Chapter 4

Quantization of Fields

Take some uniform medium that vibrates and in which waves can propagate, which could for example be a lattice of molecules. In such system, each normal mode of the vibration can be treated as an independent harmonic oscillator. Now suppose that it is small enough that quantum effects become important; then, the quantum mechanical treatment results in each oscillator containing an integer number of quanta corresponding to the wave length and oscillation frequency of the normal mode. These quanta behave in many ways as if they are particles, each carrying a definite energy and a momentum. The system is classically well-defined, and the procedure to obtain the quantum mechanical system is straightforward as we will discuss shortly. Quantum field theory of elementary particles takes such system as a *model* to describe particles where the wave length is interpreted as the momentum of each particle and the oscillation frequency as its energy. There are some important advantages to this approach. First, since it is based on a classical model, the energy of the system is well-defined by the total Hamiltonian of the system which is likely to be positive since the energy of the classical system would be positive even after quantized. Second, this will allow the theory to deal with multiple particles in a consistent way. So far, wave functions for the Klein-Gordon and the Dirac equations described a single particle; even in the hole theory, a pair creation was described as a transition of a single negative energy electron to a positive energy state. This resulted from the fact that there is only one pair of canonical observables $(\vec{x}, \vec{p})$ for such theories. In quantum field theory, there are infinite number of degrees of freedom corresponding to the position and momentum of each point of the field (or each *molecule*), and this allows the theory to describe infinite number of particles.

Let's look at how one quantizes a vibrating system. Imagine a one-dimensional string that can vibrate, where there will be a number $\phi(x)$ attached to each point x on the string. It could be the transverse or longitudinal displacement from the natural position, or anything else that can result in a wave. Classically, one would have a certain equation of motion that describes such wave, or equivalently, the Lagrangian of

the system. If the system is very small, the required quantum mechanical description of the system can be obtained from the Lagrangian of the classical system by the prescription called the *canonical quantization* which proceeds as follows. First, one chooses a set of generalized coordinates that uniquely describes the configuration of the string, $\underline{q} \equiv q_i$ $(i = 1, \ldots, n)$, and write down the Lagrangian of the system $L(\underline{q}, \underline{\dot{q}})$:

$$L(\underline{q}, \underline{\dot{q}}) = T - V \tag{4.1}$$

where T is the kinetic energy and V is the potential energy of the system. Then, the canonical momentum p_i conjugate to q_i is defined by

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} \tag{4.2}$$

and the Hamiltonian is formed as

$$H(\underline{p}, \underline{q}) \equiv \sum_i p_i \dot{q}_i - L \tag{4.3}$$

where the result is expressed as a function of $\underline{p}$ and $\underline{q}$. At this point, $q_i(t)$ and $p_i(t)$ are considered as hermitian operators (i.e. observables) in the Heisenberg picture where state vectors stay constant and all time dependences are carried by operators. Then, a set of *equal-time* commutation relations are introduced as

$$[q_i(t), p_j(t)] = i\delta_{ij}\,, \quad [q_i(t), q_j(t)] = [p_i(t), p_j(t)] = 0\,. \tag{4.4}$$

The time evolution of any observable $O(t)$ in the Heisenberg picture is given by Heisenberg's equation of motion:

$$-i\dot{O} = [H, O]\,, \tag{4.5}$$

and the matrix element of $O(t)$ at any given time for states $|\phi_{1,2}\rangle$ is then given by $\langle\phi_1|O(t)|\phi_2\rangle$, which allows the theory to be compared with measurements. Since changing the order of q_i and p_i could result in a different quantum mechanical H, this procedure is not unique; namely, there are in general more than one quantum mechanical system that has the same classical counterpart. When there is ambiguity, the choice should be made based on comparison with experiments and theoretical consistency. In practice, however, the choice is usually either quite obvious or ordering changes simply result in constant offsets that do not affect observable effects.

For the one-dimensional string, one obvious choice for the general coordinates is, for example, the transverse displacement $\phi(x)$ of each point labeled by x, or equivalently that of i-th molecule ϕ_i. An alternative choice is to describe the configuration of the string at a given time t by a superposition of normal modes:

$$\phi(x) = \sum_p c_p e^{ipx} + \text{complex conjugate} \tag{4.6}$$

which is nothing but the Fourier transform of $\phi(x)$. Each normal mode is a harmonic oscillator labeled by its wave length $2\pi/p$, and the oscillators are decoupled from each other; namely, each oscillator oscillates with a frequency uniquely given by the equation of motion without being perturbed by other oscillators. Thus, one can take the amplitudes of the normal modes as the general coordinates and proceed to quantize the system. As we will see later, the resulting quantized system becomes equivalent to the system quantized using the displacement $\phi(x)$ as the general coordinates.

So quantum field theory takes such quantized vibrating medium as a model to describe multiple particles where a creation of a particle with momentum p and energy E corresponds to an excitation of the normal mode with wave number p and angular frequency E by one quanta. It is just about the simplest way to describe multiple particles that can be created or annihilated, and it is quite amazing that such simple prescription actually works for a complicated system that consists of variety of elementary particles. Actually, in order for the model to work for the elementary particles, we need to address a few non-trivial questions:

1. Is the theory independent of the frame in which the quantization is performed?

2. Is the causality respected? Namely, are the measurements of the field at two points separated by a time-like distance independent of each other?

3. Electrons are known to obey Pauli's exclusion principle; how can it be incorporated into the theory?

The classical model described above is clearly inconsistent with special relativity since there is a spacial frame where the material is at rest. The move to a relativistic theory is similar to the case of the electromagnetic wave where the concept of the *ether* which transmits light was simply discarded and the electromagnetic fields and their transformation properties survived. After we make such transition, we will see that the answers to the item 1 and 2 are miraculously yes. Then, we will see that the third question is solved by using anticommutators rather than commutators to quantize the Dirac field which will limit the number of particles that occupy a single normal mode to one. Even after these issues are resolved, we will encounter further problems: how to formulate massless spin-1 particles (for example, photons), how to handle the infinities when calculating higher-order effects, etc. These topics will be dicussed in later chapters.

We will now start by briefly reviewing the quantum mechanics of a harmonic oscillator, and see how Pauli's exclusion principle can be incorporated using anticommutators in the quantization procedure.

4.1 Harmonic oscillator

Classical mechanics
Take a unit mass $m = 1$ attached to a spring with a spring constant ω^2. It is placed in a gravitational field and we take the vertical displacement upward from its natural position to be q. The Lagrangian, which is a function of q and $\dot{q}$, is

$$L(q, \dot{q}) = T - V = \frac{1}{2}(\dot{q}^2 - \omega^2 q^2)\,, \tag{4.7}$$

where the potential energy V is measured relative to the natural position. The Lagrange's equation of motion is then

$$\frac{d}{dt}\Big(\underbrace{\frac{\partial L}{\partial \dot{q}}}_{\dot{q}}\Big) = \underbrace{\frac{\partial L}{\partial q}}_{-\omega^2 q} \quad \to \quad \ddot{q} = -\omega^2 q \quad \text{(Lagrangian form)}\,, \tag{4.8}$$

which means $q(t) \propto e^{\pm i\omega t}$, and the general solution is (requiring that q be real)

$$q(t) = c\, e^{-i\omega t} + c^* e^{i\omega t} \tag{4.9}$$

where c is an arbitrary complex constant.

The canonical momentum is defined as

$$p \overset{\text{def}}{\equiv} \frac{\partial L}{\partial \dot{q}} = \dot{q}\,, \tag{4.10}$$

and the Hamiltonian is obtained by writing $p\dot{q} - L$ in terms of q and p only:

$$H(p, q) \overset{\text{def}}{\equiv} p\dot{q} - L = \dot{q}^2 - \frac{1}{2}(\dot{q}^2 - \omega^2 q^2) = \frac{1}{2}(\dot{q}^2 + \omega^2 q^2) \tag{4.11}$$

$$\to \quad H(p, q) = \frac{1}{2}(p^2 + \omega^2 q^2)\,. \tag{4.12}$$

Then Hamilton's equations of motions are

$$\begin{cases} \dot{q} \;=\; \dfrac{\partial H}{\partial p} \\ \dot{p} \;=\; -\dfrac{\partial H}{\partial q} \end{cases} \quad \to \quad \begin{cases} \dot{q} \;=\; p \\ \dot{p} \;=\; -\omega^2 q \end{cases} \quad \text{(Hamiltonian form)}\,, \tag{4.13}$$

where the first equation simply recovers the definition of p, which together with the second reproduces the equation of motion (4.8).

Quantization

When the system of the spring is very small then we have to use a quantum mechanical description of the system to correctly take into account quantum effects such as the quantization of energy levels. We will follow the procedure of canonical quantization to obtain such a description: namely, we regard $q(t)$ and $p(t)$ as hermitian operators in the Heisenberg picture, and impose the equal-time commutation relation

$$[q(t), p(t)] = i\,. \tag{4.14}$$

Then, the equations of motion are obtained by applying Heisenberg's equation of motion $-i\dot{O} = [H, O]$ (4.5) to $q(t)$ and $p(t)$. Using the relation $[AB, C] = A[B, C] + [A, C]B$,

$$\begin{aligned} -i\dot{q} &= [H, q\,] = \frac{1}{2}[p^2 + \omega^2 q^2, q\,] = \frac{1}{2}[p^2, q\,] \\ &= \frac{1}{2}\Big(p\underbrace{[p, q\,]}_{-i} + \underbrace{[p, q\,]}_{-i}p\Big) \\ &= -ip \qquad\qquad\qquad \rightarrow \quad \dot{q} = p\,, \end{aligned} \tag{4.15}$$

$$\begin{aligned} -i\dot{p} &= [H, p\,] = \frac{1}{2}[p^2 + \omega^2 q^2, p\,] = \frac{1}{2}[\omega^2 q^2, p\,] \\ &= \frac{\omega^2}{2}\Big(q\underbrace{[q, p\,]}_{i} + \underbrace{[q, p\,]}_{i}q\Big) \\ &= i\,\omega^2 q \qquad\qquad\qquad \rightarrow \quad \dot{p} = -\omega^2 q\,, \end{aligned} \tag{4.16}$$

thus reproducing the classical results which are now expressed as relations among operators.

A new feature created by moving to the quantized system is that the energy level, or the eigenvalue of H, can take only discrete values. The eigenvalue problem can be studied by introducing

$$\boxed{a(t) \stackrel{\text{def}}{\equiv} \frac{1}{\sqrt{2}}\Big(\sqrt{\omega}q + i\,\frac{p}{\sqrt{\omega}}\Big) \quad \rightarrow \quad a^\dagger(t) = \frac{1}{\sqrt{2}}\Big(\sqrt{\omega}q - i\,\frac{p}{\sqrt{\omega}}\Big)}\,. \tag{4.17}$$

Note that $a(t)$ is not hermitian, and thus it is not an observable. The commutator of $a(t)$ and $a^\dagger(t)$ is then

$$\begin{aligned} [a(t), a^\dagger(t)] &= \frac{1}{2}\Big[\sqrt{\omega}q + i\,\frac{p}{\sqrt{\omega}}, \sqrt{\omega}q - i\,\frac{p}{\sqrt{\omega}}\Big] \\ &= \frac{1}{2}\Big(i\underbrace{[p, q]}_{-i} - i\underbrace{[q, p]}_{i}\Big) = 1\,. \end{aligned} \tag{4.18}$$

Solving (4.17) for q and p,

$$q(t) = \frac{1}{\sqrt{2\omega}}\Big(a(t) + a^\dagger(t)\Big)\,, \quad p(t) = -i\sqrt{\frac{\omega}{2}}\Big(a(t) - a^\dagger(t)\Big)\,. \tag{4.19}$$

Using this, one sees that the commutation relation of $a(t)$ and $a^\dagger(t)$ (4.18) leads to that of q and p (4.14):

$$\begin{aligned}[q(t), p(t)] &= -\frac{i}{2}\,[\,a(t) + a^\dagger(t)\,,\, a(t) - a^\dagger(t)\,] \\ &= -\frac{i}{2}\Big(\underbrace{[a^\dagger(t), a(t)]}_{-1} - \underbrace{[a(t), a^\dagger(t)]}_{1}\Big) = i\,. \end{aligned} \tag{4.20}$$

Thus, the two commutation relations are equivalent.

The Hamiltonian can be written in terms of $a(t)$ and $a^\dagger(t)$ as follows:

$$\begin{aligned} a(t)a^\dagger(t) + a^\dagger(t)a(t) &= \frac{1}{2}\Big\{\Big(\sqrt{\omega}q + i\,\frac{p}{\sqrt{\omega}}\Big)\Big(\sqrt{\omega}q - i\,\frac{p}{\sqrt{\omega}}\Big) \\ &\qquad + \Big(\sqrt{\omega}q - i\,\frac{p}{\sqrt{\omega}}\Big)\Big(\sqrt{\omega}q + i\,\frac{p}{\sqrt{\omega}}\Big)\Big\} \\ &= \frac{1}{2}\Big(2\omega q^2 + 2\frac{p^2}{\omega}\Big) = \frac{1}{\omega}\,\underbrace{\Big(\omega^2 q^2 + p^2\Big)}_{2H}\,. \end{aligned} \tag{4.21}$$

Thus,

$$H = \frac{\omega}{2}\Big(\underbrace{a(t)a^\dagger(t)}_{1 + a^\dagger(t)a(t)\text{ by (4.18)}} + a^\dagger(t)a(t)\Big) = \omega\,\Big(a^\dagger(t)a(t) + \frac{1}{2}\Big)\,. \tag{4.22}$$

The time dependence of $a(t)$ is then given by

$$\begin{aligned} -i\dot{a}(t) &= [H, a(t)] = \omega[a^\dagger(t)a(t), a(t)] = \omega\,\underbrace{[a^\dagger(t), a(t)]}_{-1}\,a(t) \\ &\to\ \dot{a}(t) = -i\omega a(t)\,, \end{aligned} \tag{4.23}$$

which has the general solution

$$a(t) = ae^{-i\omega t} \quad \text{with} \quad a \stackrel{\text{def}}{\equiv} a(0)\,. \tag{4.24}$$

The a operator without explicit time dependence '(t)' is understood hereafter to be the value at $t = 0$. Then,

$$a(t)a^\dagger(t) = aa^\dagger\,, \quad a^\dagger(t)a(t) = a^\dagger a\,, \tag{4.25}$$

and thus,

$$\boxed{[a, a^\dagger] = 1}\,, \tag{4.26}$$

and

$$\boxed{H = \omega\Big(N + \frac{1}{2}\Big)\,, \quad N \stackrel{\text{def}}{\equiv} a^\dagger a}\,, \tag{4.27}$$

where the operator N is called the *number operator* for the reason we will see below.

We first evaluate $[a^\dagger, N]$ and $[a, N]$:

$$\begin{aligned} \underbrace{[a^\dagger, N]}_{a^\dagger N - N a^\dagger} &= [a^\dagger, a^\dagger a] = a^\dagger \underbrace{[a^\dagger, a]}_{-1} = -a^\dagger \;\;\rightarrow\;\; N a^\dagger = a^\dagger (N+1)\,, \\ \underbrace{[a, N]}_{aN - Na} &= [a, a^\dagger a] = \underbrace{[a, a^\dagger]}_{1} a = a \qquad \rightarrow \quad N a = a(N-1)\,. \end{aligned} \tag{4.28}$$

Let $|n\rangle$ be an eigenstate of N with an eigenvalue n:

$$N|n\rangle = n|n\rangle\,. \tag{4.29}$$

where n is real since N is hermitian: $N^\dagger = (a^\dagger a)^\dagger = a^\dagger a = N$. Applying (4.28) to $|n\rangle$,

$$\begin{aligned} N a^\dagger |n\rangle &= a^\dagger \underbrace{(N+1)}_{n+1} |n\rangle \;\;\rightarrow\;\; N\Big(a^\dagger|n\rangle\Big) = (n+1)\Big(a^\dagger|n\rangle\Big) \\ N a |n\rangle &= a \underbrace{(N-1)}_{n-1} |n\rangle \quad\rightarrow\quad N\Big(a|n\rangle\Big) = (n-1)\Big(a|n\rangle\Big)\,. \end{aligned} \tag{4.30}$$

Namely, $a^\dagger$ raises the eigenvalue by 1 and a lowers eigenvalue by 1, which then translates to raising or lowering the energy, or the eigenvalue of H written as (4.27), by one unit of ω. Accordingly, $a^\dagger$ is called the *creation operator*, and a the *annihilation operator*.

That $n \geq 0$ can be seen as follows:

$$\underbrace{\langle n|N|n\rangle}_{n\underbrace{\langle n|n\rangle}_{>0}} = \underbrace{\langle n|a^\dagger a|n\rangle}_{(\text{norm of } a|n\rangle) \geq 0} \quad\rightarrow\quad n \geq 0\,. \tag{4.31}$$

If n is not an integer, then the state $a^k|n\rangle$ with $k > n$ will have an eigenvalue $n-k < 0$ which contradicts the fact that n be non-negative. This can be avoided if at some point applying a results in the null state (the state with zero norm). If n is an integer, then indeed such is the case since $a|0\rangle$ is a state with zero norm as can be seen by setting $n = 0$ in (4.31):

$$\langle 0|a^\dagger a|0\rangle = 0 \quad\rightarrow\quad a|0\rangle = 0\,. \tag{4.32}$$

Thus, $n = 0$ corresponds to the state with lowest energy which we normalize as

$$\langle 0|0\rangle = 1\,. \tag{4.33}$$

All eigenstates of N can then be obtained by applying $a^\dagger$ to the ground state $|0\rangle$:

$$|n\rangle \stackrel{\text{def}}{\equiv} \frac{1}{\sqrt{n\,!}} a^{\dagger\,n}|0\rangle\,, \tag{4.34}$$

where $1/n\,!$ is a normalization factor. It can be readily shown that this state is indeed properly normalized:

$$\langle n|n\rangle = 1\,, \tag{4.35}$$

and that the operations of $a^\dagger$ and a on $|n\rangle$ are given by

$$\begin{cases} a^\dagger|n\rangle &= \sqrt{n+1}\,|n+1\rangle \\ a\,|n\rangle &= \sqrt{n}\,|n-1\rangle \end{cases}\,. \tag{4.36}$$

Exercise 4.1 *Eigenstates of harmonic oscilator.*
(*a*) *Show that the state* $|n\rangle$ *as defined in (4.34) is normalized to unity; i.e.* $\langle n|n\rangle = 1$.
(*b*) *Derive (4.36).*

From (4.19) and the time dependence of $a(t)$ (4.24), $q(t)$ can be written as

$$\boxed{q(t) = \frac{1}{\sqrt{2\omega}}\Big(ae^{-i\omega t} + a^\dagger e^{i\omega t}\Big)}\,. \tag{4.37}$$

Comparing this with the classical solution (4.9), one can see that the same quantum mechanical system could be obtained by the following procedure: First, write down the general classical solution as (4.9), identify the coefficient of $e^{-i\omega t}$ as $a/\sqrt{2\omega}$ and then introduce the commutation relation $[a, a^\dagger] = 1$. How did the factor $1/\sqrt{2\omega}$ in (4.37) come about? It originated from the definition (4.17). If $a(t)$ is defined as $(cq + ip/c)/\sqrt{2}$ where c is an arbitrary real constant, then the same derivation as (4.18) shows that $[a(t), a^\dagger(t)] = 1$ still results from $[q, p] = i$, or equivalently, $a^\dagger(t)a(t)$ is still the number operator with integer eigenvalues. However, the requirement that the Hamiltonian be a function only of $N = a^\dagger a$ leads to $c^2 = \omega$, which then uniquely fixes the definition (4.17) up to an overall sign.

It would be worthwhile at this point to emphasize an important feature of the Heisenberg picture. We have a set of basis states $|i\rangle$ $(i = 0, 1, 2, \ldots)$ which forms a complete set in the Hilbert space that represents a single oscillator. In the Heisenberg picture, these states do not vary with time. The time evolution corresponding to any of these states, or any superpositions thereof, are contained in the set of time-dependent observables such as $q(t)$, $p(t)$, etc. It is important to note that the same set

of operators with the same time dependences takes care of any states. For example, one can form a wave packet by a linear combination of the basis states where the probability to find the mass is localized as a bump near a certain position at a given time. Such a bump would follow a motion approximating a classical oscillation. In the Heisenberg picture, such motion is representated by a stationary state in the Hilbert space, and different motions (for example, different amplitudes of the oscillation) are represented by different static states. The operators and their time dependences are identical for all these states.

