0} - E_{p})\theta(p^{0})]$. We use the identity

$$\delta[f(x)] = \sum_{\{x \mid f(x) = 0\}} \frac{1}{|f'(x)|}\delta(x),$$

where the notation tells us that the sum is evaluated for those values of $x$ that make $f(x) = 0$. We take $x = p^{0}$ and $f(p^{0}) = p^{2} - m^{2} = (p^{0})^{2} - \mathbf{p}^{2} - m^{2}$. This gives us that $|f'(p^{0})| = 2|p^{0}|$ and we use the fact that the zeros of $f(p^{0})$ occur for $p^{0} = \pm(\mathbf{p}^{2} + m^{2})^{1/2} = \pm E_{p}$ to write

$$\delta(p^{2} - m^{2}) \theta(p^{0}) = \frac{1}{2E_{p}}[\delta(p^{0} - E_{p})\theta(p^{0}) + \delta(p^{0} + E_{p})\theta(p^{0})],$$

and so the result follows (since the second term in eqn 11.15 is zero).

---

4 This looks a little different from what we had in Chapter 4 when we introduced field operators. We will explain the reason for the difference in Section 11.5.

5 The condition $p^{2} = m^{2}$ means that $(p^{0})^{2} - \mathbf{p}^{2} = E_{p}^{2} - \mathbf{p}^{2} = m^{2}$.

---

**Fig. 11.1** The mass shell $p^{2} = m^{2}$ is a hyperboloid in four-momentum space. Also shown is the equivalent surface for light, $p^{2} = 0$. 
Thus we will write our Lorentz-invariant measure as

$$\frac{d^3 p}{(2\pi)^3 2E_p},$$

(11.16)

where the additional factor $\delta(p_0 - E_p)\theta(p_0)$ is there in every calculation and so we suppress it, and we have included the factor $1/(2\pi)^3$ because the mode expansion is essentially an inverse Fourier transform (we have one factor of $1/(2\pi)$ for every component of three-momentum). We are now in a position to write down integrals, for example:

$$1 = \int \frac{d^3 p}{(2\pi)^3 2E_p} |p\rangle \langle p|.$$  

(11.17)

This requires us to have Lorentz-covariant four-momentum states $|p\rangle$. We previously normalized momentum states according to

$$\langle p|q\rangle = \delta^{(3)}(p - q),$$

(11.18)

so our new four-momentum states $|p\rangle$ will need to be related to the three-momentum states $|p\rangle$ by

$$|p\rangle = (2\pi)^{3/2}(2E_p)^{1/2}|p\rangle,$$

(11.19)

and then their normalization can be written

$$\langle p|q\rangle = (2\pi)^3 2E_p \delta^{(3)}(p - q).$$

(11.20)

Similarly, to make creation operators $\hat{a}^\dagger$ appropriately normalized so that they create Lorentz-covariant states, we must define them by

$$\hat{a}^\dagger_p = (2\pi)^{3/2}(2E_p)^{1/2} \hat{a}^\dagger,$$

(11.21)

so that $\hat{a}^\dagger_p|0\rangle = |p\rangle$. In this case our mode expansion would take the form of a simple inverse Fourier transform using our Lorentz-invariant measure

$$\hat{\phi}(x) = \int \frac{d^3 p}{(2\pi)^3 (2E_p)} \frac{1}{(2E_p)^{1/2}} (\hat{a}_p e^{-ip\cdot x} + \hat{a}^\dagger_p e^{ip\cdot x}),$$

(11.22)

or writing in terms of $\hat{a}^\dagger$ and $\hat{a}_p$ rather than $\hat{a}^\dagger_p$ and $\hat{a}_p$:

$$\hat{\phi}(x) = \int \frac{d^3 p}{(2\pi)^{3/2} (2E_p)^{1/2}} (\hat{a}_p e^{-ip\cdot x} + \hat{a}^\dagger_p e^{ip\cdot x}),$$

(11.23)

which is identical to eqn 11.12.

### 11.3 What becomes of the Hamiltonian?

We can now substitute our expansion of the field operator $\hat{\phi}(x)$ into the Hamiltonian to complete our programme of canonical quantization. This will provide us with an expression for the energy operator in terms
of the creation and annihilation operators. The Hamiltonian is given by the volume integral of the Hamiltonian density,

\[ \hat{H} = \int d^3x \frac{1}{2} \left\{ \left[ \partial_0 \phi(x) \right]^2 + \left[ \mathbf{\nabla} \phi(x) \right]^2 + m^2 \phi(x)^2 \right\}. \]  \quad (11.24)

All we need do is substitute in the mode expansion and use the commutation relations to simplify. We’ll start by computing the momentum density \( \hat{\Pi}_\mu(x) = \partial_\mu \phi(x) \) which is given by

\[ \hat{\Pi}_\mu(x) = \partial_\mu \phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}(2E_p)^{1/2}} (-ip_\mu) \left( \hat{a}_p e^{-ip \cdot x} - \hat{a}_p^\dagger e^{ip \cdot x} \right). \]  \quad (11.25)

To obtain an expression for \( \partial_0 \phi(x) \), simply take the time-like component of \( \hat{\Pi}_\mu(x) \):

\[ \partial_0 \phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}(2E_p)^{1/2}} (-ip) \left( \hat{a}_p e^{-ip \cdot x} - \hat{a}_p^\dagger e^{ip \cdot x} \right), \]  \quad (11.26)

while the space-like components provide us with \( \mathbf{\nabla} \phi(x) \):

\[ \mathbf{\nabla} \phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}(2E_p)^{1/2}} (ip) \left( \hat{a}_p e^{-ip \cdot x} - \hat{a}_p^\dagger e^{ip \cdot x} \right). \]  \quad (11.27)

We now have all of the ingredients to calculate the Hamiltonian.

---

**Example 11.3**

Substitution of the mode expansion leads to a multiple integral of the form

\[ \hat{H} = \frac{1}{2} \int \frac{d^3x \: d^3p \: d^3q}{(2\pi)^3(2E_p)^{1/2}(2E_q)^{1/2}} \left[ (-E_p E_q - \mathbf{p} \cdot \mathbf{q}) \hat{a}_p \hat{a}_q e^{-i(\mathbf{p} - \mathbf{q}) \cdot \mathbf{x}} - \hat{a}_q \hat{a}_p^\dagger e^{i(\mathbf{p} - \mathbf{q}) \cdot \mathbf{x}} \right] \times \left[ \hat{a}_q e^{-iq \cdot x} - \hat{a}_q^\dagger e^{iq \cdot x} \right] + m^2 \left[ \hat{a}_p e^{-i\mathbf{p} \cdot \mathbf{x}} + \hat{a}_p^\dagger e^{i\mathbf{p} \cdot \mathbf{x}} \right] \left[ \hat{a}_q e^{-i\mathbf{q} \cdot \mathbf{x}} + \hat{a}_q^\dagger e^{i\mathbf{q} \cdot \mathbf{x}} \right]. \]  \quad (11.28)

We use our favourite trick again: first, we integrate over \( \mathbf{x} \) and use \( \int d^3x e^{i\mathbf{p} \cdot \mathbf{x}} = (2\pi)^3 \delta^{(3)}(\mathbf{p}) \) and obtain

\[ \hat{H} = \frac{1}{2} \int \frac{d^3p \: d^3q}{(2E_p)^{3/2}(2E_q)^{3/2}} \]  \quad (11.29)

\[ \times \left[ \delta^{(3)}(\mathbf{p} - \mathbf{q})(E_p E_q + \mathbf{p} \cdot \mathbf{q} + m^2)[\hat{a}_p \hat{a}_q e^{-i(E_p - E_q) t} + \hat{a}_q \hat{a}_p^\dagger e^{-i(E_p - E_q) t}] \right. \\
\left. + \delta^{(3)}(\mathbf{p} + \mathbf{q})(-E_p E_q - \mathbf{p} \cdot \mathbf{q} + m^2)[\hat{a}_p \hat{a}_q e^{-i(E_p + E_q) t} + \hat{a}_q \hat{a}_p^\dagger e^{-i(E_p + E_q) t}] \right]. \]