Fermionic oscillator

The quantized harmonic oscillator we have just studied will serve as a normal mode of a field corresponding to a given momentum where the number of quanta of the oscillator is identified as the number of particles with that momentum. It works fine for particles with a integer spin (spin 0, 1, ...: called *bosons*) for which a given state can be occupied by any number of such particles. However, particles with a half-integer spin (1/2, 3/2, ...: called *fermions*) cannot occupy an already occupied state, and some modification is needed to limit the number of quanta for a given oscillator to one. This can be accomplished by replacing the commutators of a and $a^\dagger$ by anticommutators:

$$\begin{aligned} & [a, a^\dagger] = 1, \quad [a, a] = [a^\dagger, a^\dagger] = 0 \\ \rightarrow \quad & \boxed{\{a, a^\dagger\} = 1, \quad \{a, a\} = \{a^\dagger, a^\dagger\} = 0} \,. \end{aligned} \tag{4.38}$$

The anticommutation relations $\{a, a\} = \{a^\dagger, a^\dagger\} = 0$ mean that the square of a or $a^\dagger$ is zero:

$$a^2 = 0\,, \quad a^{\dagger 2} = 0\,. \tag{4.39}$$

Define the number operator as before: $N \equiv a^\dagger a$, then,

$$\begin{aligned} N^2 &= a^\dagger a \underbrace{a^\dagger a}_{1 - a a^\dagger} = a^\dagger a - a^\dagger \underbrace{a a}_{0} a^\dagger = N \\ &\rightarrow \; N(N-1) = 0 \end{aligned} \tag{4.40}$$

which means that N has eigenvalues 0 and 1 as promised. Let $|0\rangle$ be the eigenstate with eigenvalue 0 which is assumed to be normalized as

$$\langle 0|0\rangle = 1\,. \tag{4.41}$$

Using the anticommutation relation (4.38), we have

$$N a^\dagger = \underbrace{a^\dagger a}_{1 - a a^\dagger} a^\dagger = a^\dagger - a \underbrace{a^{\dagger 2}}_{0} = a^\dagger\,. \tag{4.42}$$

Applying this to $|0\rangle$, we get

$$Na^\dagger|0\rangle = a^\dagger|0\rangle\,, \tag{4.43}$$

which indicates the identification,

$$|1\rangle = a^\dagger|0\rangle\,, \tag{4.44}$$

if it is not a null state. In fact, its norm is unity:

$$\langle 1|1\rangle = \langle 0|\underbrace{a\,a^\dagger}_{1-a^\dagger a}|0\rangle = \langle 0|(1-\underbrace{N}_{0})|0\rangle = \langle 0|0\rangle = 1\,. \tag{4.45}$$

Applying a to (4.44),

$$\begin{aligned} a\,|1\rangle &= \underbrace{aa^\dagger}_{1-a^\dagger a}\,|0\rangle = (1-\underbrace{N}_{0})|0\rangle \\ &\rightarrow\ |0\rangle = a\,|1\rangle\,. \end{aligned} \tag{4.46}$$

Equations (4.44) and (4.46) show that a and $a^\dagger$ are acting as an annihilation operator and a creation operator, respectively. If we try to lower the eigenvalue of $|0\rangle$ or raise the eigenvalue of $|1\rangle$, then the state vanishes:

$$\left\{\begin{array}{rcll} a^\dagger|1\rangle &=& a^{\dagger 2}|0\rangle &= 0 \\ a\,|0\rangle &=& a^2|1\rangle &= 0 \end{array}\right. . \tag{4.47}$$

The Hamiltonian should be proportional to the number operator up to a constant offset which may be ignored:

$$H = \omega a^\dagger a\,, \tag{4.48}$$

where the proportionality coefficient ω is identified as the energy of one quanta. Heisenberg's equation of motion (using a commutator) is assumed to be still valid; then, the time dependence of a is given by

$$\begin{aligned} -i\dot{a} &= [H,a\,] = \omega[a^\dagger a, a] = \omega(a^\dagger \underbrace{a\,a}_{0} - \underbrace{aa^\dagger}_{1-a^\dagger a}\,a) = \omega(-a + a^\dagger \underbrace{a\,a}_{0}) = -\omega a \\ &\rightarrow \dot{a} = -i\omega a\,. \end{aligned} \tag{4.49}$$

Thus, a has the same time dependence as in the boson case:

$$a(t) = ae^{-i\omega t}\,, \tag{4.50}$$

where a without explicit time dependence '(t)', in this equation and those we have been using, is simply understood to be the value of the operator at $t = 0$.

Thus, an oscillator that satisfies Pauli's exclusion principle can be constructed using anticommutators for the annihilation and creation operators a and $a^\dagger$. Such an oscillator is sometimes called a fermionic oscillator. Before leaving the topic, we note that the above formulation of the fermionic oscillator is symmetric under the simultaneous exchanges $|0\rangle \leftrightarrow |1\rangle$ and $a \leftrightarrow a^\dagger$. Namely, if we define

$$\begin{cases} a' \equiv a^\dagger \\ a'^\dagger = a \end{cases}, \quad \begin{cases} |0\rangle' \equiv |1\rangle \\ |1\rangle' \equiv |0\rangle \end{cases}, \tag{4.51}$$

then, a' and $a'^\dagger$ satisfy the same anticommutation relations (4.38) and the raising (4.44) and lowering (4.46) relations:

$$\{a', a'^\dagger\} = 1, \quad \{a', a'\} = \{a'^\dagger, a'^\dagger\} = 0\,,$$

$$a'^\dagger|0\rangle' = |1\rangle'\,, \quad a'|1\rangle' = |0\rangle'\,.$$

It is also easily verified that

$$a'^\dagger a'|i\rangle' = i\,|i\rangle' \quad (i = 0, 1)\,; \tag{4.52}$$

namely, the number operator is still $N' = a'^\dagger a'$. Thus, it is arbitrary which is the occupied state and which is the empty state at this point. Later, we will define the state with lower energy to be $|0\rangle$.

4.2 Lagrangian formulation of classical wave

As stated earlier, the first step toward quantizing a field is to find the Lagrangian of the system, then we can proceed to derive the canonical momenta from it, form the Hamiltonian, and then introduce commutation relations between the coordinates and momenta to quantize the system. Thus, let's first find the Lagrangian of a simple vibrating material. At this stage, it has nothing to do with quantum mechanics; it is just old classical mechanics.

Lagrangian density and the Euler-Lagrange equation
Consider a string of masses (m each) as shown in Figure 4.1. Each mass is attached to a spring with a spring constant k. Each mass is constrained to move vertically, and the displacement of i-th mass from the natural position is denoted as ϕ_i, which can also be labeled by the position along the string

$$\phi_i \equiv \phi(x_i)\,. \tag{4.53}$$

In order for waves to propagate, there should be some kind of coupling between the masses which is provided by a rubber band stretched with a constant tension τ.

The tension is assumed to be independent of the vertical displacements or the slope thereof which are assumed to be small. All these assumptions seem rather artificial; the only essential elements, however, are that the string has some uniform mass and that the potential has one term that is proportional to ϕ^2 and another term that is proportional to $(\partial\phi/\partial x)^2$, which is the case for our system as we will see shortly. When measured relative to the minimum potential (the natural state), these terms are the lowest order terms when the potential is expanded in terms of ϕ and $\partial\phi/\partial x$. The higher-order terms can be separately treated as self interactions of the field, and ignored at this point.

Assuming that θ_i's are small, the total potential energy of the rubber band (relative to the natural state) is

$$\begin{aligned} V_{\text{rubber}} &= \sum_i \tau \underbrace{\Delta x \Big(\overbrace{\frac{1}{\cos\theta_i} - 1}^{\theta_i^2/2} \Big)}_{\text{rubber stretch length}} = \sum_i \tau \Delta x \frac{\theta_i^2}{2} \\ &= \sum_i \tau \Delta x \frac{1}{2} \Big(\frac{\phi_i - \phi_{i-1}}{\Delta x} \Big)^2 , \end{aligned} \tag{4.54}$$

and that of the springs is

$$V_{\text{spring}} = \sum_i \frac{1}{2} k \phi_i^2 \, . \tag{4.55}$$

The Lagrangian of the system is then

$$\begin{aligned} L(\underline{\phi}, \underline{\dot{\phi}}) &= T - V_{\text{spring}} - V_{\text{rubber}} \\ &= \sum_i \Big[\frac{1}{2} m \dot{\phi}_i^2 - \frac{1}{2} k \phi_i^2 - \frac{1}{2} \tau \Delta x \Big(\frac{\phi_i - \phi_{i-1}}{\Delta x} \Big)^2 \Big] \end{aligned} \tag{4.56}$$

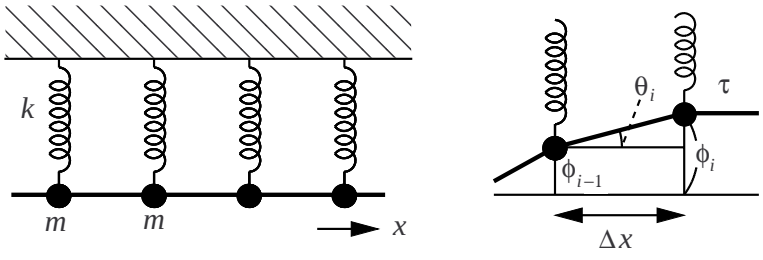

Figure 4.1: A string of springs, each with mass m and spring constant k. The masses are connected by a rubber band of tension τ.

with

$$\underline{\phi} \equiv (\phi_1, \ldots, \phi_n)\,. \tag{4.57}$$

In terms of the mass per unit length μ and the spring constant per unit length κ, m and k can be written as

$$m = \mu\Delta x\,, \quad k = \kappa\Delta x\,. \tag{4.58}$$

The Lagrangian is then

$$L(\underline{\phi}, \dot{\underline{\phi}}) = \sum_i \Delta x\Big[\frac{\mu}{2}\dot{\phi}_i^2 - \frac{\kappa}{2}\phi_i^2 - \frac{\tau}{2}\Big(\frac{\phi_i - \phi_{i-1}}{\Delta x}\Big)^2\Big]\,. \tag{4.59}$$

Lagrange's equation of motion is

$$\begin{aligned} \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{\phi}_i}\Big) &= \frac{\partial L}{\partial \phi_i} \quad (i = 1, \ldots, n) \\ \rightarrow \quad \mu\ddot{\phi}_i &= -\kappa\phi_i - \tau\frac{(\phi_i - \phi_{i-1}) - (\phi_{i+1} - \phi_i)}{(\Delta x)^2}\,, \end{aligned} \tag{4.60}$$

where the tension term in $\partial L/\partial\phi_i$ resulted from two terms corresponding to i and $i+1$. This is a set of n differential equations that are coupled. Let's take the limit $n \rightarrow \infty$ and $\Delta x \rightarrow 0$ while keeping μ and κ constant. Noting that in this limit we have

$$\frac{(\phi_i - \phi_{i-1})}{\Delta x} = \frac{\partial\phi}{\partial x}(x_i) \tag{4.61}$$

the tension term in (4.60) can be written as

$$\begin{aligned} \frac{(\phi_i - \phi_{i-1}) - (\phi_{i+1} - \phi_i)}{(\Delta x)^2} &= \frac{\dfrac{(\phi_i - \phi_{i-1})}{\Delta x} - \dfrac{(\phi_{i+1} - \phi_i)}{\Delta x}}{\Delta x} \\ &= \frac{\dfrac{\partial\phi}{\partial x}(x_i) - \dfrac{\partial\phi}{\partial x}(x_{i+1})}{\Delta x} = -\frac{\partial^2\phi}{\partial x^2}(x_i)\,, \end{aligned} \tag{4.62}$$

and the equation of motion is now written in a differential form:

$$\mu\ddot{\phi}(x) = -\kappa\phi(x) + \tau\frac{\partial^2\phi}{\partial x^2}(x)\,. \tag{4.63}$$

In the limit $n \rightarrow \infty$, the Lagrangian (4.59) can be written as an integral over x. Using (4.61),

$$L(\underline{\phi}, \dot{\underline{\phi}}) = \int dx\,\mathcal{L}\Big(\phi, \dot{\phi}, \frac{\partial\phi}{\partial x}\Big) \tag{4.64}$$

with

$$\mathcal{L}\Big(\phi, \dot{\phi}, \frac{\partial\phi}{\partial x}\Big) = \frac{\mu}{2}\dot{\phi}^2 - \frac{\kappa}{2}\phi^2 - \frac{\tau}{2}\Big(\frac{\partial\phi}{\partial x}\Big)^2\,. \tag{4.65}$$

The range of integration is either $(-\infty, \infty)$ or $(-L/2, L/2)$ where L is some large length. If we take the range to be infinity, then we assume that the function $\phi(x)$ vanishes at sufficiently large distances from origin. If we take the range to be finite, we impose periodicity condition:

$$\phi(x + L) = \phi(x)\,. \tag{4.66}$$

These conditions are required so that the boundary values do not contribute in the partial integrations that occur in what follows. The integrand $\mathcal{L}(\phi, \dot{\phi}, \partial\phi/\partial x)$ is a function of the field value $\phi(x)$, its time derivative $\dot{\phi}(x)$ and of the spacial derivative $\partial\phi/\partial x(x)$, and called the *Lagrangian density*.

The equation of motion (4.63) was obtained by taking the limit $\Delta x \to 0$ of the equation of motion for the discrete positions. It is convenient to establish a formula which gives us the continuous equation of motion directly from the Lagrangian density. To do so, we start from the action principle itself. The action S is the time integral of the Lagrangian from time t_1 to t_2 which are the start and end of the motion of interest:

$$S \equiv \int_{t_1}^{t_2} L\, dt = \int_{t_1}^{t_2} dt \int dx\, \mathcal{L}\Big(\phi, \dot{\phi}, \frac{\partial\phi}{\partial x}\Big)\,. \tag{4.67}$$

The function $\phi(t, x)$ specifies the motion of the string which may or may not be a true realizable motion, and for each motion, there is *a real number* S associated with it as defined by (4.67). The action S is thus a mapping of a function $\phi(t, x)$ to a number; such mapping is called a *functional*. The action principle then tells us that the true motion is the one that has the smallest S. Or equivalently, when we vary the function slightly around the true motion,

$$\phi'(t, x) = \phi(t, x) + \delta\phi(t, x) \tag{4.68}$$

with the variation set to zero at the start and end of the time window,

$$\delta\phi(t_1, x) = \delta\phi(t_2, x) = 0\,, \tag{4.69}$$

then the change in S should be zero to the first order in $\delta\phi$:

$$\delta S \equiv S' - S = 0\,. \tag{4.70}$$

The change in S can be obtained by integrating $\delta\mathcal{L} \equiv \mathcal{L}' - \mathcal{L}$ over t and x:

$$\delta S = \int dt dx (\mathcal{L}' - \mathcal{L}) = \int dt dx\, \delta\mathcal{L} = 0\,, \tag{4.71}$$

where the range of time integration is understood to be (t_1, t_2). In turn, the change in $\mathcal{L}$ at given time and position (t, x) occurs through the changes in ϕ, $\dot{\phi}$, and $\partial\phi/\partial x$, each evaluated at the point (t, x):

$$\delta\mathcal{L}(t, x) = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\dot{\phi}}\delta(\dot{\phi}) + \frac{\partial\mathcal{L}}{\partial\big(\frac{\partial\phi}{\partial x}\big)}\delta\Big(\frac{\partial\phi}{\partial x}\Big)\,. \tag{4.72}$$

Since $\phi' = \phi + \delta\phi$,

$$\delta(\dot{\phi}) \equiv \dot{\phi}' - \dot{\phi} = (\dot{\delta\phi}), \qquad \delta\Big(\frac{\partial\phi}{\partial x}\Big) \equiv \frac{\partial\phi'}{\partial x} - \frac{\partial\phi}{\partial x} = \frac{\partial}{\partial x}(\delta\phi)\,, \tag{4.73}$$

then using the chain rule $A(\frac{\partial}{\partial s}B) = \frac{\partial}{\partial s}(AB) - (\frac{\partial}{\partial s}A)B$ with $s = t$ or x,

$$\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\underbrace{\delta(\dot{\phi})}_{(\dot{\delta\phi})} = \underbrace{\frac{\partial}{\partial t}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\delta\phi\Big)}_{\to 0} - \Big(\frac{\partial}{\partial t}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\Big)\Big)\delta\phi \tag{4.74}$$

$$\frac{\partial\mathcal{L}}{\partial\Big(\frac{\partial\phi}{\partial x}\Big)}\underbrace{\delta\Big(\frac{\partial\phi}{\partial x}\Big)}_{\frac{\partial}{\partial x}(\delta\phi)} = \underbrace{\frac{\partial}{\partial x}\Big(\frac{\partial\mathcal{L}}{\partial\Big(\frac{\partial\phi}{\partial x}\Big)}\delta\phi\Big)}_{\to 0} - \Big(\frac{\partial}{\partial x}\Big(\frac{\partial\mathcal{L}}{\partial\Big(\frac{\partial\phi}{\partial x}\Big)}\Big)\Big)\delta\phi \tag{4.75}$$

where the first term in the right hand side of (4.74) or (4.75) vanishes upon integration over t or x, respectively, due to the boundary conditions (4.69) and (4.66). Thus, δS can now be written as

$$\delta S = \int dtdx\Big[\frac{\partial\mathcal{L}}{\partial\phi} - \frac{\partial}{\partial t}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\Big) - \frac{\partial}{\partial x}\Big(\frac{\partial\mathcal{L}}{\partial\Big(\frac{\partial\phi}{\partial x}\Big)}\Big)\Big]\delta\phi = 0\,. \tag{4.76}$$

Since δS should vanish for any variation $\delta\phi$, the integrand $[\cdots]$ should be zero at all (t, x):

$$\frac{\partial}{\partial t}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\Big) + \frac{\partial}{\partial x}\Big(\frac{\partial\mathcal{L}}{\partial\Big(\frac{\partial\phi}{\partial x}\Big)}\Big) = \frac{\partial\mathcal{L}}{\partial\phi}\,, \tag{4.77}$$

which is called the *Euler-Lagrange equation* for the Lagrangian density. Applying this to the Lagrangian density (4.65), we obtain an equation of motion

$$\frac{\partial}{\partial t}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\Big) = \mu\ddot{\phi}\,, \quad \frac{\partial}{\partial x}\Big(\frac{\partial\mathcal{L}}{\partial\Big(\frac{\partial\phi}{\partial x}\Big)}\Big) = -\tau\frac{\partial^2\phi}{\partial x^2}\,, \quad \frac{\partial\mathcal{L}}{\partial\phi} = -\kappa\phi$$

$$\to \quad \mu\ddot{\phi} - \tau\frac{\partial^2\phi}{\partial x^2} = -\kappa\phi\,, \tag{4.78}$$

which is exactly the same equation as (4.63) which was obtained directly from $L(\underline{\phi}, \underline{\dot{\phi}})$ in the limit $\Delta x \to 0$.

When the field is labeled by a continuous parameter as in $\phi(x)$, the total Lagrangian is not convenient to deal with. This is because Lagrange's equation of motion for the total Lagrangian is written in terms of each discrete coordinate ϕ_i and not in terms of $\phi(x)$. In the continuous limit, a natural object is the Lagrangian density which allows one to derive the relevant equation of motion directly.

It is easy to extend the Euler-Lagrange equation (4.77) to three dimensions. Now the field is a function of $\vec{x}$, and $\vec{\nabla}\phi$ replaces $\partial\phi/\partial x$ in the argument of the Lagrangian density:

$$\mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi) = \mathcal{L}(\phi, \partial_\mu \phi)\,.$$

Then in (4.72), $\delta\mathcal{L}$ needs to be expanded in terms of $\delta(\partial\phi/\partial y)$ and $\delta(\partial\phi/\partial z)$ in addition to $\delta(\partial\phi/\partial x)$, and following through the same derivation results in

$$\frac{\partial}{\partial t}\Big(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\Big) + \frac{\partial}{\partial x}\Big(\frac{\partial \mathcal{L}}{\partial\big(\frac{\partial\phi}{\partial x}\big)}\Big) + \frac{\partial}{\partial y}\Big(\frac{\partial \mathcal{L}}{\partial\big(\frac{\partial\phi}{\partial y}\big)}\Big) + \frac{\partial}{\partial z}\Big(\frac{\partial \mathcal{L}}{\partial\big(\frac{\partial\phi}{\partial z}\big)}\Big) = \frac{\partial \mathcal{L}}{\partial \phi}$$

or

$$\partial_\mu\Big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\Big) = \frac{\partial \mathcal{L}}{\partial \phi}\,, \tag{4.79}$$

which is the Euler-Lagrange equation now written in a space-time symmetric form.

When there are more than one fields, $\phi_1(x), \ldots, \phi_N(x)$, the Lagrangian density is simply

$$\boxed{\mathcal{L}\Big(\tilde{\phi}, \partial_\mu\tilde{\phi}\Big)\,, \qquad \tilde{\phi}(x) \equiv (\phi_1(x), \ldots, \phi_N(x))\,,\ x^\mu = (t, \vec{x})}\,. \tag{4.80}$$

Then the Lagrangian and the action are given by

$$L = \int d^3x \mathcal{L}\,, \qquad S = \int d^4x \mathcal{L}\,, \tag{4.81}$$

where the integration range over time is (t_1, t_2), and the space integral is over the entire space (infinite), or within a cube defined by $-L/2 < x_i < L/2$ $(i = 1, 2, 3)$. If the range is infinite, then the fields are assumed to vanish at sufficiently large distances from the origin:[1]

$$\tilde{\phi}(t, \vec{x}) = 0 \quad (|\vec{x}| \to \infty)\,, \tag{4.82}$$

or if the range is finite, then the periodicity is imposed:

$$\tilde{\phi}(t, x + L, y, z) = \tilde{\phi}(t, x, y + L, z) = \tilde{\phi}(t, x, y, z + L) = \tilde{\phi}(t, x, y, z)\,. \tag{4.83}$$

The corresponding Euler-Lagrange equation is obtained by taking the variation with respect to only one field ϕ_k in (4.72) while keeping other fields fixed. Then the same derivation leads to

$$\boxed{\partial_\mu\Big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_k)}\Big) = \frac{\partial \mathcal{L}}{\partial \phi_k} \quad (k = 1, \ldots, N)}\,. \tag{4.84}$$

[1]Strictly speaking, what is required of how quickly $\phi(t, \vec{x})$ vanishes as $|\vec{x}| \to \infty$ is that surface terms such as $\int d^3x \vec{\nabla}\phi$ vanish when integrated over all space. It is usually sufficient if ϕ vanishes faster than $1/\vec{x}^2$.

This form indicates that if the Lagrangian density $\mathcal{L}$ is a Lorentz scalar, then the resulting equation of motion is also Lorentz-invariant assuming that ϕ_k transforms in a well-defined way on both sides of the equation (scalar field, vector field, etc.).