Then we do the integral over \( \mathbf{q} \) to mop up the delta functions:

\[ \hat{H} = \frac{1}{2} \int \frac{d^3p}{2E_p} \left[ (E_p^2 + \mathbf{p}^2 + m^2)(\hat{a}_p \hat{a}_p^\dagger + \hat{a}_p^\dagger \hat{a}_p) \right. \\
\left. + (-E_p^2 + \mathbf{p}^2 + m^2)(\hat{a}_p^\dagger \hat{a}_p e^{iE_p t} + \hat{a}_p \hat{a}_p^\dagger e^{-iE_p t}) \right]. \]  \quad (11.30)

and since \( E_p^2 = \mathbf{p}^2 + m^2 \) this quickly simplifies to

\[ \hat{H} = \frac{1}{2} \int d^3p E_p (\hat{a}_p \hat{a}_p^\dagger + \hat{a}_p^\dagger \hat{a}_p). \]  \quad (11.31)
Using the commutation relations on the result in eqn 11.31 we obtain
\[ E = \int d^3p \: E_p \left( \hat{a}^\dagger_p \hat{a}_p + \frac{1}{2} \delta^{(3)}(0) \right). \] (11.32)

The last term should fill us with dread. The term \( \frac{1}{2} \int d^3p \: \delta^3(0) \) will give us an infinite contribution to the energy! However we should keep in mind that by evaluating the total energy \( E \) we're asking the theory to tell us about something we can't measure. This is a nonsensical question and we have paid the price by getting a nonsensical answer. What we can measure is differences in energy between two configurations. This is fine in our formalism, since upon taking the differences between two energies the \( \frac{1}{2} \delta^{(3)}(0) \) factors will cancel obligingly. However, it's still rather unsatisfactory to have infinite terms hanging around in all of our equations. We need to tame this infinity.

### 11.4 Normal ordering

To get around the infinity encountered at the end of the last section we define the act of **normal ordering**. Given a set of free fields, we define the normal ordered product as
\[ N \left[ \hat{A} \hat{B} \hat{C}^\dagger \ldots \hat{X}^\dagger \hat{Y} \hat{Z} \right] = \left( \begin{array}{c} \text{Operators rearranged with all} \\ \text{creation operators on the left} \end{array} \right). \] (11.33)

The \( N \) is not an operator in the sense that is doesn't act on a state to provide some new information. Instead, normal ordering is an interpretation that we use to eliminate the meaningless infinities that occur in field theories. So the rule goes that 'if you want to tell someone about a string of operators in quantum field theory, you have to normal order them first'. If you don't, you're talking nonsense.\(^6\)

Rearranging operators is fine for Bose fields, but swapping the order of Fermi fields results in the expression picking up a minus sign. We therefore need to multiply by a factor \((-1)^P\), where \( P \) is the number of permutations needed to normally order a product of operators.

#### Example 11.4

Some examples of normal ordering a string of Bose operators follow. We first consider a single mode operator
\[ N \left[ \hat{a}^\dagger \hat{a} \right] = \hat{a}^\dagger \hat{a}, \quad N \left[ \hat{a} \right] = \hat{a}^\dagger \hat{a}, \] (11.34)
and
\[ N \left[ \hat{a}^\dagger \hat{a}^\dagger \hat{a} \right] = \hat{a}^\dagger \hat{a}^\dagger \hat{a}. \] (11.35)

Next, consider the case of many modes
\[ N \left[ \hat{a}_p \hat{a}_q \hat{a}_r \right] = \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r. \] (11.36)

The order of \( \hat{a}_p \) and \( \hat{a}_r \) doesn't matter since they commute.