Hamiltonian density

Let's start from the one-dimensional discrete system whose total Lagrangian was given by $L(\underline{\phi}, \underline{\dot{\phi}})$ (4.59) with $\underline{\phi} = \phi_i$ $(i = 1, \ldots, n)$. One should not confuse the index i in ϕ_i, which labels the position along the string, and the index k in $\phi_k(t, \vec{x})$ which labels the field. Here, we have only one field. The canonical momentum corresponding to the i-th general coordinate ϕ_i is

$$\pi_i \stackrel{\text{def}}{\equiv} \frac{\partial L}{\partial \dot{\phi}_i} = \Delta x \mu \dot{\phi}_i \,, \tag{4.85}$$

and the Hamiltonian is by definition

$$\begin{aligned} H &\equiv \sum_i \pi_i \dot{\phi}_i - L \\ &= \sum_i \Delta x \mu \dot{\phi}_i^2 - \sum_i \Delta x \Big[\frac{\mu}{2} \dot{\phi}_i^2 - \frac{\kappa}{2} \phi_i^2 - \frac{\tau}{2} \Big(\frac{\phi_i - \phi_{i-1}}{\Delta x} \Big)^2 \Big] \\ &= \sum_i \Delta x \Big[\frac{\mu}{2} \dot{\phi}_i^2 + \frac{\kappa}{2} \phi_i^2 + \frac{\tau}{2} \Big(\frac{\phi_i - \phi_{i-1}}{\Delta x} \Big)^2 \Big] \,. \end{aligned} \tag{4.86}$$

This is kinetic energy + potential energy, and thus it is the total energy of the system. In the continuous limit, it becomes

$$H = \int dx \Big[\frac{\mu}{2} \dot{\phi}^2 + \frac{\kappa}{2} \phi^2 + \frac{\tau}{2} \Big(\frac{\partial \phi}{\partial x} \Big)^2 \Big] \,. \tag{4.87}$$

The same result can be obtained by the following general procedure: first, the canonical field conjugate to ϕ is defined using the Lagrangian density $\mathcal{L}(\phi, \dot{\phi}, \partial\phi/\partial x)$ as

$$\pi(x) \stackrel{\text{def}}{\equiv} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \mu \dot{\phi}(x) \,, \tag{4.88}$$

and the *Hamiltonian density* $\mathcal{H}$ is defined by

$$\begin{aligned} \mathcal{H}\Big(\pi, \phi, \frac{\partial \phi}{\partial x} \Big) &\stackrel{\text{def}}{\equiv} \pi \dot{\phi} - \mathcal{L}\Big(\phi, \dot{\phi}, \frac{\partial \phi}{\partial x} \Big) \\ &= \mu \dot{\phi}^2 - \Big[\frac{\mu}{2} \dot{\phi}^2 - \frac{\kappa}{2} \phi^2 - \frac{\tau}{2} \Big(\frac{\partial \phi}{\partial x} \Big)^2 \Big] \\ &= \frac{\mu}{2} \dot{\phi}^2 + \frac{\kappa}{2} \phi^2 + \frac{\tau}{2} \Big(\frac{\partial \phi}{\partial x} \Big)^2 \end{aligned} \tag{4.89}$$

The total Hamiltonian is then its space integral:

$$H = \int dx \mathcal{H} \,, \tag{4.90}$$

which gives the same Hamiltonian as (4.87).

Extension to multiple fields in 3-dimension is straightforward. For a set of fields $\tilde{\phi} = (\phi_1(x), \ldots, \phi_N(x))$, the conjugate fields are defined by

$$\boxed{\pi_k(x) \stackrel{\text{def}}{\equiv} \frac{\partial \mathcal{L}\left(\tilde{\phi}, \partial_\mu \tilde{\phi}\right)}{\partial \dot{\phi}_k} \qquad (k = 1, \ldots, N)}\,; \tag{4.91}$$

then the total Hamiltonian is then given by

$$\boxed{H = \int d^3x \mathcal{H}}\,, \tag{4.92}$$

with the Hamiltonian density defined by

$$\boxed{\mathcal{H}\left(\tilde{\pi}, \tilde{\phi}, \vec{\nabla}\tilde{\phi}\right) \stackrel{\text{def}}{\equiv} \sum_k \pi_k \dot{\phi}_k - \mathcal{L}}\,. \tag{4.93}$$

Noether currents

One advantage of Lagrangian formulation is that symmetry of the system it describes is evident in the form of the Lagrangian. If certain transformation leaves the action invariant, then the resulting equations of motion will be the same before and after the transformation and thus the law of physics it describes will stay the same. In general, transformations that leave the action invariant are easy to spot in Lagrangian.

There exists a general theorem, called *Noether's theorem*, that states that for each transformation that leaves the action invariant, there exists a conserved quantity. We now discuss such conserved quantities corresponding to the symmetry under the space-time translations, which turn out to be the total energy and momentum of the system.

That the equations of motion (or the laws of physics) is the same anytime anywhere means that the Lagrangian does not depend on space-time explicitly. That is actually what we have been assuming since there is no explicit t in the argument of $\mathcal{L}(\phi, \partial_\mu \phi)$. Suppose a function $\phi(x)$, $x^\mu = (t, \vec{x})$, is given; namely, the motion of the field is specified. Then the value of the Lagrangian density $\mathcal{L}$ is uniquely defined at each space-time point x^μ. The difference between the values of $\mathcal{L}$ at x^μ and at $x^\mu + \epsilon^\mu$ is then given by

$$\delta \mathcal{L} = (\partial_\mu \mathcal{L})\, \epsilon^\mu \,, \tag{4.94}$$

where $\partial_\mu \mathcal{L}$ is understood to be the space-time dependence of $\mathcal{L}(\phi, \partial_\mu \phi)$ through the space-time dependence of $\phi(x)$ and $\partial_\mu \phi$. Thus, the same quantity can be written in terms of the changes in $\phi(x)$ and $\partial_\mu \phi(x)$. Using

$$\delta\phi \equiv \phi(x+\epsilon) - \phi(x) = (\partial_\nu \phi)\epsilon^\nu \tag{4.95}$$

$$\delta(\partial_\mu \phi) \equiv \partial_\mu \phi(x+\epsilon) - \partial_\mu \phi(x) = \partial_\nu(\partial_\mu \phi)\epsilon^\nu \tag{4.96}$$

as well as the Euler-Lagrange equation (4.79), we have

$$\begin{aligned}
\delta\mathcal{L} &= \underbrace{\frac{\partial \mathcal{L}}{\partial \phi}} \overbrace{\delta\phi}^{(\partial_\nu\phi)\epsilon^\nu} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \overbrace{\delta(\partial_\mu \phi)}^{(\partial_\nu\partial_\mu\phi)\epsilon^\nu} \\
&\quad \partial_\mu\Big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\Big) \text{ by (4.79)} \\
&= \Big[\partial_\mu\Big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\Big)\partial_\nu \phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial_\mu\partial_\nu \phi\Big]\epsilon^\nu \\
&= \partial_\mu\Big[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial_\nu \phi\Big]\epsilon^\nu \,.
\end{aligned} \tag{4.97}$$

Note that this holds only for the true motion $\phi(x)$ since we have used the Euler-Lagrange equation. Equating (4.94) and (4.97), we can factor out ϵ^ν and the derivative as

$$\delta\mathcal{L} = \partial_\mu\Big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial_\nu \phi\Big)\epsilon^\nu = \partial_\mu \mathcal{L} \underbrace{\epsilon^\mu}_{g^\mu{}_\nu \epsilon^\nu}$$

$$\rightarrow \quad \partial_\mu\Big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial_\nu \phi - \mathcal{L}\, g^\mu{}_\nu\Big)\epsilon^\nu = 0\,. \tag{4.98}$$

Since ϵ^ν is small but otherwise arbitrary, we have four conserved currents:

$$\partial_\mu J^{\mu\nu} = 0 \quad (\nu = 0, 1, 2, 3)\,, \tag{4.99}$$

with

$$J^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial^\nu \phi - \mathcal{L}\, g^{\mu\nu}\,. \tag{4.100}$$

These conserved currents are called the *Noether currents* of the space-time translations. Then, the space integral of the time component of each current is conserved:

$$\partial_0 J^{0\nu} = -\partial_i J^{i\nu}$$

$$\rightarrow \quad \frac{\partial}{\partial t}\Big(\int d^3x J^{0\nu}\Big) = \int d^3x\, \partial_0 J^{0\nu} = -\int d^3x\, \partial_i J^{i\nu} = -\int_A da_i\, J^{i\nu} = 0 \tag{4.101}$$

where A is the boundary surface, da_i is the i-th component of the area element and the last equality is due to the boundary condition (4.82) or (4.83). The conserved quantities are thus

$$P^\nu \stackrel{\text{def}}{\equiv} \int d^3x J^{0\nu} \quad (\nu = 0, 1, 2, 3)\,. \tag{4.102}$$

We can see that P^0 is nothing but the Hamiltonian (namely, the total energy):

$$P^0 \equiv \int d^3x J^{00} = \int d^3x \underbrace{\Big(\overbrace{\frac{\partial \mathcal{L}}{\partial \dot{\phi}}}^{\pi} \dot{\phi} - \mathcal{L} \Big)}_{\mathcal{H}} = H\,. \tag{4.103}$$

Then, it is natural to identify the space component $\vec{P}$ given by

$$P^i \equiv \int d^3x J^{0i} = \int d^3x \Big(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \partial^i \phi \Big) = \int d^3x\, \pi\, \partial^i \phi \tag{4.104}$$

as the total momentum of the system. In fact, this identification is required if we can prove that P^ν is a 4-vector, which, as it turns out, is not trivial. The main problem is that P^ν's are defined as space integrals in a given frame. In the following, we will assume that $\mathcal{L}$ is a Lorentz scalar and show that P^ν transforms as a 4-vector under an infinitesimal Lorentz transformation which is sufficient for all proper and orthochronous transformations.

The following proof does not require that ϕ is a scalar, as long as fields are combined in $\mathcal{L}$ to make it a scalar quantity. That $\mathcal{L}$ is a Lorentz scalar means

$$\mathcal{L}'(x') = \mathcal{L}(x)\,, \quad x' = \Lambda x\,. \tag{4.105}$$

Then, since the quantity $\delta\mathcal{L}$ in (4.97) is a Lorentz scalar (it is the difference of the values of $\mathcal{L}$ at two event points, x and $x+\epsilon$) and ∂_μ and ϵ^ν are 4-vectors, the quantity inside the square bracket should transform as a tensor. Together with the fact that the metric tensor is itself a tensor (1.65), we see that $J^{\mu\nu}$ defined by (4.100) transforms as a tensor:

$$J'^{\mu\nu}(x') = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta J^{\alpha\beta}(x)\,. \tag{4.106}$$

Since P'^ν and P^ν are constants of motion, we choose to evaluate them at $t' = 0$ and $t = 0$, respectively:

$$P'^\nu \equiv \int d^3x' J'^{0\nu}(0, \vec{x}')\,, \quad P^\nu \equiv \int d^3x J^{0\nu}(0, \vec{x})\,. \tag{4.107}$$

For a proper transformation, we have $\det \Lambda = 1$ and thus

$$d^4x' = (\det \Lambda) d^4x = d^4x\,. \tag{4.108}$$

Also, using the property of the delta function

$$\delta(f(s)) = \sum_i \frac{1}{|f'(s_i)|}\delta(s - s_i)\,, \quad (s_i\text{: solutions of } f(s) = 0)\,, \tag{4.109}$$

we have

$$\delta(t') = \delta(\Lambda^0{}_\rho\, x^\rho) = \frac{1}{\Lambda^0{}_0}\delta\Big(t + \frac{\Lambda^0{}_i x^i}{\Lambda^0{}_0}\Big)\,. \tag{4.110}$$

Then, P'^ν can be written as

$$\begin{aligned} P'^\nu &\equiv \int d^3x'\, J'^{0\nu}(0,\vec{x}') = \int \overbrace{d^4x'}^{d^4x}\, \underbrace{\delta(t')}_{\frac{1}{\Lambda^0{}_0}\delta\left(t+\frac{\Lambda^0{}_i x^i}{\Lambda^0{}_0}\right)}\, \overbrace{J'^{0\nu}(t',\vec{x}')}^{\Lambda^0{}_\alpha\Lambda^\nu{}_\beta\, J^{\alpha\beta}(t,\vec{x})} \\ &= \frac{1}{\Lambda^0{}_0}\int dt d^3x\, \delta\Big(t + \frac{\Lambda^0{}_i x^i}{\Lambda^0{}_0}\Big)\, \Lambda^0{}_\alpha\Lambda^\nu{}_\beta\, J^{\alpha\beta}(t,\vec{x}) \\ &= \frac{1}{\Lambda^0{}_0}\int d^3x\, \Lambda^0{}_\alpha\Lambda^\nu{}_\beta\, J^{\alpha\beta}\Big(-\tfrac{\Lambda^0{}_i\, x^i}{\Lambda^0{}_0}, \vec{x}\Big)\,. \end{aligned} \tag{4.111}$$

If Λ is an infinitesimal transformation, it can be written as

$$\Lambda^\mu{}_\nu = g^\mu{}_\nu + \omega^\mu{}_\nu\,, \tag{4.112}$$

where ω is a small parameter. In particular, we have [see (1.88)]

$$\omega^0{}_0 = 0\,, \quad \to \quad \Lambda^0{}_0 = 1\,, \tag{4.113}$$

which means

$$-\frac{\Lambda^0{}_i\, x^i}{\Lambda^0{}_0} = -(g^0{}_i + \omega^0{}_i)\, x^i = -\omega^0{}_i\, x^i\,. \tag{4.114}$$

Then, to the first order in ω and expanding $J^{\alpha\beta}(t,\vec{x})$ about $t = 0$, P'^ν becomes

$$\begin{aligned} P'^\nu &= \int d^3x\, (g^0{}_\alpha + \omega^0{}_\alpha)(g^\nu{}_\beta + \omega^\nu{}_\beta)\, J^{\alpha\beta}(-(\omega^0{}_i\, x^i), \vec{x}) \\ &= \int d^3x\, (g^0{}_\alpha g^\nu{}_\beta + g^0{}_\alpha\omega^\nu{}_\beta + \omega^0{}_\alpha g^\nu{}_\beta)\Big[J^{\alpha\beta}(0,\vec{x}) - (\partial_0 J^{\alpha\beta})(\omega^0{}_i\, x^i)\Big] \\ &= \int d^3x\, \Big[J^{0\nu} + \omega^\nu{}_\beta J^{0\beta} + \underbrace{\omega^0{}_\alpha J^{\alpha\nu}}_{\omega^0{}_i J^{i\nu}} - \underbrace{(\partial_0 J^{0\nu})}_{-\partial_j J^{j\nu}}(\omega^0{}_i\, x^i)\Big] \quad \text{by (4.99)} \\ &= \underbrace{\int d^3x J^{0\nu}}_{P^\nu} + \omega^\nu{}_\beta \underbrace{\int d^3x J^{0\beta}}_{P^\beta} + \int d^3x\, \omega^0{}_i J^{i\nu} + \int d^3x\, \omega^0{}_i\, \underbrace{(\partial_j J^{j\nu})x^i}_{\underbrace{\partial_j(J^{j\nu}x^i)}_{\to 0} - J^{j\nu}\underbrace{\partial_j x^i}_{\delta_{ij}}} \end{aligned} \tag{4.115}$$

$$
\begin{aligned}
&= P^\nu + \omega^\nu{}_\beta P^\beta + \int d^3x \cancel{\omega^0{}_i} J^{i\nu} - \int d^3x \cancel{\omega^0{}_i} J^{i\nu} \\
&= (g^\nu{}_\beta + \omega^\nu{}_\beta) P^\beta \\
&= \Lambda^\nu{}_\beta P^\beta \,, \qquad (4.115)
\end{aligned}
$$

where all $J^{\mu\nu}$'s are understood to be evaluated at $(0, \vec{x})$ unless otherwise indicated. Thus, we have shown that P^ν transforms as a 4-vector under an infinitesimal Lorentz transformation. The proof can be easily extended to a finite transformation by dividing the finite tranformation into a large number of infinitesimal transformations. The four conserved quantities P^ν therefore form a 4-vector, and since the time component has been shown to be the total energy, we identify P^ν as the total energy-momentum 4-vector. The tensor quantity $J^{\mu\nu}$ is called the *energy-momentum tensor*.

Extension to multiple fields is again straightforward: for a set of *real* fields $\phi_k(x)$ ($k = 1, \ldots, N$), the energy momentum tensor is given by

$$
\boxed{J^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_k)} \partial^\nu \phi_k - \mathcal{L}\, g^{\mu\nu} \,.}, \qquad (4.116)
$$

and the total energy and momentum are [using $\partial^i = -\nabla_i$ in (4.104)]

$$
\boxed{P^0 = H \,, \quad \vec{P} = -\int d^3x\, \pi_k \vec{\nabla} \phi_k}, \qquad (4.117)
$$

where π_k and H are given by (4.91) through (4.93), and summations over k are implied.

Transition to Lorentz invariance

We are jumping ahead too fast, since we formed our Lagrangian density for a classical system which is at rest in certain frame and yet assumed that the resulting Lagrangian density is a Lorentz scaler. Let's go back to the equation of motion and discuss how we can make the system consistent with special relativity. If we set $\mu = \tau = 1$ and $k = m^2$ in the equation of motion of the string (4.63), we obtain

$$
\frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} + m^2 \phi = 0 \,. \qquad (4.118)
$$

Then, we see that its three-dimensional extension is nothing but the Klein-Gordon equation $(\partial^2 + m^2)\phi = 0$. Namely, for the specific choice of parameters, the angular frequency ω and the wave number k of the plane-wave solutions satisfy $\omega^2 - k^2 = m^2$, and the equation of motion becomes Lorentz-invariant *provided* that we define the field $\phi(x)$ to be a scalar field:

$$
\phi'(x') = \phi(x) \quad (x' = \Lambda x) \,. \qquad (4.119)
$$

Just as we discarded the idea of *ether* that transmits light and took the electromagnetic fields and their transformation properties as the reality, we forget the classical string at this point and take the field $\phi(x)$ and its transformation property (4.119) as reality. As we will see shortly in the next section, the Lagrangian density indeed becomes a Lorentz scaler with this transformation property of ϕ and the choice of parameters $\mu = \tau = 1$ and $k = m^2$..

4.3 Quantization of the Klein-Gordon field

Canonical quantization

We are now ready to quantize the Klein-Gordon field following the standard canonical quantization procedure. First, the Lagrangian density of the Klein-Gordon field is obtained by setting $\mu = \tau = 1$ and $k = m^2$ in (4.65) and extending it to 3-dimension

$$\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2}\Big(\dot{\phi}^2 - (\vec{\nabla}\phi)^2 - m^2\phi^2\Big) \tag{4.120}$$

or

$$\boxed{\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2}\Big(\partial_\mu \phi\, \partial^\mu \phi - m^2\phi^2\Big)}. \tag{4.121}$$

If the above extension to the three-dimensional case is somewhat unclear, it is justified by the fact that this Lagrangian density indeed leads to the Klein-Gordon equation: the Euler-Lagrange equation (4.84) gives

$$\underbrace{\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right)}_{\partial^\mu \phi \,:\, \text{note the factor of 2}} - \overbrace{\frac{\partial \mathcal{L}}{\partial \phi}}^{-m^2\phi} = (\partial^2 + m^2)\phi = 0\,. \tag{4.122}$$

From the definition (4.91), the conjugate field is

$$\boxed{\pi(x) \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}\,,} \tag{4.123}$$

and the Hamiltonian density defined by (4.93) is then

$$\boxed{\mathcal{H}(\pi, \phi, \vec{\nabla}\phi) = \frac{1}{2}\Big(\pi^2 + (\vec{\nabla}\phi)^2 + m^2\phi^2\Big)}. \tag{4.124}$$

Note that the field $\phi(x)$ is a real number as in the case of the one-dimensional string. So far, this is just classical mechanics. The only difference from the string is

that the choice of parameters as well as the assumption that $\phi(x)$ is a scalar has made the Lagrangian density a Lorentz scaler; namely, when $\mathcal{L}$ is regarded as a function of x through $\phi(x)$, the value of $\mathcal{L}$ evaluated at x is the same as the value of $\mathcal{L}$ in the trassformed frame at $x' = \Lambda x$: Using $\phi'(x') = \phi(x)$ and $\partial'_\mu = \Lambda_\mu{}^\alpha \partial_\alpha$,

$$\begin{aligned}\mathcal{L}'(x') &\equiv \frac{1}{2}\Big(\underbrace{\partial'_\mu \phi'(x')\,\partial^{\mu'}\phi'(x')}_{\underbrace{\Lambda_\mu{}^\alpha \Lambda^\mu{}_\beta}_{g^\alpha{}_\beta}\,\partial_\alpha\phi(x)\partial^\beta\phi(x)} - m^2\phi'^2(x')\Big) \\ &= \frac{1}{2}\Big(\partial_\mu\phi(x)\,\partial^\mu\phi(x) - m^2\phi^2(x)\Big) = \mathcal{L}(x)\,. \end{aligned} \tag{4.125}$$

Now we will turn the system into a quantum mechanical system by regarding $\phi(t,\vec{x})$ and $\pi(t,\vec{x})$ as operators in the Heisenberg picture, and introducing commutation relations among them. Here, ϕ's and π's at different positions are considered independent observables.

For simplicity, let's start from the discrete one-dimensional case. The operator conjugate to ϕ_i is then given by (4.85) (with $\mu = 1$):

$$\pi_i(t) = \Delta x \dot{\phi}_i(t)\,. \tag{4.126}$$

The equal-time commutation relations are then well-defined; namely, the commutator of ϕ_i and π_i is i if they are conjugate to each other (i.e. have same index i) and zero otherwise:

$$[\phi_i(t), \pi_j(t)] = i\delta_{ij}\,, \tag{4.127}$$

$$[\phi_i(t), \phi_j(t)] = [\pi_i(t), \pi_j(t)] = 0\,. \tag{4.128}$$

The continuous limit of the conjugate field is defined directly from the lagrangian density and is given by (4.88) with $\mu = 1$:

$$\pi(t,\vec{x}) = \dot{\phi}(t,\vec{x}) \tag{4.129}$$

Together with the definition $\phi_i(t) \equiv \phi(t, x_i)$ (4.53), the relations connecting the continuous and discrete operators are:

$$\phi_i(t) = \phi(t, x_i)\,, \quad \pi_i(t) = \Delta x\, \pi(t, x_i)\,. \tag{4.130}$$

Thus, the commutation relation (4.127) can be written as

$$[\phi(t, x_i), \pi(t, x_j)] = i\frac{\delta_{ij}}{\Delta x}\,. \tag{4.131}$$

That this behaves as a delta function can be seen by integrating over all x:

$$\int dx[\phi(t,x),\pi(t,x')] \equiv \sum_i \Delta x \underbrace{[\phi(t,x_i),\pi(t,x_j)]}_{i\delta_{ij}/\Delta x} = \sum_i i\delta_{ij} = i\,, \tag{4.132}$$

which identifies $[\phi(t,x),\pi(t,x')]$ as $i\delta(x-x')$; thus, in the continuous limit we have

$$\begin{aligned}[\phi(t,x),\pi(t,x')] &= i\delta(x-x')\,,\\ [\phi(t,x),\phi(t,x')] &= [\pi(t,x),\pi(t,x')] = 0\,.\end{aligned}$$

Extension to the three-dimensional case is straightforward and given by

$$\boxed{\begin{aligned}[\phi(t,\vec{x}),\pi(t,\vec{x}')] &= i\delta^3(\vec{x}-\vec{x}')\,,\\ [\phi(t,\vec{x}),\phi(t,\vec{x}')] &= [\pi(t,\vec{x}),\pi(t,\vec{x}')] = 0\,.\end{aligned}}\,, \tag{4.133}$$

with

$$\delta^3(\vec{x}-\vec{x}') \stackrel{\text{def}}{\equiv} \delta(x-x')\delta(y-y')\delta(z-z')\,. \tag{4.134}$$

Note that we are using the Heisenberg picture where all time dependences are contained in the operators, and that the commutation relations above are defined at a given time t.

Heisenberg's equations of motion

We are now in the realm of quantum mechanics. We have infinite number of operators $\phi(t,\vec{x})$ representing the field values at different positions in space, and their conjugate operators $\pi(t,\vec{x})$. Then, the time dependence of any operator which is a function of ϕ and π is given by Heisenberg's equation of motion and the commutation relations (4.133). One natural question is then what kind of equation of motion does ϕ satisfy, is it the Klein-Gordon equation? As we will see now, the answer is yes. Heisenberg's equations of motion for ϕ and π are

$$-i\dot{\phi}(t,\vec{x}) = [H,\phi(t,\vec{x})]\,, \tag{4.135}$$

$$-i\dot{\pi}(t,\vec{x}) = [H,\pi(t,\vec{x})]\,, \tag{4.136}$$

where all times are understood to be the same. Note that what appears in the commutators is the total Hamiltonian and not the Hamiltonian density. Using the Hamiltonian density (4.124) and $H = \int d^3x\mathcal{H}$, the above equations are written as (dropping 't' for simplicity)

$$-i\dot{\phi}(\vec{x}) = \Big[\int d^3x'\frac{1}{2}\Big(\pi^2(\vec{x}') + (\vec{\nabla}'\phi(\vec{x}'))^2 + m^2\phi^2(\vec{x}')\Big)\,,\phi(\vec{x})\Big]\,, \tag{4.137}$$

$$-i\dot{\pi}(\vec{x}) = \Big[\int d^3x'\frac{1}{2}\Big(\pi^2(\vec{x}') + (\vec{\nabla}'\phi(\vec{x}'))^2 + m^2\phi^2(\vec{x}')\Big)\,,\pi(\vec{x})\Big]\,. \tag{4.138}$$

Some mathematical preparations are in order. For the operator function $\phi(\vec{x})$, the definition of $\nabla_i\phi(x)$ is

$$\frac{\partial\phi}{\partial x^i}(\vec{x}) \stackrel{\text{def}}{\equiv} \frac{\phi(\vec{x}+\epsilon\,\hat{e}_i)-\phi(\vec{x})}{\epsilon} \tag{4.139}$$

where ϵ is a small real number and $\hat{e}_i$ is the unit vector in the direction of the i-th axis. We note that $\partial_i\phi$ is nothing but the difference between neighboring $\phi(\vec{x})$'s. Since $\phi(\vec{x})$ and $\phi(\vec{x}')$ commute at equal time for any $\vec{x}$ and $\vec{x}'$, we then see that $\partial_i\phi(\vec{x}')$ commutes with $\phi(\vec{x})$ (at a given time t):

$$[\vec{\nabla}'\phi(\vec{x}'),\phi(\vec{x})] = 0\,. \tag{4.140}$$

Note that the same argument cannot be applied to the time derivative $\dot{\phi}$ since the commutation relation $[\phi(t,\vec{x}),\phi(t,\vec{x}')]=0$ is for a given time t; namely, commutators such as $[\phi(t+dt,\vec{x}),\phi(t,\vec{x}')]$ are not defined at this point, and in fact it is not in general zero as we will see later.

Next, if $f(\vec{x})$ is an arbitrary operator that is a function of $\vec{x}$, then from the definition of $\partial_i\phi$ (4.139),

$$\begin{aligned}
\int d^3x' f(\vec{x}')\Big[\frac{\partial\phi}{\partial x'^i}(\vec{x}'),\pi(\vec{x})\Big] &= \int d^3x' f(\vec{x}')\Big[\frac{\phi(\vec{x}'+\epsilon\,\hat{e}_i)-\phi(\vec{x}')}{\epsilon},\pi(\vec{x})\Big] \\
&= \frac{1}{\epsilon}\int d^3x' f(\vec{x}')\Big\{\underbrace{[\phi(\vec{x}'+\epsilon\,\hat{e}_i),\pi(\vec{x})]}_{i\delta^3(\vec{x}'+\epsilon\,\hat{e}_i-\vec{x})} - \underbrace{[\phi(\vec{x}'),\pi(\vec{x})]}_{i\delta^3(\vec{x}'-\vec{x})}\Big\} \\
&= \frac{i}{\epsilon}\Big(f(\vec{x}-\epsilon\,\hat{e}_i)-f(\vec{x})\Big) \\
&= -i\,\frac{\partial f}{\partial x^i}(\vec{x})\,.
\end{aligned} \tag{4.141}$$

In the above, note that the same result is obtained when the operator $f(\vec{x}')$ is placed after the commutator.

Then, (4.137) becomes

$$\begin{aligned}
-i\dot{\phi}(\vec{x}) &= \frac{1}{2}\int d^3x'\Big[\pi^2(\vec{x}') + \underbrace{(\vec{\nabla}'\cancel{\phi(\vec{x}')})^2}_{\text{by (4.140)}} + m^2\cancel{\phi^2(\vec{x}')}\,,\phi(\vec{x})\Big] \\
&= \frac{1}{2}\int d^3x'\ \underbrace{[\pi^2(\vec{x}'),\phi(\vec{x})]}_{\pi(\vec{x}')\ \underbrace{[\pi(\vec{x}'),\phi(\vec{x})]}_{-i\delta^3(\vec{x}-\vec{x}')} + \underbrace{[\pi(\vec{x}'),\phi(\vec{x})]}_{-i\delta^3(\vec{x}-\vec{x}')}\ \pi(\vec{x}')} \\
&= -i\pi(\vec{x})\,;
\end{aligned} \tag{4.142}$$

namely,

$$\boxed{\dot{\phi}(t,\vec{x}) = \pi(t,\vec{x})}, \tag{4.143}$$

reproducing (4.129), this time derived from Heisenberg's equation of motion. The second equation of motion (4.138) is now,

$$\begin{aligned} -i\dot{\pi}(\vec{x}) &= \frac{1}{2}\int d^3x' \Big[\cancel{\pi^2}(\vec{x}') + (\vec{\nabla}'\phi(\vec{x}'))^2 + m^2\phi^2(\vec{x}')\,, \pi(\vec{x})\Big] \\ &= \underbrace{\frac{1}{2}\int d^3x' \Big[(\vec{\nabla}'\phi(\vec{x}'))^2, \pi(\vec{x})\Big]}_{\equiv I_1} + \underbrace{\frac{m^2}{2}\int d^3x' \Big[\phi^2(\vec{x}'), \pi(\vec{x})\Big]}_{\equiv I_2}\,. \end{aligned} \tag{4.144}$$

Using (4.141), the first term becomes

$$\begin{aligned} I_1 &= \frac{1}{2}\sum_i \int d^3x' \Big[(\partial_i'\phi(\vec{x}'))^2, \pi(\vec{x})\Big] \\ &= \frac{1}{2}\sum_i \underbrace{\int d^3x' \Big\{\partial_i'\phi(\vec{x}')\,[\partial_i'\phi(\vec{x}'), \pi(\vec{x})]}_{-i\partial_i\partial_i\phi(\vec{x})} + \underbrace{[\partial_i'\phi(\vec{x}'), \pi(\vec{x})]\,\partial_i'\phi(\vec{x}')}_{\to\ -i\partial_i\partial_i\phi(\vec{x})}\Big\} \\ &= -i\nabla^2\phi(\vec{x})\,, \end{aligned} \tag{4.145}$$

while the second term is simpler:

$$\begin{aligned} I_2 &= \frac{m^2}{2}\int d^3x' \Big\{\phi(\vec{x}')\underbrace{[\phi(\vec{x}'), \pi(\vec{x})]}_{i\delta(\vec{x}'-\vec{x})} + \underbrace{[\phi(\vec{x}'), \pi(\vec{x})]}_{i\delta(\vec{x}'-\vec{x})}\phi(\vec{x}')\Big\} \\ &= im^2\phi(\vec{x})\,. \end{aligned} \tag{4.146}$$

Putting the pieces together, we obtain

$$\dot{\pi}(t,\vec{x}) = \nabla^2\phi(t,\vec{x}) - m^2\phi(t,\vec{x})\,, \tag{4.147}$$

which, combined with $\pi = \dot{\phi}$ (4.143), leads to

$$\begin{aligned} \ddot{\phi}(t,\vec{x}) &= \nabla^2\phi(t,\vec{x}) - m^2\phi(t,\vec{x})\,, \\ \to \quad & \boxed{(\partial^2 + m^2)\phi(t,\vec{x}) = 0}, \end{aligned} \tag{4.148}$$

which is nothing but the Klein-Gordon equation. Thus, we have shown that Heisenberg's equations of motion for $\phi(x)$ and $\pi(x)$ together with the commutation relations among $\phi(x)$ and $\pi(x)$ lead to the Klein-Gordon equation which is now satisfied by the operator field $\phi(t,\vec{x})$ in the Heisenberg picture.

Momentum expansion

We will now express the operator field $\phi(t,\vec{x})$ in terms of normal modes and introduce creation and annihilation operators associated with each mode. In performing the

expansion, it is convenient to impose the periodicity (4.83), and focus our attention on the (large) box $-L/2 < x_i < L/2$ $(i = 1, 2, 3)$. Then at a given time t, any operator function $\phi(t, \vec{x})$ can be expanded as

$$\phi(t, \vec{x}) = \sum_{\vec{p}} C_{\vec{p}}(t) e^{i\vec{p}\cdot\vec{x}}\,, \tag{4.149}$$

where $C_{\vec{p}}$ are operators and $\vec{p}$ takes discrete values that satisfy

$$\begin{gathered}\vec{p} = (p_x, p_y, p_z) = \Big(\frac{2\pi n_x}{L}, \frac{2\pi n_y}{L}, \frac{2\pi n_z}{L}\Big) \\ (-\infty < n_x, n_y, n_z < \infty : \text{integers})\,.\end{gathered} \tag{4.150}$$

As we will show below, the set of functions $\{e^{i\vec{p}\cdot\vec{x}}\}$ forms a complete orthonormal set in the box:

$$\begin{aligned}\int d^3x (e^{i\vec{p}\cdot\vec{x}})^* e^{i\vec{p}'\cdot\vec{x}} &= V\delta_{\vec{p},\vec{p}'} \\ \sum_{\vec{p}} (e^{i\vec{p}\cdot\vec{x}})^* e^{i\vec{p}\cdot\vec{x}'} &= V\delta^3(\vec{x} - \vec{x}')\,,\end{aligned} \tag{4.151}$$

where $\delta_{\vec{p},\vec{p}'}$ is defined as

$$\delta_{\vec{p},\vec{p}'} \overset{\text{def}}{\equiv} \begin{cases} 1 & \text{if } \vec{p} = \vec{p}' \\ 0 & \text{otherwise} \end{cases}\,. \tag{4.152}$$

In the above and hereafter, the space integral is understood to be within the box of volume $V \equiv L^3$ unless otherwise stated. The first of (4.151) is easy to verify: since if $\vec{p} = \vec{p}'$ then integrand is 1 and the integral is V, while if $\vec{p} \neq \vec{p}'$ then the integrand oscillates and the integral vanishes. The second of (4.151) can be proven by converting the sum over $\vec{p}$ to an integral using

$$\sum_{\vec{p}} = \frac{V}{(2\pi)^3}\int d^3p \qquad (L \to \infty)\,. \tag{4.153}$$

This can be seen as follows: the possible values of $\vec{p} = (2\pi n_x/L, 2\pi n_y/L, 2\pi n_z/L)$ are on a cubic grid of cell size $2\pi/L$. Each cell of volume $dv = (2\pi/L)^3$ contains exactly one grid point, say, at its center. The integral of an arbitrary function $f(\vec{p})$ over $\vec{p}$ can be written in the limit $2\pi/L \to 0$ as

$$\int d^3p f(\vec{p}) = \sum_{\vec{p}} \underbrace{dv}_{(2\pi/L)^3} f(\vec{p}) = \frac{(2\pi)^3}{V}\sum_{\vec{p}} f(\vec{p})\,, \tag{4.154}$$

which establishes the correspondence (4.153). Then, we have

$$\sum_{\vec{p}} (e^{i\vec{p}\cdot\vec{x}})^* e^{i\vec{p}\cdot\vec{x}'} = \frac{V}{(2\pi)^3}\underbrace{\int d^3p\, e^{i\vec{p}\cdot(\vec{x}'-\vec{x})}}_{(2\pi)^3\delta^3(\vec{x}-\vec{x}')} = V\delta^3(\vec{x} - \vec{x}')\,, \tag{4.155}$$

which proves the second orthogonality of (4.151)

The expansion (4.149) is unique since the coefficient $C_{\vec{p}}$ is uniquely given by

$$\int d^3x (e^{i\vec{p}\cdot\vec{x}})^* \phi(t,\vec{x}) = \sum_{\vec{p}'} C_{\vec{p}'}(t) \underbrace{\int d^3x (e^{i\vec{p}\cdot\vec{x}})^* e^{i\vec{p}'\cdot\vec{x}}}_{V\delta_{\vec{p},\vec{p}'}} = V C_{\vec{p}}\,,$$

$$\rightarrow \quad C_{\vec{p}} = \frac{1}{V}\int d^3x (e^{i\vec{p}\cdot\vec{x}})^* \phi(t,\vec{x})\,, \tag{4.156}$$

and it is complete since the original function ϕ is recovered as

$$\begin{aligned} \sum_{\vec{p}} C_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} &= \frac{1}{V}\sum_{\vec{p}}\int d^3x' (e^{i\vec{p}\cdot\vec{x}'})^* \phi(t,\vec{x}') e^{i\vec{p}\cdot\vec{x}} \\ &= \frac{1}{V}\int d^3x' \underbrace{\sum_{\vec{p}} (e^{i\vec{p}\cdot\vec{x}'})^* e^{i\vec{p}\cdot\vec{x}}}_{V\delta^3(\vec{x}-\vec{x}')} \phi(t,\vec{x}') = \phi(t,\vec{x}) \end{aligned} \tag{4.157}$$

Since ϕ is an observable, it is hermitian, and thus

$$\phi^\dagger = \phi \quad \rightarrow \quad \underbrace{\sum_{\vec{p}} C^\dagger_{\vec{p}}\, e^{-i\vec{p}\cdot\vec{x}}}_{\sum_{\vec{p}} C^\dagger_{-\vec{p}} e^{i\vec{p}\cdot\vec{x}} \text{ (relabeled } \vec{p}\to-\vec{p})} = \sum_{\vec{p}} C_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}}\,. \tag{4.158}$$

Equating the coefficients of the orthonormal functions, the hermiticity condition is

$$C^\dagger_{-\vec{p}}(t) = C_{\vec{p}}(t) \tag{4.159}$$

Thus, any hermitian field can be expanded at any given time as (4.149) with the hermiticity condition (4.159). Now we will see below that the time dependence of the field, past and future, is then uniquely determined if $\phi(t,\vec{x})$ is to satisfy the Klein-Gordon equation. Substituting (4.149) in $\ddot{\phi} = \nabla^2\phi - m^2\phi$,

$$\sum_{\vec{p}} \ddot{C}_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} = \sum_{\vec{p}} \left[(i\vec{p})^2 - m^2\right] C_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}}$$

$$\rightarrow \quad \ddot{C}_{\vec{p}} = -{p^0}^2 C_{\vec{p}}\,, \tag{4.160}$$

where p^0 is defined as a positive number given by

$$p^0 \equiv \sqrt{\vec{p}^{\,2} + m^2} \geq 0\,. \tag{4.161}$$

The general solution is then

$$C_{\vec{p}}(t) = A_{\vec{p}}\, e^{-ip^0 t} + B_{\vec{p}}\, e^{ip^0 t}\,, \tag{4.162}$$

where $A_{\vec{p}}$ and $B_{\vec{p}}$ are constant operators. The hermitian condition (4.159) is now

$$A^\dagger_{-\vec{p}}\, e^{ip^0 t} + B^\dagger_{-\vec{p}}\, e^{-ip^0 t} = A_{\vec{p}}\, e^{-ip^0 t} + B_{\vec{p}}\, e^{ip^0 t} \tag{4.163}$$

which should hold for all t; namely,

$$A^\dagger_{-\vec{p}} = B_{\vec{p}}\,, \quad B^\dagger_{-\vec{p}} = A_{\vec{p}} \tag{4.164}$$

where the second condition is actually the same as the first. We then obtain the general solution of the Klein-Gordon equation by substituting (4.162) in the expansion (4.149) and using $A^\dagger_{-\vec{p}} = B_{\vec{p}}$:

$$\begin{aligned}
\phi(t,\vec{x}) &= \sum_{\vec{p}} \Big(A_{\vec{p}}\, e^{-ip^0 t} + A^\dagger_{-\vec{p}}\, e^{ip^0 t}) e^{i\vec{p}\cdot\vec{x}} \\
&= \sum_{\vec{p}} \Big(A_{\vec{p}}\, e^{-ip^0 t + i\vec{p}\cdot\vec{x}} + \underbrace{A^\dagger_{-\vec{p}}\, e^{ip^0 t + i\vec{p}\cdot\vec{x}}}_{\text{relabel } \vec{p} \to -\vec{p}} \Big) \\
&= \sum_{\vec{p}} \Big(A_{\vec{p}}\, e^{-ip\cdot x} + A^\dagger_{\vec{p}}\, e^{ip\cdot x} \Big)
\end{aligned} \tag{4.165}$$

where we have used the 4-vector notation $p \cdot x = p^0 t - \vec{p}\cdot\vec{x}$. Note that in relabeling $\vec{p} \to -\vec{p}$, the value of p^0 did not change because of the definition (4.161).

In analogy to the single harmonic oscillator case (4.37), we define the operator $a_{\vec{p}}$ as

$$A_{\vec{p}} = \frac{1}{\sqrt{2p^0 V}} a_{\vec{p}}\,. \tag{4.166}$$

We will see below that $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$ act as the annihilation and the creation operator for a particle with momentum $\vec{p}$. The normalization factor $1/\sqrt{2p^0 V}$ is such that the commutation relations among ϕ and π (4.133) leads to the commutation relation $[a_{\vec{p}}, a^\dagger_{\vec{p}}] = \delta_{\vec{p},\vec{p}'}$ while satisfying $\pi = \dot{\phi}$.

Putting (4.165) and (4.166) together, the expansion can be written as

$$\boxed{\phi(x) = \sum_{\vec{p}} \Big(a_{\vec{p}}\, e_{\vec{p}}(x) + a^\dagger_{\vec{p}}\, e^*_{\vec{p}}(x) \Big)} \tag{4.167}$$

with the *normal-mode functions* defined by

$$\boxed{e_{\vec{p}}(x) \stackrel{\text{def}}{\equiv} \frac{e^{-ip\cdot x}}{\sqrt{2p^0 V}} \quad (x \equiv (t,\vec{x}),\ p^0 = \sqrt{\vec{p}^{\,2} + m^2} \geq 0)} \tag{4.168}$$

This expansion is general and unique; that is, you give me any solution of the Klein-Gordon equation, then I will look at its spacial dependence at a fixed time t and expand it uniquely as (4.149), then the time dependence for the past and future is already uniquely determined as in (4.162) in order to satisfy the Klein-Gordon equation, and together with the hermitian condition (4.164), it results in the expansion above. Note that p^0 is a function of $\vec{p}$, $e_{\vec{p}}(x)$ is just a number (namely, not an operator), and $a_{\vec{p}}$ is an operator that does not depend on t. The conjugate field can be obtained by (4.143):

$$\pi(x) = \dot{\phi}(x) = \sum_{\vec{p}} (-ip^0)\Big(a_{\vec{p}}\, e_{\vec{p}}(x) - a^\dagger_{\vec{p}}\, e^*_{\vec{p}}(x)\Big)\,, \tag{4.169}$$

which is also hermitian.