For Fermi fields, \( N[\hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_r] = -\hat{c}_q^\dagger \hat{c}_p \hat{c}_r \), the sign change occurring because we have performed a single permutation (and hence an odd number of them).
Step V: We can finally complete the programme of canonical quantization. With the normal ordering interpretation the old expression

$$\hat{H} = \frac{1}{2} \int d^3 p \ E_p (\hat{a}_p \hat{a}_p^\dagger + \hat{a}_p^\dagger \hat{a}_p),$$  \hspace{2cm} (11.37)

becomes

$$N[\hat{H}] = \frac{1}{2} \int d^3 p \ E_p N[\hat{a}_p \hat{a}_p^\dagger + \hat{a}_p^\dagger \hat{a}_p]$$

$$= \frac{1}{2} \int d^3 p \ E_p 2\hat{a}_p^\dagger \hat{a}_p$$

$$= \int d^3 p \ E_p \hat{\rho}_p,$$  \hspace{2cm} (11.38)

where $\hat{\rho}_p = \hat{a}_p^\dagger \hat{a}_p$ is the number operator. Acting on a state it tells you how many excitations there are in that state with momentum $p$.

We now have a Hamiltonian operator that makes sense. It turns out that the Hamiltonian for the scalar field theory is exactly that which we obtained for independent particles in Chapter 3. This isn’t so surprising since we started with a Lagrangian describing waves that didn’t interact.

What we’ve seen though, is that the excited states of the wave equation can be thought of as particles possessing quantized momenta. These particles could be called scalar phions.\(^7\) They are Bose particles with spin\(^8\) $S = 0$.

---

Example 11.5

We have claimed that the vacuum energy is unobservable, so may be ignored. However, changes in vacuum energy terms are physically significant and lead to measurable effects. Such a change results if you adjust the boundary conditions for your field and this is the basis of the Casimir effect.\(^9\) This is a small, attractive force between two closely spaced metal plates, which results from the vacuum energy of the electromagnetic field. The essential physics can be understood using a toy model involving a massless scalar field in one dimension.

Consider two metal plates I and II separated by a distance $L$. We put a third plate (III) in between them, a distance $x$ from plate I (see Fig. 11.2). We will derive the force on plate III resulting from the field on either side of it. The presence of the plates forces the field to be quantized according to $k_n = n\pi / x$ or $n\pi / (L - x)$. The dispersion is $E_n = k_n$ and so the total zero-point energy is given by

$$E = \sum_{n=1}^{\infty} \left[ \frac{1}{2} \left( \frac{n\pi}{x} \right)^2 + \frac{1}{2} \left( \frac{n\pi}{L-x} \right)^2 \right] = f(x) + f(L-x),$$  \hspace{2cm} (11.39)

that is, $\frac{1}{2} \hbar \omega_n$ per mode.\(^11\) These sums both diverge, just as we expect since we are evaluating the infinite vacuum energy.

However, real plates can’t reflect radiation of arbitrarily high frequency: the highest energy modes leak out. To take account of this we cut off these high-energy modes thus:

$$\frac{n\pi}{2x} - \frac{n\pi}{2x},$$

which removes those modes with wavelength significantly smaller than the cut-off $\alpha$. The value of this cut-off is arbitrary, so we hope that it won’t feature in any measurable quantity.

---

7 This is because they are excitations of the scalar field $\phi$ (Greek letter ‘phi’).

8 We shall see that spin information is encoded by multiplying the creation and annihilation operators by objects that tell us about the particle’s spin polarization. These objects are vectors for $S = 1$ particles and spinors for a $S = 1/2$ particles. For $S = 0$ they do not feature.

9 The effect was predicted by the Dutch physicist Hendrik Casimir (1909–2000).

10 This approach follows Zee. A more detailed treatment may be found in Itzykson and Zuber, Chapter 3.

---

Fig. 11.2 Two metal plates (I and II) separated by a distance $L$, with a third plate (III) inserted in between.