In order to show that $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$ indeed are creation and annihilation operators, we need the orthonormality relations of the normal-mode functions:

$$\begin{aligned}
\int d^3x\, e^*_{\vec{p}}(x) e_{\vec{p}\,\prime}(x) &= \frac{1}{2V\sqrt{p^0 p^{0\prime}}} \int d^3x\, e^{ip\cdot x} e^{-ip'\cdot x} \\
&= \frac{e^{i(p^0 - p^{0\prime})t}}{2V\sqrt{p^0 p^{0\prime}}} \underbrace{\int d^3x (e^{i\vec{p}\cdot\vec{x}})^* e^{i\vec{p}\,'\cdot\vec{x}}}_{V\delta_{\vec{p},\vec{p}\,'}} \\
&= \frac{\delta_{\vec{p},\vec{p}\,'}}{2p^0}\,,
\end{aligned} \tag{4.170}$$

where in the last step we have set $p^0 = p^{0\prime}$ which is allowed because of the delta function $\delta_{\vec{p},\vec{p}\,'}$. The integral of $e_{\vec{p}}(x) e_{\vec{p}\,\prime}(x)$ can be obtained similarly:

$$\int d^3x\, e_{\vec{p}}(x) e_{\vec{p}\,\prime}(x) = e^{-2ip^0 t}\, \frac{\delta_{\vec{p},-\vec{p}\,'}}{2p^0}\,. \tag{4.171}$$

This is a little awkward; it would be much better if the above gives zero instead of the delta function with a phase factor. This can be accomplished by using the differential operator $\overset{\leftrightarrow}{\partial}_0$ introduced in (2.42):

$$a \overset{\leftrightarrow}{\partial}_0 b \equiv a(\partial_0 b) - (\partial_0 a) b\,, \tag{4.172}$$

where a and b are arbitrary functions of t. We then have

$$\int d^3x\, e_{\vec{p}}(x) i \overset{\leftrightarrow}{\partial}_0 e_{\vec{p}\,\prime}(x) = (p^{0\prime} - p^0) \underbrace{\int d^3x\, e_{\vec{p}}(x) e_{\vec{p}\,\prime}(x)}_{e^{-2ip^0 t}\delta_{\vec{p},-\vec{p}\,'}/2p^0} = 0 \tag{4.173}$$

where the delta function forced $p^{0\prime} = p^0$. With $\overset{\leftrightarrow}{\partial}_0$, the orthonormality (4.170) also becomes simpler:

$$\int d^3x\, e^*_{\vec{p}}(x) i \overset{\leftrightarrow}{\partial}_0 e_{\vec{p}'}(x) = (p^{0\prime} + p^0) \underbrace{\int d^3x\, e^*_{\vec{p}}(x) e_{\vec{p}'}(x)}_{\delta_{\vec{p},\vec{p}'}/2p^0} = \delta_{\vec{p},\vec{p}'}\,. \tag{4.174}$$

Taking complex conjugate of these relations, we have a set of orthonormality relations given by

$$\boxed{\begin{array}{lll} \int d^3x\, e^*_{\vec{p}}(x) i \overset{\leftrightarrow}{\partial}_0 e_{\vec{p}'}(x) &=& \delta_{\vec{p},\vec{p}'}\,, \\ \int d^3x\, e_{\vec{p}}(x) i \overset{\leftrightarrow}{\partial}_0 e_{\vec{p}'}(x) &=& 0\,, \end{array} \qquad \begin{array}{lll} \int d^3x\, e_{\vec{p}}(x) i \overset{\leftrightarrow}{\partial}_0 e^*_{\vec{p}'}(x) &=& -\delta_{\vec{p},\vec{p}'}\,, \\ \int d^3x\, e^*_{\vec{p}}(x) i \overset{\leftrightarrow}{\partial}_0 e^*_{\vec{p}'}(x) &=& 0\,. \end{array}} \tag{4.175}$$

In the classical (i.e. non-quantized) Klein-Gordon theory, the probability density is given by $j_0 = \phi^* i \overset{\leftrightarrow}{\partial}_0 \phi$ (2.41). Thus, the normal-mode functions $e_{\vec{p}}(x)$ are normalized to be unit probability in the volume V as defined classically.

These relations allow us to express $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$ in terms of ϕ and π: using the expansion (4.167),

$$\int d^3x\, e^*_{\vec{p}}(x) i \overset{\leftrightarrow}{\partial}_0 \phi(x) = \sum_{\vec{p}'} \Big[a_{\vec{p}'} \underbrace{\int d^3x\, e^*_{\vec{p}} i \overset{\leftrightarrow}{\partial}_0 e_{\vec{p}'}}_{\delta_{\vec{p},\vec{p}'}} + a^\dagger_{\vec{p}'} \underbrace{\int d^3x\, e^*_{\vec{p}} i \overset{\leftrightarrow}{\partial}_0 e^*_{\vec{p}'}}_{0} \Big] = a_{\vec{p}}\,, \tag{4.176}$$

where we have used the linearity of the operator $\overset{\leftrightarrow}{\partial}_0$ to take the sum over momentum out of the integral. Writing out the right-hand side explicitly

$$a_{\vec{p}} = \int d^3x\, e^*_{\vec{p}}(x) \big[i\dot{\phi} + p^0 \phi(x) \big] = \int d^3x\, e^*_{\vec{p}}(x) \big[p^0 \phi(x) + i\pi(x) \big]\,. \tag{4.177}$$

Taking the hermitian conjugate of this relation, and noting that ϕ and π are hermitian, we get

$$a^\dagger_{\vec{p}} = \int d^3x\, e_{\vec{p}}(x) \big[p^0 \phi(x) - i\pi(x) \big]\,. \tag{4.178}$$

Note that the relations (4.177) and (4.178) are valid for any t.

Then, the commutation relations among $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$ follow directly from those among ϕ and π (4.133): with the notation

$$x \equiv (t, \vec{x}), \quad x' \equiv (t, \vec{x}') \qquad \text{(equal time)} \tag{4.179}$$

we have

$$[a_{\vec{p}}, a^\dagger_{\vec{p}'}] \;=\; \Big[\int d^3x\, e^*_{\vec{p}}(x) \big(p^0 \phi(x) + i\pi(x) \big)\,, \int d^3x' e_{\vec{p}'}(x') \big(p^{0\prime} \phi(x') - i\pi(x') \big) \Big]$$

$$
\begin{aligned}
&= \int d^3x \int d^3x' e^*_{\vec{p}}(x) e_{\vec{p}'}(x') \underbrace{\left[p^0\phi(x) + i\pi(x)\,, p^{0\prime}\phi(x') - i\pi(x')\right]}_{\substack{ip^{0\prime}\ \underbrace{[\pi(x),\phi(x')]}_{-i\delta^3(\vec{x}-\vec{x}')} \\ -ip^0\ \underbrace{[\phi(x),\pi(x')]}_{i\delta^3(\vec{x}-\vec{x}')}}} \\
&= \underbrace{\int d^3x\, e^*_{\vec{p}}(x) e_{\vec{p}'}(x)}_{\delta_{\vec{p},\vec{p}'}/2p^0}\, p^0 + \underbrace{\int d^3x\, e^*_{\vec{p}}(x) e_{\vec{p}'}(x)}_{\delta_{\vec{p},\vec{p}'}/2p^0 = \delta_{\vec{p},\vec{p}'}/2p^{0\prime}}\, p^{0\prime} \\
&= \delta_{\vec{p},\vec{p}'}\,. \qquad (4.180)
\end{aligned}
$$

Similarly, we can see that $[a_{\vec{p}}, a_{\vec{p}'}]$ vanishes, which results from the two terms with delta functions in (4.180) cancelling each other. Then, its hermitian conjugate $[a^\dagger_{\vec{p}}, a^\dagger_{\vec{p}'}]$ also vanishes. Thus, the commutation relations of ϕ and π resulted in those of $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$ given by

$$
\boxed{\begin{array}{c} [a_{\vec{p}}, a^\dagger_{\vec{p}'}] = \delta_{\vec{p},\vec{p}'} \\ [a_{\vec{p}}, a_{\vec{p}'}] = [a^\dagger_{\vec{p}}, a^\dagger_{\vec{p}'}] = 0 \end{array}}\,. \qquad (4.181)
$$

which are exactly the commutation relations for the annihilation and creation operators for a set of independent harmonic oscillators where each harmonic oscillator is labeled by $\vec{p}$; namely, for a given $\vec{p}$ we have $[a, a^\dagger] = 1$, and operators belonging to different harmonic oscillators commute. Then by the same argument as in the case of a single harmonic oscillator (4.28), this commutation relation between a and $a^\dagger$ singlehandedly leads to the fact that $a^\dagger$ raises the eigenvalue of the number operator $N = a^\dagger a$ by one and a lowers by one, thus justifying the interpretation that they are indeed creation and annihilation operators. The commutation relations among ϕ and π (4.133) can be re-derived from those of $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$, and it is left as an exercise.

Exercise 4.2 *Quantization conditions of hermitian spin-0 field.*
Start from the commutation relations among creation and annihilation operators (4.181) and derive those among fields (4.133) using the momentum expansions.

Total energy and momentum

We have seen that a hermitian field $\phi(t, \vec{x})$ that satisfies the Klein-Gordon equation can be regarded as a set of harmonic oscillators each labeled by $\vec{p}$ and with associated pair of annihilation and creation operators $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$. Then each harmonic oscillator or normal mode can have an integer number of quanta each of which we regard as a particle with momentum $\vec{p}$ and energy $p^0 = \sqrt{\vec{p}^{\,2} + m^2}$. If this picture is correct, then the total energy and momentum should be simply the sum of energies and momenta of the particles in each normal modes summed over all normal modes. We will now see that it is indeed the case.

First, let's examine in what kind of space the operators such as $a_{\vec{p}}$, ϕ, π, etc. are acting. This is in fact the Hilbert space we are dealing with. Since each normal mode is an independent harmonic oscillator, there is a number operator for each $\vec{p}$:

$$N_{\vec{p}} \equiv a^\dagger_{\vec{p}} a_{\vec{p}}\,, \tag{4.182}$$

and the corresponding set of eigenstates denoted as $|n_{\vec{p}}\rangle_{\vec{p}}$:

$$N_{\vec{p}}|n_{\vec{p}}\rangle_{\vec{p}} = n_{\vec{p}}|n_{\vec{p}}\rangle_{\vec{p}}\,. \tag{4.183}$$

Namely, $|n_{\vec{p}}\rangle_{\vec{p}}$ is a state of the harmonic oscillator labeled by $\vec{p}$ that contains $n_{\vec{p}}$ quanta. Then just as in the case of the single harmonic oscillator (4.34), the state $|n_{\vec{p}}\rangle_{\vec{p}}$ can be constructed by repeatedly applying $a^\dagger_{\vec{p}}$ on the ground state of the oscillator $|0\rangle_{\vec{p}}$:

$$|n_{\vec{p}}\rangle_{\vec{p}} = \frac{(a^\dagger_{\vec{p}})^{n_{\vec{p}}}}{\sqrt{n_{\vec{p}}!}}|0\rangle_{\vec{p}} \tag{4.184}$$

Let $|\{n_{\vec{p}}\}\rangle$ be the overall state with $n_{\vec{p}}$ quanta in the harmonic oscillator labeled by $\vec{p}$ where $\vec{p}$ runs over all possible values. It is the direct product of the states of all harmonic oscillators:

$$|\{n_{\vec{p}}\}\rangle = \prod_{\vec{p}} |n_{\vec{p}}\rangle_{\vec{p}} = \left[\prod_{\vec{p}} \frac{(a^\dagger_{\vec{p}})^{n_{\vec{p}}}}{\sqrt{n_{\vec{p}}!}}\right]|0\rangle\,, \tag{4.185}$$

where the vacuum state $|0\rangle$ is the product of all ground states:

$$|0\rangle \stackrel{\text{def}}{\equiv} \prod_{\vec{p}} |0\rangle_{\vec{p}}\,. \tag{4.186}$$

Since each $|0\rangle_{\vec{p}}$ is normalized, the over-all vacuum is also normalized:

$$\langle 0|0\rangle = 1\,. \tag{4.187}$$

The Hilbert space is then spanned by the states $|\{n_{\vec{p}}\}\rangle$ with all possible combinations of $\{n_{\vec{p}}\}$. It is a very large space. When there are more than one types of particles, then the space should be extended even larger: the basis states in this case are $|\{n_{r,\vec{p}}\}\rangle$, where $n_{r,\vec{p}}$ is the number of quanta in the normal mode corresponding to the particle type r and momentum $\vec{p}$. These states with definite numbers of quanta in each particle type and momentum are called the *Fock states*. For now, we assume that there is only one type of particle.

What we would like to show is then, for the operator P^μ that represents the total energy-momentum, its action on a Fock state $|\{n_{\vec{p}}\}\rangle$ is

$$P^\mu|\{n_{\vec{p}}\}\rangle = \Big(\sum_{\vec{p}} n_{\vec{p}}\, p^\mu\Big)|\{n_{\vec{p}}\}\rangle\,, \tag{4.188}$$

where $p^\mu \equiv (p^0, \vec{p})$ with $p^0 = \sqrt{\vec{p}^{\,2} + m^2}$; namely, the total energy-momentum 4-vector becomes the sum of those for all the existing particles in the universe. This would be accomplished if the total energy-momentum operator can be written as

$$P^\mu = \sum_{\vec{p}} p^\mu N_{\vec{p}}\,, \tag{4.189}$$

where $N_{\vec{p}}$ is the number operator given by (4.182). Now what should we take for the energy-momentum operator P^μ to start with? In the case of classical field, we had the conserved 4-vector (4.102) obtained from the Noether current of space-time translation, which is a good candidate to use after regarding the fields in the expressions as operators. In fact, the Hamiltonian of the quantized Klein-Gordon field (4.124) was constructed exactly in such manner and it was shown that such Hamiltonian indeed leads to the Klein-Gordon equation through Heisenberg's equations of motion. Thus, we will start from the Hamiltonian density (4.124), regard the fields as operators, and express the total Hamiltonian $H = \int d^3\mathcal{H}$ in terms of $a_{\vec{p}}$'s and $a^\dagger_{\vec{p}}$'s.

Let's first use a partial integration and the Klein-Gordon equation to simplify the expression (4.124) for H:

$$\begin{aligned} H &= \int d^3x \frac{1}{2}\Big(\pi^2 + \underbrace{(\vec{\nabla}\phi)^2}_{\underbrace{\vec{\nabla}\cdot(\phi\vec{\nabla}\phi)}_{\to 0} - \phi\underbrace{\nabla^2\phi}_{\ddot{\phi} + m\cancel{^2}\phi \,:\, \text{by the K-G eq.}}} + m\cancel{^2}\phi^2\Big) = \frac{1}{2}\int d^3x\Big(\underbrace{\pi^2}_{\dot{\phi}^2} - \phi\ddot{\phi}\Big) \\ &= \frac{1}{2}\int d^3x \phi\, i\overleftrightarrow{\partial}_0 (i\partial_0\phi)\,, \end{aligned} \tag{4.190}$$

where we have used the fact that the surface term vanishes when integrated over all space:

$$\int d^3x \vec{\nabla}\cdot(\phi\vec{\nabla}\phi) = \int_{\text{boundary}} d\vec{a}\cdot(\phi\vec{\nabla}\phi) = 0\,. \tag{4.191}$$

Then, substituting the expansions (4.167) and using the orthonormality relations (4.175), we have

$$\begin{aligned} H &= \frac{1}{2}\int d^3x \Big[\sum_{\vec{p}}(a_{\vec{p}}\, e_{\vec{p}} + a^\dagger_{\vec{p}}\, e^*_{\vec{p}})\Big] i\overleftrightarrow{\partial}_0 \Big[\sum_{\vec{p}'} p^{0\prime}(a_{\vec{p}'}\, e_{\vec{p}'} - a^\dagger_{\vec{p}'}\, e^*_{\vec{p}'})\Big] \\ &= \frac{1}{2}\int d^3x \sum_{\vec{p},\vec{p}'} p^{0\prime}\Big(a_{\vec{p}}\, a_{\vec{p}'}\underbrace{e_{\vec{p}}\, i\overleftrightarrow{\partial}_0 e_{\vec{p}'}}_{\to 0} - a_{\vec{p}}\, a^\dagger_{\vec{p}'}\underbrace{e_{\vec{p}}\, i\overleftrightarrow{\partial}_0 e^*_{\vec{p}'}}_{\to\, -\delta_{\vec{p},\vec{p}'}} \\ &\qquad + a^\dagger_{\vec{p}}\, a_{\vec{p}'}\underbrace{e^*_{\vec{p}}\, i\overleftrightarrow{\partial}_0 e_{\vec{p}'}}_{\to\, \delta_{\vec{p},\vec{p}'}} - a^\dagger_{\vec{p}}\, a^\dagger_{\vec{p}'}\underbrace{e^*_{\vec{p}}\, i\overleftrightarrow{\partial}_0 e^*_{\vec{p}'}}_{\to 0}\Big) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{\vec{p}} p^0 (\underbrace{a_{\vec{p}}\, a^\dagger_{\vec{p}}}_{a^\dagger_{\vec{p}} a_{\vec{p}} + 1} + a^\dagger_{\vec{p}} a_{\vec{p}}) \\
&= \sum_{\vec{p}} p^0 \Big(a^\dagger_{\vec{p}} a_{\vec{p}} + \frac{1}{2} \Big) \,. \qquad (4.192)
\end{aligned}
$$

Thus, the total Hamiltonian is now a simple function of the number operators $N_{\vec{p}} = a^\dagger_{\vec{p}} a_{\vec{p}}$. This is almost exactly what we wanted except the term $\sum_{\vec{p}} p^0/2$ which apparently is infinity. It is the energy of the state with $n_{\vec{p}} = 0$ for all $\vec{p}$, namely, the energy of the vacuum $|0\rangle$. If the object we are dealing with is a lattice of atoms, then there is a maximum value for $|\vec{p}|$ where the wavelength becomes smaller than the lattice spacing and waves stop propagating. In that case, this ground state energy is finite. In our case of the Klein-Gordon field, theory may break down at very high energy, and there may be a natural cutoff that would make the ground state energy finite. In any case, such a constant offset of energy does not have observable effects on the phenomena we are interested in, and thus we choose to simply discard it. Then, our Hamiltonian is

$$
H = \sum_{\vec{p}} p^0 a^\dagger_{\vec{p}} a_{\vec{p}} \qquad (4.193)
$$

which gives the time component of (4.189).

We will encounter similar situations where we want to discard constant offsets corresponding to vacuum expectation values; thus, it is convenient to introduce a procedure that forces the vacuum expectation value to be zero. This can be accomplished by the so-called *normal ordering* denoted by ‘: :’. Whatever is inside the colons is a polynomial of annihilation and creation operators, and the normal ordering reorders each term such that all annihilation operators (a’s) are to the right of all creation operators ($a^\dagger$’s). For example,

$$
\begin{aligned}
: a_{\vec{p}} a^\dagger_{\vec{p}'} : &= a^\dagger_{\vec{p}'} a_{\vec{p}} \,, \qquad : a^\dagger_{\vec{p}'} a_{\vec{p}} : = a^\dagger_{\vec{p}'} a_{\vec{p}} \,, \\
: a_{\vec{p}_1} a^\dagger_{\vec{p}_2} a^\dagger_{\vec{p}_3} : &= a^\dagger_{\vec{p}_2} a^\dagger_{\vec{p}_3} a_{\vec{p}_1} \,, \qquad \text{etc.}
\end{aligned} \qquad (4.194)
$$

Recalling that $a_{\vec{p}}$ annihilates the vacuum state [see (4.32)], we have

$$
a_{\vec{p}}|0\rangle = 0 \quad \rightarrow \quad \langle 0| a^\dagger_{\vec{p}} = 0 \,, \qquad (4.195)
$$

which means that the vacuum expectation value of any normal ordered product of $a_{\vec{p}}$’s and $a^\dagger_{\vec{p}}$’s is zero:

$$
\langle 0| : \cdots : |0\rangle = 0 \,, \qquad (4.196)
$$

because there is at least one $a_{\vec{p}}$ directly to the left of $|0\rangle$ or at least one $a^\dagger_{\vec{p}}$ directly to the right of $\langle 0|$. Since $a_{\vec{p}}$’s commute among themselves as do $a^\dagger_{\vec{p}}$’s, the normal

ordering uniquely defines the resulting operator. One has to be careful, however, not to use commutation relations *before* the normal ordering is performed since it defeats the very purpose of the normal ordering, namely to discard the constant term arising from the commutator:

$$: a_{\vec{p}} a^\dagger_{\vec{p}'} : = : \delta_{\vec{p},\vec{p}'} + a^\dagger_{\vec{p}'} a_{\vec{p}} : = \delta_{\vec{p},\vec{p}'} + a^\dagger_{\vec{p}'} a_{\vec{p}} \qquad \text{(don't do this)}\,. \tag{4.197}$$

Normal ordering is understood to be simple reordering of a's and $a^\dagger$'s. Then, the total Hamiltonian can be written using the normal ordering symbol as

$$H = : \int d^3x \frac{1}{2}\Big(\pi^2 + (\vec{\nabla}\phi)^2 + m^2\phi^2\Big) : = \sum_{\vec{p}} p^0\, N_{\vec{p}}\,. \tag{4.198}$$

From now on, when we deal with terms in a Hamiltonian or Lagrangian, normal ordering is implicitly assumed.

For the total momentum, we start from $\vec{P}$ given by (4.104), regard the fields as operators, and use the momentum expansion to write it in terms of a's and $a^\dagger$'s. In doing so, it is convenient to write it using $\overset{\leftrightarrow}{\partial}_0$. First, we note that the implicit normal ordering means

$$: (\vec{\nabla}\pi)\phi : = : \phi(\vec{\nabla}\pi) : \ . \tag{4.199}$$

Then, the total momentum can be expressed as

$$\begin{aligned}
\vec{P} &= -\int d^3x\, \pi\vec{\nabla}\phi = -\frac{1}{2}\int d^3x(\pi\vec{\nabla}\phi + \underbrace{\pi\vec{\nabla}\phi}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \underbrace{\vec{\nabla}(\pi\phi)} - \underbrace{(\vec{\nabla}\pi)\phi} \\
&\qquad\qquad\qquad\qquad\qquad \text{surface term } \to 0 \quad \phi(\vec{\nabla}\pi) \\
&= -\frac{1}{2}\int d^3x(\dot{\phi}\vec{\nabla}\phi - \phi\vec{\nabla}\dot{\phi}) \\
&= \frac{1}{2}\int d^3x\, \phi\, i\overset{\leftrightarrow}{\partial}_0(-i\vec{\nabla}\phi)\,.
\end{aligned} \tag{4.200}$$

At this point, we notice the pattern: both H and $\vec{P}$ are written in the form

$$\frac{1}{2}\int d^3x \phi i\overset{\leftrightarrow}{\partial}_0 \mathcal{O}\phi\,, \tag{4.201}$$

where $\mathcal{O} = i\partial_0$ for H, and $\mathcal{O} = -i\vec{\nabla}$ for $\vec{P}$. Once written in this form, the rest is nearly identical to the Hamiltonian case (4.192), and the result is

$$\vec{P} = - : \int d^3x\, \pi\vec{\nabla}\phi : = \sum_{\vec{p}} \vec{p}\, N_{\vec{p}}\,. \tag{4.202}$$

Exercise 4.3 *Complete the derivation above.*

Note that in the case of classical fields there is no difference between $\pi\vec{\nabla}\phi$ and $(\nabla\phi)\pi$, but for the quantized fields they differ since π and ϕ do not commute. The normal ordering, however, resolves such ambiguity.

We have seen in the classical case that the quantities P^ν given by (4.102) are conserved. Now we are dealing with quantized fields, and a natural question is if the operators P^ν are indeed constants of motion. Time variations of operators are given by Heisenberg's equation of motion, and thus immediately we see that $P^0 = H$ is conserved since H commutes with itself. It is easy to see that $\vec{P}$ is also conserved:

$$-i\dot{\vec{P}} = [H, \vec{P}] = \Big[\sum_{\vec{p}} p^0\, N_{\vec{p}}\,, \sum_{\vec{p}} \vec{p}\, N_{\vec{p}}\Big] = 0\,, \tag{4.203}$$

since $N_{\vec{p}}$ commutes with all other $N_{\vec{p}}$'s including itself. Similarly, the components of total momentum $\vec{P}$ commute among themselves:

$$\Big[P^i, P^j\Big] = 0 \quad (i, j = 1, 2, 3)\,. \tag{4.204}$$

Thus, all the four components of P^μ commute among themselves, and thus can simultaneously good quantum numbers of a system.

Space-time translation

This is a good place to show that the operators P^ν given in general by (4.117) are the generators of space-time translations; namely, if $F(\phi(x), \pi(x))$ is an arbitrary function (namely, polynomial) of ϕ and $\pi \equiv \dot{\phi}$, then

$$e^{iP\cdot a} F(x) e^{-iP\cdot a} = F(x + a) \tag{4.205}$$

provided that ϕ satisfies the equation of motion, where the operator $F(x)$ is a function of $x = (t, \vec{x})$ through $\phi(x)$ and $\pi(x)$ and a^μ is a real 4-vector constant. We will assume the number of fields to be one, but the derivation can be readily extended to multiple fields.

We first evaluate the commutator of $\vec{P}$ and ϕ, π (all at equal time):

$$\begin{aligned}
[\vec{P}, \phi(x)] &= -\int d^3x' [\pi(\vec{x}')\, \overbrace{\vec{\nabla}'\phi(\vec{x}'), \phi(\vec{x})}^{\text{commute by (4.140)}}\,] \\
&= -\int d^3x'\, \underbrace{[\pi(\vec{x}'), \phi(\vec{x})]}_{-i\delta^3(\vec{x}' - \vec{x})}\, \vec{\nabla}'\phi(\vec{x}') \\
&= i\vec{\nabla}\phi(\vec{x})
\end{aligned} \tag{4.206}$$

and

$$\begin{aligned}[\vec{P}, \pi(x)] &= -\int d^3x' \, [\pi(\vec{x}')\vec{\nabla}'\phi(\vec{x}'), \pi(\vec{x})] \\ &= -\int d^3x' \, \pi(\vec{x}') \, [\vec{\nabla}'\phi(\vec{x}'), \pi(\vec{x})] \\ &= i\vec{\nabla}\pi(\vec{x}) \quad \text{by (4.141)}\,. \end{aligned} \tag{4.207}$$

Together with Heisenberg's equations of motion $-i\dot{\phi} = [P^0, \phi]$ and $-i\dot{\pi} = [P^0, \pi]$, we have

$$\begin{array}{l} [P^\mu, \phi(x)] = -i\partial^\mu\phi(x) \\ [P^\mu, \pi(x)] = -i\partial^\mu\pi(x) \end{array}\,. \tag{4.208}$$

Then, it is easy to show that, for an arbitrary polynomial of ϕ and π, $F(\phi, \pi)$,

$$\boxed{[P^\mu, F(\phi, \pi)] = -i\partial^\mu F(\phi, \pi)}\,. \tag{4.209}$$

Exercise 4.4 *Space-time translation operators.*
Use the property of the energy-momentum operator P^μ (4.208) and show that (4.209) holds for any $F(\phi(x), \pi(x))$ that can be expanded as

$$F(\phi, \pi) = \sum_{n,m=0}^{\infty} c_{n,m}\phi^n\pi^m \tag{4.210}$$

where $c_{n,m}$ are c-numbers.
(*hint: Prove and use*

$$[A, B^n] = B^{n-1}[A, B] + B^{n-2}[A, B]B + ... + [A, B]B^{n-1}.$$

Comment: note that the above can be extended to any number of different fields.)

For a small translation $a^\mu = \epsilon^\mu$, we have (to the first order in ϵ),

$$\begin{aligned} e^{iP\cdot\epsilon}F(x)e^{-iP\cdot\epsilon} &= (1 + iP\cdot\epsilon)F(1 - iP\cdot\epsilon) \\ &= F + i(P^\mu\epsilon_\mu F - FP^\mu\epsilon_\mu) \\ &= F + i\underbrace{[P^\mu, F]}_{-i\partial^\mu F}\epsilon_\mu \\ &= F(x) + (\partial^\mu F)\epsilon_\mu = F(x + \epsilon)\,. \end{aligned} \tag{4.211}$$

Applying again $e^{iP\cdot\epsilon}$ from the left and $e^{-iP\cdot\epsilon}$ from the right, and repeating n times total, we obtain

$$e^{iP\cdot n\epsilon}F(x)e^{-iP\cdot n\epsilon} = F(x + n\epsilon)\,. \tag{4.212}$$

Taking the limit $n \to \infty$ while keeping $a = n\epsilon$, we obtain (4.205).

Thus, we have shown that the total energy-momentum operator P^μ indeed generates space-time translation where the time translation is nothing but the time evolution according to Heisenberg's equation of motion. Note that we started from the definition of P^μ for a general field and used only the commutation relations between ϕ and π which happens to be valid for all bosons. We will see later that (4.205) also holds for spin-1/2 fields; it is in fact a general feature of quantum field theory.

4.4 Two scalar fields of the same mass

In the case of a hermitian scalar field, we had one pair of creation and annihilation operators, $a^\dagger_{\vec{p}}$ and $a_{\vec{p}}$, for a given momentum $\vec{p}$. Thus, it cannot have more than one degree of freedom for a given $\vec{p}$; namely, it cannot represent a particle with spin greater than zero, or a system of a particle and its antiparticle. We will now discuss a framework that can describe a spin-0 particle and its antiparticle that are distinct. It is accomplished by combining two hermitian Klein-Gordon fields with same mass into one non-hermitian field. The fact that the two fields have same mass introduces a symmetry to the combined Lagrangian, and the conserved quantity associated with the symmetry will turn out to be the 'charge' of the particles where a particle and an antiparticle have charges with opposite signs and same magnitude. A hermitian or non-hermitian field in quantum field theory corresponds respectively to a real or complex field in classical field theory. These terminologies are often used interchangeably.

Lagrangian
Consider two real fields with same mass m, $\phi_1(x)$ and $\phi_2(x)$, within the framework of classical field theory. The Lagrangian density for each is given by (4.121):

$$\mathcal{L}_k(\phi_k, \dot{\phi}_k) = \frac{1}{2}(\partial_\mu \phi_k \partial^\mu \phi_k - m^2 \phi_k^2) \qquad (k = 1, 2;\ \text{no sum over } k)\,. \tag{4.213}$$

Assuming there is no interaction, the Lagrangian of the system is just the sum of the two:

$$\begin{aligned} \mathcal{L}(\tilde{\phi}, \dot{\tilde{\phi}}) &= \mathcal{L}_1(\phi_1, \dot{\phi}_1) + \mathcal{L}_2(\phi_2, \dot{\phi}_2) \\ &= \frac{1}{2}\Big[\partial_\mu \phi_1 \partial^\mu \phi_1 + \partial_\mu \phi_2 \partial^\mu \phi_2 - m^2(\phi_1^2 + \phi_2^2)\Big] \end{aligned} \tag{4.214}$$

where $\tilde{\phi} = (\phi_1, \phi_2)$. The Euler-Lagrange equations of this Lagrangian indeed lead to correct equations of motion:

$$\underbrace{\frac{\partial \mathcal{L}}{\partial \phi_k}}_{-m^2\phi_k} = \partial_\mu \underbrace{\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_k)}}_{\partial^\mu \phi_k} \quad \rightarrow \quad (\partial_\mu \partial^\mu + m^2)\phi_k = 0 \qquad (k = 1, 2)\,. \tag{4.215}$$

The conjugate fields are, by the definition (4.91),

$$\pi_k \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}_k} = \dot{\phi}_k \qquad (k = 1, 2)\,. \tag{4.216}$$

The Hamiltonian then becomes just the sum of those of the two systems: with $\tilde{\pi} = (\pi_1, \pi_2)$, we have

$$\mathcal{H}(\tilde{\pi}, \tilde{\phi}) \equiv \sum_k \pi_k \dot{\phi}_k - \mathcal{L} = \sum_k (\pi_k \dot{\phi}_k - \mathcal{L}_k) = \sum_k \mathcal{H}_k(\pi_k, \phi_k)\,, \tag{4.217}$$

and

$$H = H_1 + H_2\,, \quad \text{with} \quad H_k = \int d^3x \mathcal{H}_k(\pi_k, \phi_k) \quad (k = 1, 2)\,. \tag{4.218}$$

Define two complex fields ϕ and $\phi^\dagger$ by

$$\boxed{\begin{array}{rcl} \phi & \stackrel{\text{def}}{\equiv} & \dfrac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \\ \phi^\dagger & \stackrel{\text{def}}{\equiv} & \dfrac{1}{\sqrt{2}}(\phi_1 - i\phi_2) \end{array}}\,. \tag{4.219}$$

Then, we have

$$\phi^\dagger \phi = \frac{1}{2}(\phi_1 - i\phi_2)(\phi_1 + i\phi_2) = \frac{1}{2}(\phi_1^2 + \phi_2^2) \tag{4.220}$$

and

$$\partial_\mu \phi^\dagger \partial^\mu \phi = \frac{1}{2}(\partial_\mu \phi_1 - i\partial_\mu \phi_2)(\partial^\mu \phi_1 + i\partial^\mu \phi_2) = \frac{1}{2}(\partial_\mu \phi_1 \partial^\mu \phi_1 + \partial_\mu \phi_2 \partial^\mu \phi_2)\,. \tag{4.221}$$

The Lagrangian of the system (4.214) can then be written as

$$\boxed{\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi}\,. \tag{4.222}$$

Since ϕ_1 and ϕ_2 both satisfy the Klein-Gordon equation, so do ϕ and $\phi^\dagger$:

$$(\partial^2 + m^2)\phi = 0\,, \qquad (\partial^2 + m^2)\phi^\dagger = 0\,, \tag{4.223}$$

which can also be obtained by *regarding ϕ and $\phi^\dagger$ as independent* in (4.222) and applying the Euler-Lagrange equation:

$$\underbrace{\frac{\partial \mathcal{L}}{\partial \phi}}_{-m^2\phi^\dagger} = \partial_\mu \underbrace{\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}}_{\partial^\mu \phi^\dagger} \quad \rightarrow \quad (\partial_\mu \partial^\mu + m^2)\phi^\dagger = 0 \tag{4.224}$$

$$\underbrace{\frac{\partial \mathcal{L}}{\partial \phi^\dagger}}_{-m^2\phi} = \partial_\mu \underbrace{\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\dagger)}}_{\partial^\mu \phi} \quad \rightarrow \quad (\partial_\mu \partial^\mu + m^2)\phi = 0\,. \tag{4.225}$$

Again regarding ϕ and $\phi^\dagger$ as independent, the corresponding conjugate fields, π and $\pi^\dagger$, are by definition given by

$$\pi \stackrel{\text{def}}{\equiv} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}^\dagger\,, \qquad \pi^\dagger \stackrel{\text{def}}{\equiv} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}\,, \tag{4.226}$$

or using (4.219) and (4.216)

$$\begin{aligned} \pi &= \frac{1}{\sqrt{2}}(\pi_1 - i\pi_2) \\ \pi^\dagger &= \frac{1}{\sqrt{2}}(\pi_1 + i\pi_2) \end{aligned} \tag{4.227}$$

Note that at this point, π and $\pi^\dagger$ are defined as the conjugate fields of ϕ and $\phi^\dagger$, respectively, and considered independent; the relations (4.219) and (4.227), however, are self-consistent when the symbol † is considered as complex conjugation or, in the case of quantized fields, hermitian conjugation. The correct Hamiltonian can be obtained also by regarding ϕ and $\phi^\dagger$ as independent:

$$\begin{aligned} \mathcal{H} &= \pi\dot{\phi} + \pi^\dagger\dot{\phi}^\dagger - \mathcal{L} \\ &= \frac{1}{2}\Big[(\pi_1 - i\pi_2)(\dot{\phi}_1 + i\dot{\phi}_2) + (\pi_1 + i\pi_2)(\dot{\phi}_1 - i\dot{\phi}_2)\Big] - \mathcal{L} \\ &= \sum_k \pi_k\dot{\phi}_k - \mathcal{L}\,. \end{aligned} \tag{4.228}$$

How could ϕ and $\phi^\dagger$ be considered independent without destroying the whole logical structure, since they seem to be clearly related by complex conjugation? Actually, there is a good reason why such procedure works. If we regard ϕ and $\phi^\dagger$ as independent, and π and $\pi^\dagger$ as independent, then the relations (4.219) and (4.227) formally constitute a canonical transformation generated by

$$F = \frac{1}{\sqrt{2}}\Big[\pi(\phi_1 + i\phi_2) + \pi^\dagger(\phi_1 - i\phi_2)\Big]\,, \tag{4.229}$$

where it can be readily verified that

$$\phi = \frac{\partial F}{\partial \pi}\,, \quad \phi^\dagger = \frac{\partial F}{\partial \pi^\dagger}\,, \quad \text{and} \quad \pi_k = \frac{\partial F}{\partial \phi_k} \quad (k = 1, 2) \tag{4.230}$$

reproduce (4.219) and (4.227). Thus, it is no surprise that the two 'coordinate systems' describe the identical dynamical system using the same $\mathcal{H}$ and $\mathcal{L}$ expressed in terms of appropriate fields.

Quantization

The quantization proceeds as usual. We regard ϕ_k and π_k as operator fields, and introduce the equal-time commutators given by

$$\begin{array}{rcl}[\phi_k(t,\vec{x}),\pi_{k'}(t,\vec{x}')] &=& i\delta_{kk'}\delta^3(\vec{x}-\vec{x}') \\ [\phi_k(t,\vec{x}),\phi_{k'}(t,\vec{x}')] &=& [\pi_k(t,\vec{x}),\pi_{k'}(t,\vec{x}')] = 0\,;\end{array} \tag{4.231}$$

namely, the same as (4.133) except that if ϕ or π belong to different fields, then they always commute. Heisenberg's equation of motion for ϕ_k is

$$-i\phi_k(x) = [H,\phi_k(x)] = [H_1+H_2,\phi_k(x)] = [H_k,\phi_k(x)] \tag{4.232}$$

which, together with the corresponding equation for π_k and following exactly the same derivation as (4.135) through (4.148), leads to the Klein-Gordon equation for ϕ_k:

$$(\partial^2+m^2)\phi_k = 0 \qquad (k=1,2)\,. \tag{4.233}$$

Then, the two fields ϕ_k can be momentum-expanded using the normal-mode functions $e_{\vec{p}}(x)$ (4.168) together with the creation and annihilation operators for each field:

$$\phi_k(x) = \sum_{\vec{p}}\left(a_{k\vec{p}}\,e_{\vec{p}}(x) + a^\dagger_{k\vec{p}}\,e^*_{\vec{p}}(x)\right) \qquad (k=1,2)\,. \tag{4.234}$$

Note that we can use the same $e_{\vec{p}}(x)$ for both fields since they have the same mass which results in the same value of p^0 for a given momentum $\vec{p}$. Then, as before, the commutation relations (4.231) lead to

$$\begin{array}{c}[a_{k\vec{p}},a^\dagger_{k'\vec{p}'}] = \delta_{kk'}\delta_{\vec{p},\vec{p}'} \\ [a_{k\vec{p}},a_{k'\vec{p}'}] = [a^\dagger_{k\vec{p}},a^\dagger_{k'\vec{p}'}] = 0\,.\end{array} \tag{4.235}$$

The equal-time commutation relations among ϕ_k and π_k (4.231) and the definitions of the non-hermitian fields (4.219) and (4.227) readily lead to

$$\boxed{\begin{array}{c}[\phi(t,\vec{x}),\pi(t,\vec{x}')] = i\delta^3(\vec{x}-\vec{x}') \\ [\phi^\dagger(t,\vec{x}),\pi^\dagger(t,\vec{x}')] = i\delta^3(\vec{x}-\vec{x}') \\ \text{all others} = 0\end{array}}\,. \tag{4.236}$$

Note that this is what one would get if ϕ and $\phi^\dagger$ (and π and $\pi^\dagger$) are regarded as independent fields and the standard quantization procedure is followed.

The momentum expansion of ϕ is obtained simply from those of ϕ_1 and ϕ_2:

$$\begin{aligned}\phi(x) &\equiv \frac{1}{\sqrt{2}}\Big(\phi_1(x)+i\phi_2(x)\Big)\\ &= \frac{1}{\sqrt{2}}\Big[\sum_{\vec{p}}(a_{1\vec{p}}\,e_{\vec{p}}+a_{1\vec{p}}^\dagger e_{\vec{p}}^*)+i\sum_{\vec{p}}(a_{2\vec{p}}\,e_{\vec{p}}+a_{2\vec{p}}^\dagger e_{\vec{p}}^*)\Big]\\ &= \sum_{\vec{p}}\Big[\frac{1}{\sqrt{2}}(a_{1\vec{p}}+ia_{2\vec{p}})e_{\vec{p}}+\frac{1}{\sqrt{2}}(a_{1\vec{p}}^\dagger+ia_{2\vec{p}}^\dagger)e_{\vec{p}}^*)\Big]\,. \end{aligned} \tag{4.237}$$

Defining a set of new operators by

$$\begin{aligned} a_{\vec{p}} &\stackrel{\text{def}}{\equiv} \frac{1}{\sqrt{2}}(a_{1\vec{p}}+ia_{2\vec{p}})\,,\\ b_{\vec{p}} &\stackrel{\text{def}}{\equiv} \frac{1}{\sqrt{2}}(a_{1\vec{p}}-ia_{2\vec{p}}) \quad\to\quad b_{\vec{p}}^\dagger=\frac{1}{\sqrt{2}}(a_{1\vec{p}}^\dagger+ia_{2\vec{p}}^\dagger)\,,\end{aligned} \tag{4.238}$$

the field ϕ can then be expanded as

$$\boxed{\phi(x)=\sum_{\vec{p}}\Big(a_{\vec{p}}\,e_{\vec{p}}(x)+b_{\vec{p}}^\dagger e_{\vec{p}}^*(x)\Big)}\,. \tag{4.239}$$

The commutation relations among a's and b's easily follow from those among a_k's:

$$\boxed{\begin{array}{c}[a_{\vec{p}},a_{\vec{p}'}^\dagger]=\delta_{\vec{p},\vec{p}'}\,,\quad [b_{\vec{p}},b_{\vec{p}'}^\dagger]=\delta_{\vec{p},\vec{p}'}\\ \text{all others}=0\end{array}}\,. \tag{4.240}$$

This indicates that $a_{\vec{p}}$ and $b_{\vec{p}}$ act as annihilation operators of some particles. In particular, $a_{\vec{p}}^\dagger a_{\vec{p}}$ would be the number operator of the particle 'a' and $b_{\vec{p}}^\dagger b_{\vec{p}}$ that of the particle 'b'.

Exercise 4.5 *Quantization of charged Klein-Gordon field.*
(*a*) *Derive the commutation relations among a's and b's (4.240) from those among a_1's and a_2's: (4.235).*
(*b*) *Derive the equal-time commutators among ϕ and π: (4.236) from those among ϕ_k and π_k (4.231).*

Following the same derivation as (4.190) through (4.193), the total Hamiltonian of the two fields, H_1 and H_2, can be written as

$$H_1=\sum_{\vec{p}}p^0a_{1\vec{p}}^\dagger a_{1\vec{p}}\,,\quad H_2=\sum_{\vec{p}}p^0a_{2\vec{p}}^\dagger a_{2\vec{p}}\,. \tag{4.241}$$

The total Hamiltonian of the system is simply the sum of the two:

$$H = H_1 + H_2 = \sum_{\vec{p}} p^0 (a^\dagger_{1\vec{p}} a_{1\vec{p}} + a^\dagger_{2\vec{p}} a_{2\vec{p}}) \,. \tag{4.242}$$

Noting that

$$\begin{aligned} a^\dagger_{\vec{p}} a_{\vec{p}} + b^\dagger_{\vec{p}} b_{\vec{p}} &= \frac{1}{2}(a^\dagger_{1\vec{p}} - i a^\dagger_{2\vec{p}})(a_{1\vec{p}} + i a_{2\vec{p}}) + \frac{1}{2}(a^\dagger_{1\vec{p}} + i a^\dagger_{2\vec{p}})(a_{1\vec{p}} - i a_{2\vec{p}}) \\ &= a^\dagger_{1\vec{p}} a_{1\vec{p}} + a^\dagger_{2\vec{p}} a_{2\vec{p}} \,, \end{aligned} \tag{4.243}$$

we have

$$\rightarrow \quad H = \sum_{\vec{p}} p^0 (a^\dagger_{\vec{p}} a_{\vec{p}} + b^\dagger_{\vec{p}} b_{\vec{p}}) \,, \tag{4.244}$$

which shows that the total energy is the sum of the energies of all 'a-particles' and 'b-particles'. Similarly, the total momentum defined by (4.117) is just the sum of those of two fields: using (4.202),

$$\begin{aligned} \vec{P} &= -\int d^3x \, \pi_k \vec{\nabla} \phi_k = \underbrace{-\int d^3x \, \pi_1 \vec{\nabla} \phi_1}_{\sum_{\vec{p}} \vec{p} \, a^\dagger_{1\vec{p}} a_{1\vec{p}}} \underbrace{-\int d^3x \, \pi_2 \vec{\nabla} \phi_2}_{\sum_{\vec{p}} \vec{p} \, a^\dagger_{2\vec{p}} a_{2\vec{p}}} \\ &= \sum_{\vec{p}} \vec{p} (a^\dagger_{1\vec{p}} a_{1\vec{p}} + a^\dagger_{2\vec{p}} a_{2\vec{p}}) \,. \end{aligned} \tag{4.245}$$

Then, (4.243) allows us to write it as

$$\vec{P} = \sum_{\vec{p}} \vec{p} (a^\dagger_{\vec{p}} a_{\vec{p}} + b^\dagger_{\vec{p}} b_{\vec{p}}) \,, \tag{4.246}$$

which shows that the total momentum is the sum of the momenta of all a-particles and b-particles. Next, we will see that the two types of particles have opposite 'charge', and that the total 'charge' of the universe is conserved.