11 Remember, for photons $\omega_n = c k_n$, and the zero-point energy is $\frac{1}{2} \hbar \omega_n$, so in units in which $\hbar = c = 1$ this becomes $\frac{1}{2} k_n$, and so $E = \sum_n \frac{1}{2} k_n$. Equation 11.39 contains two such sums.
The sums may now be evaluated:

\[
f(x) = \sum_n \frac{n\pi}{2x} e^{-n\pi a/x} \\
    = \frac{1}{2} \frac{\partial}{\partial a} \sum_n e^{-n\pi a/x} \\
    = \frac{1}{2} \frac{\partial}{\partial a} \frac{1}{1 - e^{-\pi a/x}} \\
    = \frac{\pi}{2x} \frac{e^{-\pi a/x}}{2x (1 - e^{-\pi a/x})^2} \approx \frac{x}{2\pi a^2} - \frac{1}{24x} + O(a^2).
\]

(11.41)

The total energy between plates I and II is

\[
E = f(x) + f(L - x) = \frac{L}{2\pi a^2} - \frac{\pi}{24} \left( \frac{1}{x} - \frac{1}{L - x} \right) + O(a^2).
\]

(11.42)

If \( x \) is small compared to \( L \), then we find a force

\[
F = -\frac{\partial E}{\partial x} = -\frac{\pi}{24x^2},
\]

(11.43)

which is independent of \( a \), as we hoped. Thus, there is an attractive force between the closely spaced plates I and III. This is the Casimir force. We can understand this force intuitively by realizing that as the two plates are pulled together we lose the high-energy modes. This reduces the energy between the plates and leads to an attractive force. A more quantum-field-theory friendly interpretation is that the effect results from quantum fluctuation in the vacuum, in which particles are spontaneously created and annihilated. These processes give rise to the vacuum Feynman diagrams described in Chapter 19.

11.5 The meaning of the mode expansion

In the next chapter we‘ll turn the crank on our canonical quantization machine for the second simplest field theory that we can imagine: that of the complex scalar field. Before doing that, let‘s have a closer look at the mode expansion for this theory. Our first guess for the field operator might have been the simple Fourier transform

\[
\hat{\phi}^-(x) = \int \frac{d^3p}{(2\pi)^\frac{3}{2}} \frac{1}{(2E_p)^\frac{1}{2}} \delta_p e^{-ip \cdot x},
\]

(11.44)

which looks like the one we had in Chapter 4, where the \( e^{-ip \cdot x} \) tells us that the particles are incoming. Unfortunately, this expansion won‘t work for the relativistic theory we‘re considering. The problem is the existence of negative energy solutions in the relativistic equations of motion: each momentum state gave rise to two energies \( E_p = \pm (p^2 + m^2) \) and we can‘t just leave out half of the solutions. Looking back at our discussion of the Klein–Gordon equation we saw the resolution of this problem, and we‘ll employ the same solution here.

What we need is a mode expansion that includes these negative energy modes

\[
\hat{\phi}(x) = \sum_p \left( \text{positive } E_p \right) + \sum_p \left( \text{negative } E_p \right).
\]

(11.45)
The expansion is carried out in terms of incoming plane waves $e^{-ip\cdot x}$. Now recall Feynman’s interpretation of the negative frequency modes in which negative energies are assumed to be meaningless, all energies are set to be positive and the signs of three-momenta for such modes are flipped. In this picture, the formerly negative energy states are interpreted as *outgoing* antiparticles. The positive energy modes continue to represent incoming particles. Incoming here means, not only a factor $e^{-ip\cdot x}$, but also that the particle is annihilated: it comes into the system and is absorbed by it. Conversely, outgoing means a factor $e^{ip\cdot x}$ and that the particle is created. We therefore interpret the mode expansion as

$$\phi(x) = \sum_p \left( \begin{array}{c} \text{incoming positive } E_p \\ \text{particle annihilated} \end{array} \right) + \sum_p \left( \begin{array}{c} \text{outgoing positive } E_p \\ \text{antiparticle created} \end{array} \right).$$

(11.46)

The resulting expansion of a field annihilation operator is then

$$\hat{\phi}(x) = \int \frac{d^3 p}{(2\pi)^{3/2}} \frac{1}{(2E_p)^{1/2}} \left( \hat{a}_p e^{-ip\cdot x} + \hat{b}^+_p e^{ip\cdot x} \right),$$

(11.47)

where the particle and antiparticle energies are given by $E_p = -(p^2 + m^2)^{1/2}$. The rules are that $\hat{a}_p$ annihilates particles, $\hat{a}^+_p$ creates particles and the operators $\hat{b}_p$ and $\hat{b}^+_p$ respectively destroy and create antiparticles. (For our scalar field theory the particles and antiparticles are the same: we say that ‘each particle is its own antiparticle’. This means that $\hat{a}_p$ can be thought of as either an operator that annihilates a particle or one that annihilates an antiparticle.)