Noether current of internal symmetry

We have seen that the symmetry under space-time translation led to the conservation of total energy and momentum. Here, we encounter a slightly different kind of conserved quantity arising from an internal symmetry which crept into the system because the two fields ϕ_1 and ϕ_2 have the same mass. We start from non-quantized fields.

We first note that the Lagrangian (4.222) is invariant under the phase rotation

$$\phi'(x) = e^{i\theta} \phi(x) \quad \rightarrow \quad \phi'^\dagger(x) = e^{-i\theta} \phi^\dagger(x) \,, \tag{4.247}$$

where θ is a real parameter and the space-time points are equal on the both sides of the equalities. In fact,

$$\mathcal{L}' \equiv \partial_\mu \phi'^\dagger \partial^\mu \phi' - m^2 \phi'^\dagger \phi' = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi = \mathcal{L}\,. \tag{4.248}$$

Note that ϕ and $\phi^\dagger$ are not treated as independent in this procedure; indeed, it is critical that $\phi^\dagger$ is rotated by a phase exactly opposite to that of ϕ. In terms of ϕ_1 and ϕ_2, the phase rotation is written as

$$\begin{aligned}
\phi' &\equiv \frac{1}{\sqrt{2}}(\phi_1' + i\phi_2') \\
&= e^{i\theta}\phi = (\cos\theta + i\sin\theta)\frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \\
&= \frac{1}{\sqrt{2}}\big[(\cos\theta\phi_1 - \sin\theta\phi_2) + i(\sin\theta\phi_1 + \cos\theta\phi_2)\big]\,.
\end{aligned} \tag{4.249}$$

Equating the real parts and the imaginary parts, respectively, we have

$$\begin{pmatrix} \phi_1' \\ \phi_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}\,, \tag{4.250}$$

which is a rotation in the space spanned by ϕ_1 and ϕ_2. Such transformations form a group called $SO(2)$ - a group formed by 2×2 *o*rthogonal matrices with det $= 1$ (*s*pecial). This group is then apparently equivalent to the group formed by simple phase rotations $e^{i\theta}$ which are 1×1 *u*nitary matrices - called $U(1)$. Thus, $SO(2)$ and $U(1)$ have the same group structure:

$$SO(2) \sim U(1)\,. \tag{4.251}$$

Let's try Noether's mathematical trick again to extract a conserved current associated with this symmetry. For the case of space-time translation, the Lagrangian density $\mathcal{L}$ changed its value under the transformation since it is a function of x through $\phi(x)$. This time, the value of $\mathcal{L}$ is unchanged under the transformation. For a small θ, the changes in the fields and their derivatives are

$$\begin{aligned}
\delta\phi = i\theta\phi\,, &\qquad \delta\phi^\dagger = -i\theta\phi^\dagger\,, \\
\delta(\partial_\mu\phi) = i\theta\partial_\mu\phi\,, &\quad \delta(\partial_\mu\phi^\dagger) = -i\theta\partial_\mu\phi^\dagger
\end{aligned} \tag{4.252}$$

Then, the change in $\mathcal{L}(\phi, \phi^\dagger, \partial_\mu\phi, \partial_\mu\phi^\dagger)$, which should be zero, can be written as

$$\delta\mathcal{L} = \underbrace{\frac{\partial\mathcal{L}}{\partial\phi}}_{\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\text{ by (4.224)}} \overbrace{i\theta\phi}^{\delta\phi} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \overbrace{i\theta\partial_\mu\phi}^{\delta(\partial_\mu\phi)} + \underbrace{\frac{\partial\mathcal{L}}{\partial\phi^\dagger}}_{\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\dagger)}\text{ by (4.225)}} \overbrace{-i\theta\phi^\dagger}^{\delta\phi^\dagger} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\dagger)} \overbrace{-i\theta\partial_\mu\phi^\dagger}^{\delta(\partial_\mu\phi^\dagger)}$$

$$= i\theta\Big[\underbrace{\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\phi+\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu\phi}_{\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\phi\right)}\underbrace{-\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\dagger)}\phi^\dagger-\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\dagger)}\partial_\mu\phi^\dagger}_{-\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\dagger)}\phi^\dagger\right)}\Big]$$

$$= i\theta\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\phi-\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\dagger)}\phi^\dagger\Big)=0\,. \tag{4.253}$$

Thus, we have a conserved current given by

$$\boxed{\partial_\mu J^\mu=0\,,\qquad J^\mu=i\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\dagger)}\phi^\dagger-\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\phi\Big)\,,} \tag{4.254}$$

where the factor 'i' is arbitrary at this point. The conserved quantity, which we generically call 'charge', is the space integral of the time component:

$$\boxed{Q=\int d^3x J^0\,,\qquad \dot{Q}=0}\,. \tag{4.255}$$

So far, we have not assumed any particular functional form of the Lagrangian. For the Klein-Gordon Lagrangian (4.222), we have

$$J^\mu=i\Big[(\partial^\mu\phi)\phi^\dagger-(\partial^\mu\phi^\dagger)\phi\Big]=\phi^\dagger i\overleftrightarrow{\partial^\mu}\phi\,, \tag{4.256}$$

and

$$Q=\int d^3x J^0=\int d^3x\,\phi^\dagger i\overleftrightarrow{\partial^0}\phi\,, \tag{4.257}$$

which is exactly the same conserved current as (2.41) obtained earlier by directly constructing a conserved quantity from the Klein-Gordon equation. This time, the same conserved current was obtained from the $U(1)$ symmetry of the Lagrangian through Noether's trick.

At this point, we consider the fields that appear in Q as operators, and apply the momentum expansion to write it in terms of creation and annihilation operators. A simple calculation gives

$$Q=\int d^3x\,\phi^\dagger i\overleftrightarrow{\partial^0}\phi=\sum_{\vec{p}}(a^\dagger_{\vec{p}}a_{\vec{p}}-b^\dagger_{\vec{p}}b_{\vec{p}}) \tag{4.258}$$

or

$$Q=N_a-N_b \tag{4.259}$$

where

$$N_a\overset{\text{def}}{\equiv}\sum_{\vec{p}}a^\dagger_{\vec{p}}a_{\vec{p}}\,,\qquad N_b\overset{\text{def}}{\equiv}\sum_{\vec{p}}b^\dagger_{\vec{p}}b_{\vec{p}}\,, \tag{4.260}$$

are the operators that represents the total number of a-particles and b-particles, respectively. Then, the Heisenberg equation of motion tells us that Q is indeed a constant of motion:

$$\begin{aligned} -i\dot{Q} &= [H, Q] \\ &= [\sum_{\vec{p}} p^0 (a^\dagger_{\vec{p}} a_{\vec{p}} + b^\dagger_{\vec{p}} b_{\vec{p}}), \sum_{\vec{p}} (a^\dagger_{\vec{p}} a_{\vec{p}} - b^\dagger_{\vec{p}} b_{\vec{p}})\,] \\ &= 0\,, \end{aligned} \tag{4.261}$$

where we have used the fact that the number operators $a^\dagger_{\vec{p}} a_{\vec{p}}$ and $b^\dagger_{\vec{p}} b_{\vec{p}'}$ all commute among them.

Thus, the total number of a-particles minus that of b-particles is conserved. The expressions of the total charge (4.260) indicates that an a-particle carries charge $+1$ and a b-particle carries charge -1 regardless of momentum. Together with the fact that the two types of particles have the same mass, a natural interpretation is that they are antiparticles of each other. Which is plus and which is minus is arbitrary at this point, and so is the absolute value of the charge; we could have multiplied any constant to (4.256). This charge is not necessarily the electric charge. Electric charge is something that couples to photons, and we have not introduced any interaction yet. Here, the charge simply refers to the quantum numbers attached to the two types of particles. Then, why cannot ϕ_1 and ϕ_2 be interpreted as antiparticles of each other? This can be seen by writing the conserved charge Q in terms of a_k's and $a^\dagger_k$'s using (4.238). The result is

$$Q = i \sum_{\vec{p}} (a^\dagger_{1\vec{p}} a_{2\vec{p}} - a^\dagger_{2\vec{p}} a_{1\vec{p}}) \tag{4.262}$$

which does not contain number operators of particles 1 and 2. One should keep in mind, however, that the two interpretations - particles a and b, or particles 1 and 2 - are equivalent.

Exercise 4.6 *Conserved quantities of charged spin-0 field.*
(*a*) *Start from the Langrangian density in terms of ϕ and $\phi^\dagger$:*

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi\,,$$

and treat ϕ and $\phi^\dagger$ as independent variables to show that the hamiltonian is given by

$$H = \int d^3x \Big(\dot{\phi}^\dagger \dot{\phi} + \vec{\nabla}\phi^\dagger \cdot \vec{\nabla}\phi + m^2 \phi^\dagger \phi \Big)\,.$$

Compare with the result obtained by regarding ϕ_1 and ϕ_2 as independent variables and verify that they are identical. Up to here, regard the fields as non-quantized.
(*b*) *Express H above in terms of creation and annihilation operators a's and b's. Use the momentum expansions in terms of a's and b's and commutators among them. Do not refer to a_1's and a_2's.*

c) Similarly, show that the total charge operator Q given by

$$Q = \int d^3x J^0 = \int d^3x \phi^\dagger i \overleftrightarrow{\partial}^0 \phi \,,$$

can be written in terms of a's and b's as

$$Q = \sum_{\vec{p}} (a^\dagger_{\vec{p}} a_{\vec{p}} - b^\dagger_{\vec{p}} b_{\vec{p}}) \,.$$

When we studied the Klein-Gordon equation without field quantization, we had two problems: negative energy and negative probability. Now we can see how they are solved by quantizing the field. First, the energy is now representated by the Hamiltonian H and (4.244) shows that both a-particles and b-particles contribute positively. Second, the probability current is now interpreted as the charge current and thus there is no inconsistency in the total charge having a negative value.

For a real Klein-Gordon field, by definition we cannot rotate the phase of the field as done in the derivation of the conserved charge. Nothing stops us, however, from taking the field ϕ in (4.253) as real; then, the entire derivation that leads to the conserved current (4.253) is still valid. Using a real field ϕ in the conserved current, however, gives $J^\mu = 0$. Thus, a conserved charge current does not exist for a hermitian Klein-Gordon field, which is consistent with the interpretation that a hermitian Klein-Gordon field represents a particle that is antiparticle of itself. Then, what happened to the probability current? Can one define the probability to find a particle at a given point x at a given time t in the framework of quantum field theory? This brings us to the next topic: the state $\phi(x)|0\rangle$ and the important related issue of microscopic causality.

4.5 Microscopic causality

The state $\phi(x)|0\rangle$

Let's take a hermitian Klein-Gordon field $\phi(x)$ as the simplest example. In order to define the probability to find a particle at a given space-time point $x = (t, \vec{x})$, one needs a state where at time x^0 a particle is localized at position $\vec{x}$. Is $\phi(t, \vec{x})|0\rangle$ such a state? If that is the case, then the inner product of $\phi(t, \vec{x})|0\rangle$ with $\phi(t, \vec{y})|0\rangle$, where times are the same, should be zero unless $\vec{x} = \vec{y}$. We will see that such is not the case. For convenience, we first divide ϕ into the positive-energy part and the negative-energy part:

$$\phi(x) = \phi_a(x) + \phi^\dagger_a(x) \,, \tag{4.263}$$

with

$$\phi_a(x) \stackrel{\text{def}}{\equiv} \sum_{\vec{p}} a_{\vec{p}}\, e_{\vec{p}}(x) \,, \qquad \phi^\dagger_a(x) = \sum_{\vec{p}} a^\dagger_{\vec{p}}\, e^*_{\vec{p}}(x) \,. \tag{4.264}$$

Since $a_{\vec{p}}|0\rangle = 0$ and $\langle 0|a^\dagger_{\vec{p}} = 0$, ϕ_a facing the vacuum on its right or $\phi^\dagger_a$ facing the vacuum on its left annihilates the state. Thus, the inner product of $\phi(x)|0\rangle$ and $\phi(y)|0\rangle$ can be written as (using the hermiticity $\phi^\dagger = \phi$)

$$\begin{aligned}
\langle 0|\phi^\dagger(x)\phi(y)|0\rangle &= \langle 0|\phi(x)\phi(y)|0\rangle = \langle 0|\phi_a(x)\phi^\dagger_a(y)|0\rangle \\
&= \langle 0|\sum_{\vec{p}} a_{\vec{p}}\, e_{\vec{p}}(x) \sum_{\vec{p}'} a^\dagger_{\vec{p}'}\, e^*_{\vec{p}'}(y)|0\rangle \\
&= \langle 0|\sum_{\vec{p},\vec{p}'} \underbrace{a_{\vec{p}}\, a^\dagger_{\vec{p}'}}_{\substack{\delta_{\vec{p},\vec{p}'} + a^\dagger_{\vec{p}'} a_{\vec{p}} \\ \to 0}} e_{\vec{p}}(x) e^*_{\vec{p}'}(y)|0\rangle \\
&= \sum_{\vec{p}} e_{\vec{p}}(x) e^*_{\vec{p}}(y)\, \underbrace{\langle 0|0\rangle}_{1}\,, \qquad (4.265)
\end{aligned}$$

The last expression can be written as an integral over $\vec{p}$ using (4.153):

$$\begin{aligned}
\sum_{\vec{p}} e_{\vec{p}}(x) e^*_{\vec{p}}(y) &= \frac{V}{(2\pi)^3}\int d^3p \frac{e^{-ip\cdot x}}{\sqrt{2p^0 V}} \frac{e^{ip\cdot y}}{\sqrt{2p^0 V}} \\
&= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2p^0} e^{-ip\cdot(x-y)} = \Delta_+(x-y)\,, \qquad (4.266)
\end{aligned}$$

where we have defined the function $\Delta_+(z)$ to be

$$\Delta_+(z) \stackrel{\text{def}}{\equiv} \frac{1}{(2\pi)^3}\int \frac{d^3p}{2p^0} e^{-ip\cdot z}\,, \qquad (4.267)$$

where z is a real 4-vector and $p^0 \equiv \sqrt{\vec{p}^2 + m^2}$. Thus, in terms of this function, we have

$$\langle 0|\phi(x)\phi(y)|0\rangle = \Delta_+(x-y)\,. \qquad (4.268)$$

Up to this point, x and y are arbitrary. Now let's assume $x^0 = y^0$ (equal-time). Then, the time component of $z \equiv x - y$ becomes zero and

$$\begin{aligned}
\Delta_+(0,\vec{z}) &= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2p^0} e^{i\vec{p}\cdot\vec{z}} \\
&= \frac{1}{(2\pi)^3}\int \frac{2\pi p^2 dp\, d\cos\theta}{2\sqrt{p^2+m^2}} e^{ipr\cos\theta} \qquad (4.269)
\end{aligned}$$

where $p \equiv |\vec{p}|$ and $r \equiv |\vec{z}|$. This looks awful, but actually it is a well-defined real function of r and can be expressed in terms of the modified Bessel function $K_1(z)$

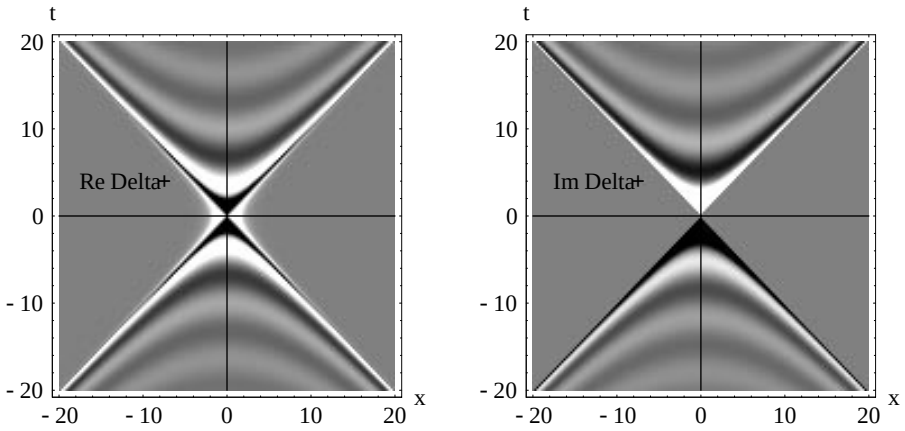

Figure 4.2: The real and imaginary parts of the function $\Delta_+(x)$. The predominant gray tone in the space-like region represents zero, darker is negative and lighter is positive. Units of x and t are $1/m$. The function is real and positive in the space-like region, and in general complex in the time-like region.

which is a solution of the differential equation $z^2X'' + zX' - (z^2 + n^2)X = 0$ with $n = 1$:

$$\Delta_+(0,\vec{z}) = \frac{m}{4\pi^2 r} K_1(mr) \;\overset{r\to\infty}{\sim}\; \frac{\sqrt{m}}{(2\pi r)^{3/2}}\, e^{-mr}\,. \tag{4.270}$$

This is a rapidly decreasing function of r, but the important fact is that it is not zero. Thus, the state $\phi(t,\vec{x})|0\rangle$ cannot be interpreted as the state in which a particle is localized at $(t,\vec{x})$; it is sharply peaked at that point, but there is some spill-over. A profile of the function $\Delta_+(x)$ is shown in Figure 4.2. For completeness, the function $\Delta_+(x)$ in the entire region is given by

$$(x^0 > 0)\;\; \Delta_+(x) = \begin{cases} \dfrac{m}{8\pi\sqrt{x^2}}\Big[Y_1(m\sqrt{x^2}) + iJ_1(m\sqrt{x^2})\Big] & (x^2 > 0) \\ \dfrac{m}{4\pi^2\sqrt{-x^2}} K_1(m\sqrt{-x^2}) & (x^2 < 0) \end{cases} \tag{4.271}$$

$$(x^0 < 0)\;\; : \text{use} \quad \Delta_+(-x) = \Delta_+^*(x) \quad [\text{follows directly from (4.267)}]\,,$$

where J_1 and Y_1 are the standard Bessel functions that are solutions of $z^2X'' + zX' + (z^2 - n^2)X = 0$ with $n = 1$. The function $\Delta_+(x)$ is real and positive for $x^2 < 0$, and complex and oscillatory for $x^2 > 0$.

Let's step back a little and examine more carefully what we have here. First, we note that the norm of the state $\phi(x)|0\rangle$ is infinity:

$$\langle 0|\phi^\dagger(x)\,\phi(x)|0\rangle = \Delta_+(0) = \int \frac{d^3p}{(2\pi)^3 2p^0} = \infty\,. \tag{4.272}$$

Undaunted, let's proceed. We are in the Heisenberg picture, thus the states are constant over time and the operators carry all time dependences. The state $\phi(t,\vec{x})|0\rangle$ then represents a particle nearly localized at $(t,\vec{x})$, but through the time-varying operators, it also represents the entire history of the evolution from the infinite past to the infinite future. Then what is the inner product of two such states $\phi(x)|0\rangle$ and $\phi(y)|0\rangle$? It may be easier to visualize it if we move to the Schrödinger picture. For simplicity, assume $y^0 = 0$ and $x^0 > y^0$. At $t = y^0 = 0$, states and operators in the two pictures are taken to be identical. Namely, the operators and states in the Schrödinger picture are defined by

$$\phi_S(\vec{x}) \equiv \phi(0,\vec{x})\,, \quad \phi_S(\vec{y}) \equiv \phi(0,\vec{y})\,, \tag{4.273}$$

$$|\vec{x}\rangle_S \equiv \phi_S(\vec{x})|0\rangle\,, \quad |\vec{y}\rangle_S \equiv \phi_S(\vec{y})|0\rangle\,. \tag{4.274}$$

where the subscript 'S' indicates the Schrödinger picture. The time evolution of the state $|\vec{y}\rangle_S$ is given by

$$|t,\vec{y}\rangle_S = e^{-iHt}|\vec{y}\rangle_S\,. \tag{4.275}$$

Using the space-time translation formula (4.205), $\phi(x^0,\vec{x})$ is related to $\phi(0,\vec{x})$ by

$$\phi(x^0,\vec{x}) = e^{iHx^0}\underbrace{\phi(0,\vec{x})}_{\phi_S(\vec{x})}e^{-iHx^0}\,. \tag{4.276}$$

Then, the inner product $\langle 0|\phi^\dagger(x)\phi(y)|0\rangle$ can be written as (with $y^0 = 0$)

$$\begin{aligned}\langle 0|\phi^\dagger(x^0,\vec{x})\phi(y^0,\vec{y})|0\rangle &= \underbrace{\langle 0|e^{iHx^0}}_{\langle 0|}\phi_S^\dagger(\vec{x})\underbrace{e^{-iHx^0}\overbrace{\phi(0,\vec{y})|0\rangle}^{|0,\vec{y}\rangle_S}}_{|x^0,\vec{y}\rangle_S}\\ &= {}_S\langle\vec{x}\,|x^0,\vec{y}\rangle_S\,,\end{aligned} \tag{4.277}$$

where we have used $H = 0$ for the vacuum (H is normal ordered), which leads to

$$e^{-iHx^0}|0\rangle_S = |0\rangle_S\,. \tag{4.278}$$

If we interpret the state $|0,\vec{y}\rangle_S$ as a creation of a particle at position $\vec{y}$ at $t = 0$, then the state $|x^0,\vec{y}\rangle_S$ is its time-evolved state at time x^0. The inner product above would then be the amplitude to find the particle at position $\vec{x}$. Thus, $\langle 0|\phi^\dagger(x)\phi(y)|0\rangle$ can be loosely interpreted as the amplitude of a particle created at $(y^0,\vec{y})$ to propagate to $(x^0,\vec{x})$.