As we will see, in other theories we will need separate operators to perform these two distinct roles, but in scalar field theory there is no difference. Thus in this case $\hat{a}_p = \hat{b}_p$, because in a scalar field theory a particle is its own antiparticle: particles and antiparticles are the same. This will not necessarily be the case for other field theories, an example of which will be described in the next chapter.

---

**Example 11.6**

Let’s see what happens if we hit the vacuum state with our field annihilation operator. Since $\hat{a}^+_p |0\rangle = |p\rangle$ we have

$$\hat{\phi}(x)|0\rangle = \int \frac{d^3 p}{(2\pi)^{3/2}} \frac{1}{(2E_p)^{1/2}} e^{ip\cdot x} |p\rangle.$$  

(11.48)

We have therefore created an outgoing superposition of momentum states. Pick out one of these states and make an amplitude by folding in a relativistically normalized state $\langle q | = (2\pi)^{3/2} (2E_p)^{1/2} |q\rangle$:

$$(2\pi)^{3/2} (2E_p)^{1/2} \langle q | \hat{\phi}(x)|0\rangle = \int d^3 p e^{ip\cdot x} \langle q | p\rangle = \int d^3 p e^{ip\cdot x} \delta^{(3)}(q - p) = e^{iq\cdot x}.$$  

(11.49)

In the language of second quantization, $e^{iq\cdot x}$ tells us how much amplitude there is in the $q$th momentum mode if we create a scalar particle at spacetime point $x$. Note that $e^{ip\cdot x} = e^{iE_p t - q\cdot x}$ and so

$$\int d^3 p e^{ip\cdot x} \delta^{(3)}(q - p) = e^{i(E_p t - q\cdot x)} = e^{iq\cdot x},$$

(11.50)

demonstrating once again that our integral over three-momentum coordinates nevertheless results in a single particle in a four-momentum state.
It is important to note that canonical quantization will not succeed in diagonalizing all field theories. Roughly speaking it works only for those Lagrangians which can be written as quadratic in a field and its derivatives.\textsuperscript{14} (We will revisit the mathematics of this point in Chapter 23.) The result of canonical quantization is a system described by single particles in momentum states which don’t interact with each other. For this reason Lagrangians that may be canonically quantized are called \textbf{non-interacting theories}. In contrast, those theories which cannot be diagonalized with canonical quantization are called \textbf{interacting theories}; these will be described in terms of single particles in momentum states which do interact with each other. Interacting theories are the subject of much of the rest of this book. For now we will continue to look at non-interacting theories and in the next chapter we examine some more uses of the canonical quantization routine that we have built.

**Chapter summary**

- Canonical quantization is an automated method for turning a classical field theory into a quantum field theory.
- Field operators are formed from mode expansions with amplitudes made of creation and annihilation operators.
- Normal ordering is needed to make sense of quantum field theories.
- The results of each step of the canonical quantization procedure for the scalar field theory are shown in the box in the margin.

**Exercises**

(11.1) One of the criteria we had for a successful theory of a scalar field was that the commutator for space-like separations would be zero. Let’s see if our scalar field has this feature. Show that

\[
|\hat{\phi}(x), \hat{\phi}(y)| = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} \left( e^{-ip \cdot (x-y)} - e^{-ip \cdot (y-x)} \right).
\]

(11.2) Show that, at equal times \( x^0 = y^0 \),

\[
[\hat{\phi}(x), \hat{\phi}(y)] = \frac{i}{2} \int \frac{d^3p}{(2\pi)^3} \left( e^{ip \cdot (x-y)} + e^{-ip \cdot (x-y)} \right).
\]

In this expression there’s nothing stopping us swapping the sign of \( p \) in the second term, and show that this leads to

\[
[\hat{\phi}(x), \hat{\phi}(y)] = i\delta(3)(x-y).
\]