Then, what do we make of the fact that the inner product is non-zero even if the separation between x and y are space-like (namely, cannot be reached from one to the other by a velocity less than the speed of light)? This apparent paradox arises

since the state $\phi(x)|0\rangle$ is actually not a state a particle is localized at $\vec{x}$ at time t. Thus, the non-zero amplitude outside the light cone does not mean that a particle can actually move faster than speed of light; rather, if we interpret $\phi(x)|0\rangle$ as the state where a particle is created at x, then we have to accept that there is non-zero 'propagation' outside the light cone. It is largely a matter of terminology.

Causality

The 'propagation' outside of the light cone that we have just seen is still alarming. Does it violate the principle of special relativity which states that any physical event at x should not affect another physical event at y if the x and y are outside the light cone of each other? Physical events are detected through observables. A hermitian field $\phi(x)$ can be viewed as an observable; in fact, it was introduced by quantizing the actual string motion. Then, causality requires that two observables $\phi(x)$ and $\phi(y)$ commute if the separation of x and y is space-like; namely,

$$[\phi(x), \phi(y)] = 0\,, \quad \text{if } (x-y)^2 < 0 \quad (?) \tag{4.279}$$

Note that the quantization condition (4.133) we had before, $[\phi(t,\vec{x}), \phi(t,\vec{y})] = 0$, is an *equal-time* commutator; now, x^0 and y^0 are in general different.

First, divide ϕ into the negative-energy and positive-energy parts as in (4.263). Since $a_{\vec{p}}$'s commute with $a_{\vec{p}}$'s and $a^\dagger_{\vec{p}}$'s commute with $a^\dagger_{\vec{p}}$'s,

$$[\phi_a(x), \phi_a(y)] = [\phi_a^\dagger(x), \phi_a^\dagger(y)] = 0\,, \tag{4.280}$$

for any x and y. Then, the commutator (4.279) can be written as

$$\begin{aligned}
[\phi(x), \phi(y)] &= [\phi_a(x) + \phi_a^\dagger(x)\,, \phi_a(y) + \phi_a^\dagger(y)] \\
&= [\phi_a(x), \phi_a^\dagger(y)] + \underbrace{[\phi_a^\dagger(x), \phi_a(y)]}_{-[\phi_a(y), \phi_a^\dagger(x)]} \\
&= [\phi_a(x), \phi_a^\dagger(y)] - (x \leftrightarrow y)\,.
\end{aligned} \tag{4.281}$$

Using the expansion (4.264) and the definition of the function Δ_+ (4.266), we have

$$\begin{aligned}
[\phi_a(x), \phi_a^\dagger(y)] &= [\sum_{\vec{p}} a_{\vec{p}}\, e_{\vec{p}}(x)\,, \sum_{\vec{p}'} a^\dagger_{\vec{p}'} e^*_{\vec{p}'}(y)] \\
&= \sum_{\vec{p},\vec{p}'} e_{\vec{p}}(x) e^*_{\vec{p}'}(y) \underbrace{[a_{\vec{p}}\,, a^\dagger_{\vec{p}'}]}_{\delta_{\vec{p},\vec{p}'}} \\
&= \sum_{\vec{p}} e_{\vec{p}}(x) e^*_{\vec{p}}(y) = \Delta_+(x-y)\,.
\end{aligned} \tag{4.282}$$

Thus,

$$[\phi(x), \phi(y)] = i\Delta(x-y)\,, \tag{4.283}$$

with

$$i\Delta(z) \stackrel{\text{def}}{\equiv} \Delta_+(z) - \Delta_+(-z)\,. \tag{4.284}$$

Since $\Delta_+(-x) = \Delta_+^*(x)$ (4.271), $\Delta_+(z) - \Delta_+(-z)$ is purely imaginary, and the factor i is added in the definition to make $\Delta(x)$ real. Our task now is to show that $\Delta(z) = 0$ for $z^2 < 0$ which establishes the causality (4.279). To do so, we will show that $\Delta(z)$ is Lorentz-invariant under proper and orthochronous transformations; or more precisely, it has the same value when one replaces z by Λz where Λ is proper and orthochronous. It should then be a function of the only Lorentz-invariant quantity formed by z: $z^2 \equiv z^0 - \vec{z}^2$. Then, for a given value of $z^2 < 0$, we take $z' = (0, \vec{z}')$ with $\vec{z}'^2 = -z^2 > 0$ for which $\Delta(z)$ should have the same value; namely

$$\Delta(z) = \Delta(z') \qquad \text{with } z' = (0, \vec{z}') \text{ and } \vec{z}'^2 = -z^2\,. \tag{4.285}$$

With $z' = x - y$, the commutator (4.283) with $z' = (0, \vec{z}')$ becomes the equal-time commutator $[\phi(t, \vec{x}), \phi(t, \vec{y})]$ which we know is zero. Thus, the critical factors are that $\Delta(z)$ is Lorentz-invariant under proper and orthochronous transformations and that any space-like 4-vector z can be transformed by a proper and orthochronous transformation such that its time component becomes zero. Now, all we need to demonstrate is the Lorentz invariance of Δ (or equivalently, of Δ_+).

The Lorentz invariance of $\Delta_+(x)$ can be shown using the following useful identity:

$$\boxed{\int d^4p\, \delta(p^2 - m^2)\theta(p^0) \cdots = \int \frac{d^3p}{2p^0} \cdots}, \tag{4.286}$$

where '$\cdots$' is any function of p^μ, by definition $p^2 \equiv {p^0}^2 - \vec{p}^2$, and $\theta(p^0)$ is a step function that limits the integration to positive p^0:

$$\theta(p^0) = \begin{cases} 1 & (p^0 \geq 0) \\ 0 & (p^0 < 0) \end{cases}\,. \tag{4.287}$$

In (4.286), p^0 is an integration variable on the left-hand side, while on the right-hand side, p^0 is fixed to $\sqrt{\vec{p}^2 + m^2}$.

Proof of (4.286)
We will apply the property of the delta function (4.109) to $f(x) = x^2 - (\vec{p}^2 + m^2)$. The roots of $f(x) = 0$ are

$$x_1 = +\sqrt{\vec{p}^2 + m^2}\,, \quad x_2 = -\sqrt{\vec{p}^2 + m^2}\,. \tag{4.288}$$

Using $f'(x) = 2x$,

$$\begin{aligned}\delta(f(x)) &= \sum_i \frac{1}{|f'(x_i)|}\delta(x - x_i) \\ &= \frac{\delta(x - x_1)}{2|x_1|} + \frac{\delta(x - x_2)}{2|x_2|}\,. \end{aligned} \tag{4.289}$$

Then, for an arbitrary function $F(x)$, we have

$$\begin{aligned}&\int_0^\infty dx\,\delta\Big(x^2 - (\vec{p}^{\,2} + m^2)\Big)F(x) \\ &\quad= \int_0^\infty dx\Bigg(\frac{\delta(x - \sqrt{\vec{p}^{\,2} + m^2})}{2\sqrt{\vec{p}^{\,2} + m^2}} + \underbrace{\frac{\delta(x + \sqrt{\vec{p}^{\,2} + m^2})}{2\sqrt{\vec{p}^{\,2} + m^2}}}_{\text{does not contribute}}\Bigg)F(x) \\ &\quad= \frac{F(\sqrt{\vec{p}^{\,2} + m^2})}{2\sqrt{\vec{p}^{\,2} + m^2}}\,. \end{aligned} \tag{4.290}$$

Renaming x as p^0 and using the definition $p^2 \equiv {p^0}^2 - \vec{p}^{\,2}$, this can be written as

$$\int_{-\infty}^\infty dp^0 \delta(p^2 - m^2)\theta(p^0)F(p^0) = \left.\frac{F(p^0)}{2p^0}\right|_{p^0=\sqrt{\vec{p}^{\,2}+m^2}}\,, \tag{4.291}$$

where $\theta(p^0)$ is inserted and the integration range is extended to $(-\infty, \infty)$. Note that p^0 on the left-hand side is an integration variable and not fixed to $\sqrt{\vec{p}^{\,2} + m^2}$. Assuming that F is a function also of $\vec{p}$ and integrating over $\vec{p}$, we obtain (4.286). ∎

Using (4.286), $\Delta_+(x)$ is now written as

$$\Delta_+(x) = \frac{1}{(2\pi)^3}\int d^4p\,\delta(p^2 - m^2)\theta(p^0)e^{-ip\cdot x} \tag{4.292}$$

where the integration $d^4p \equiv dp^0 d^3p$ is now over the entire four-dimensional space; in particular, the p^0 range is from $-\infty$ to ∞ thanks to the step function $\theta(p^0)$. Now, is this Lorentz-invariant? Namely, does $\Delta_+(x') = \Delta_+(x)$ hold if $x' = \Lambda x$ where Λ is a proper and orthochronous transformation? Let's see; we take $x' = \Lambda x$ as the argument and relabel $p^\mu \to p^{\mu\prime}$ which is allowed since they are dummy integration variables:

$$\Delta_+(x') = \frac{1}{(2\pi)^3}\int d^4p'\,\delta(p'^2 - m^2)\theta(p'^0)e^{-ip'\cdot x'}\,. \tag{4.293}$$

Then, we change the integration variables to p which is related to p' by

$$p = \Lambda^{-1} p' \,. \tag{4.294}$$

Since $x' = \Lambda x$ and $p' = \Lambda p$, we have

$$p'^2 = p^2 \,, \quad p' \cdot x' = p \cdot x \,, \tag{4.295}$$

and since Λ is proper and orthochronous,

$$d^4 p' = \underbrace{\det \Lambda}_{1} d^4 p = d^4 p \,, \quad \theta(p'^0) = \theta(p^0) \,. \tag{4.296}$$

The last relation $\theta(p'^0) = \theta(p^0)$ may not be trivial. This relation holds because as long as $p^2 = m^2 > 0$, as guaranteed by $\delta(p^2 - m^2)$, an orthochronous transformation cannot change the sign of the energy p^0. Putting all together, $\Delta_+(x')$ is now

$$\Delta_+(x') = \frac{1}{(2\pi)^3} \int d^4 p \, \delta(p^2 - m^2) \theta(p^0) e^{-ip \cdot x} = \Delta_+(x) \,. \tag{4.297}$$

Thus, $i\Delta(x) = \Delta_+(x) - \Delta_+(-x)$ is Lorentz-invariant; namely, it is a function only of x^2, and this completes the proof of the causality. At this point, it is worthwhile to point out the difference between the Lorentz-invariance of a scalar field $\phi(x)$ and that of $\Delta_+(x)$. Both are functions of x, but the scalar field changes its functional form under a Lorentz transformation: $\phi'(x') = \phi(x)$, while $\Delta(x)$ does not:

$$\Delta(x') = \Delta(x) \,. \tag{4.298}$$

This distinction is important; otherwise, any scalar field $\phi(x)$ would be a function only of x^2 !

We have seen that measurements of the Klein-Gordon field at two space-time points are indeed independent if they are separated by a space-like distance. It is interesting to see how this has happened. Since $i\Delta(x-y) = \Delta_+(x-y) - \Delta_+(y-x) = 0$ for space-like $x - y$, we see that $\Delta_+(x - y) = \Delta_+(y - x)$ for a space-like $x - y$. On the other hand, $\Delta_+(x - y)$ had a loose interpretation as the amplitude for a particle created at y to propagate to x, and according to the formula (4.271) (or Figure 4.2) it is non-zero even when $x - y$ is space-like. Then, the causality is accomplished by the amplitude of propagation from y to x cancelling out that of propagation from x to y.

Exercise 4.7 *Microscopic causality of charged Klein-Gordon field.*
(*a*) *Evaluate the commutator*

$$[\phi(x) \,, \phi^\dagger(y)]$$

where x and y are arbitrary space-time points, and express it in terms of the function Δ_+ (or Δ).
(b) In the case of hermitian field, there was an interpretation for the vanishing of the commutator for x and y separated by a space-like distance: cancellation of the amplitude for a particle to propagate from x to y and that for the propagation from y to x. Find a similar interpretation for this case.

Lorentz invariance of the quantization procedure
The procedure of canonical quantization is performed in a given frame using equal-time commutation relations. Then, a natural question is whether quantizations performed in different frames result in consistent quantized systems. In order to study this, it is convenient to re-write the sum over momentum as an integral. The reason is that the cube of volume $V = L^3$ is obviously not Lorentz-invariant and thus makes it difficult to connect formulations in different frames.

We start from the momentum expansion of $\phi(x)$ (4.167) and use the identity $\sum_{\vec{p}} = V/(2\pi)^3 \int d^3p$ (4.153):

$$\phi(x) = \sum_{\vec{p}} \Big(a_{\vec{p}}\, e_{\vec{p}}(x) + a^\dagger_{\vec{p}}\, e^*_{\vec{p}}(x) \Big) = \frac{V}{(2\pi)^3} \int d^3p \Big(a_{\vec{p}}\, e_{\vec{p}}(x) + a^\dagger_{\vec{p}}\, e^*_{\vec{p}}(x) \Big) \,, \tag{4.299}$$

and define new annihilation operators and normal-mode functions by

$$\tilde{a}_{\vec{p}} \stackrel{\text{def}}{\equiv} \sqrt{\frac{V}{(2\pi)^3}}\, a_{\vec{p}} \,, \qquad \tilde{e}_{\vec{p}}(x) \stackrel{\text{def}}{\equiv} \sqrt{\frac{V}{(2\pi)^3}}\, e_{\vec{p}}(x) = \frac{e^{-ip\cdot x}}{\sqrt{(2\pi)^3\, 2p^0}} \tag{4.300}$$

to write it as

$$\phi(x) = \int d^3p \Big(\tilde{a}_{\vec{p}}\, \tilde{e}_{\vec{p}}(x) + \tilde{a}^\dagger_{\vec{p}}\, \tilde{e}^*_{\vec{p}}(x) \Big) \,. \tag{4.301}$$

Note that in changing the sum to an integral in (4.299), the operator value of ϕ stays the same. It can be readily verified that the new normal-mode functions satisfy the orthonormal relations given by

$$\begin{aligned} &\int d^3x\, \tilde{e}^*_{\vec{p}}(x) i \overleftrightarrow{\partial}_0 \tilde{e}_{\vec{p}\,\prime}(x) = \delta^3(\vec{p} - \vec{p}\,') \,, \quad \int d^3x\, \tilde{e}_{\vec{p}}(x) i \overleftrightarrow{\partial}_0 \tilde{e}^*_{\vec{p}\,\prime}(x) = -\delta^3(\vec{p} - \vec{p}\,') \,, \\ &\int d^3x\, \tilde{e}_{\vec{p}}(x) i \overleftrightarrow{\partial}_0 \tilde{e}_{\vec{p}\,\prime}(x) = 0 \,, \qquad\qquad \int d^3x\, \tilde{e}^*_{\vec{p}}(x) i \overleftrightarrow{\partial}_0 \tilde{e}^*_{\vec{p}\,\prime}(x) = 0 \,. \end{aligned} \tag{4.302}$$

What is the commutator $[\tilde{a}_{\vec{p}}\,, \tilde{a}^\dagger_{\vec{p}\,\prime}]$? Since the discrete-$\vec{p}$ version is $\delta_{\vec{p},\vec{p}\,\prime}$, one expects that this is proportional to $\delta^3(\vec{p} - \vec{p}\,')$. In fact, using the definition (4.300), we find

$$\int d^3p\, [\tilde{a}_{\vec{p}}\,, \tilde{a}^\dagger_{\vec{p}\,\prime}] = \underbrace{\int d^3p\, \frac{V}{(2\pi)^3}}_{\sum_{\vec{p}}} \underbrace{[a_{\vec{p}}\,, a^\dagger_{\vec{p}\,\prime}]}_{\delta_{\vec{p},\vec{p}\,\prime}} = 1 \,; \tag{4.303}$$

namely, $[\tilde{a}_{\vec{p}}\,, \tilde{a}^\dagger_{\vec{p}'}]$ acts as $\delta^3(\vec{p}-\vec{p}')$:

$$[\tilde{a}_{\vec{p}}\,, \tilde{a}^\dagger_{\vec{p}'}] = \delta^3(\vec{p}-\vec{p}')\,. \tag{4.304}$$

Since $a_{\vec{p}}$ and $a_{\vec{p}'}$ commute regardless of $\vec{p}$ and $\vec{p}'$ in the discrete-$\vec{p}$ version, they will commute in the continuous-$\vec{p}$ version also:

$$[\tilde{a}_{\vec{p}}\,, \tilde{a}_{\vec{p}'}] = [\tilde{a}^\dagger_{\vec{p}}, \tilde{a}^\dagger_{\vec{p}'}] = 0\,. \tag{4.305}$$

We have seen before that $a^\dagger_{\vec{p}}$ creates one particle in the volume V. To see what $\tilde{a}^\dagger_{\vec{p}}$ does, we write the total energy (4.193) in the integral form (without changing its operator value):

$$H = \underbrace{\sum_{\vec{p}}}_{V/(2\pi)^3 \int d^3p}\, p^0\, \underbrace{a^\dagger_{\vec{p}} a_{\vec{p}}}_{(2\pi)^3/V\ \tilde{a}^\dagger_{\vec{p}}\tilde{a}_{\vec{p}}} = \int d^3p\, p^0 \tilde{a}^\dagger_{\vec{p}}\tilde{a}_{\vec{p}}\,, \tag{4.306}$$

and apply it to the state $\tilde{a}^\dagger_{\vec{p}}|0\rangle$:

$$H\tilde{a}^\dagger_{\vec{p}}|0\rangle = \Big(\int d^3p' p^{0\prime} \tilde{a}^\dagger_{\vec{p}'} \underbrace{\tilde{a}_{\vec{p}'}\Big)\tilde{a}^\dagger_{\vec{p}}}_{\delta^3(\vec{p}-\vec{p}') + \cancel{\tilde{a}^\dagger_{\vec{p}'}\tilde{a}_{\vec{p}'}}}|0\rangle = p^0\, \tilde{a}^\dagger_{\vec{p}}|0\rangle\,. \tag{4.307}$$

Thus, the total energy of the state $\tilde{a}^\dagger_{\vec{p}}|0\rangle$ is p^0, indicating that $\tilde{a}^\dagger_{\vec{p}}$ creates one particle with momentum $\vec{p}$ in the entire universe. This is not surprising since $\tilde{a}^\dagger_{\vec{p}}$ is proportional to $a^\dagger_{\vec{p}}$, and thus $\tilde{a}^\dagger_{\vec{p}}|0\rangle$ has the same eigenvalue as $a^\dagger_{\vec{p}}|0\rangle$; simply, the volume V is now the entire universe. It is usually clear whether we are using discrete $\vec{p}$ or continuous $\vec{p}$; thus, we will hereafter omit the '˜' on $a_{\vec{p}}$ and $e_{\vec{p}}(x)$.

Now, let's get back to the question of Lorentz-invariance of the quantization procedure, and define exactly what we mean by the statement. In a frame K, we take the Lagrangian of the Klein-Gordon field ϕ, find the conjugate field π, and introduce the commutation relations among them. When the field is momentum-expanded, the quantization condition becomes equivalent to the commutation relations among the expansion coefficients, namely among $a_{\vec{p}}$'s and $a^\dagger_{\vec{p}}$'s:

$$\begin{array}{c}[a_{\vec{p}}\,, a^\dagger_{\vec{p}'}] = \delta^3(\vec{p}-\vec{p}') \\ \text{all else} = 0\end{array} \quad \leftrightarrow \quad \begin{array}{c}[\phi(t,\vec{x}), \pi(t,\vec{x}')] = i\delta^3(\vec{x}-\vec{x}') \\ \text{all else} = 0\end{array} \tag{4.308}$$

Exercise 4.8 *Continous $\vec{p}$ formalisms.*
Start from the momentum expansion of the charged scalar field

$$\phi(x) = \int d^3p\, (a_{\vec{p}} e_{\vec{p}}(x) + b^\dagger_{\vec{p}} e^*_{\vec{p}}(x)) \tag{4.309}$$

with

$$e_{\vec{p}}(x) \equiv \frac{1}{\sqrt{(2\pi)^3 2p^0}} e^{-ip\cdot x} \tag{4.310}$$

and the commutation relations among creation and annihilation operators

$$\begin{gathered} [a_{\vec{p}}\,, a^\dagger_{\vec{p}'}] = [b_{\vec{p}}\,, b^\dagger_{\vec{p}'}] = \delta^3(\vec{p}-\vec{p}') \\ \textit{all else} = 0 \end{gathered} \tag{4.311}$$

to derive

$$\begin{aligned} [\phi(t,\vec{x}), \pi(t,\vec{x}')] &= i\delta^3(\vec{x}-\vec{x}')\,, \\ [\phi(t,\vec{x}), \phi(t,\vec{x}')] &= [\pi(t,\vec{x}), \pi(t,\vec{x}')] = 0\,, \\ [\phi(t,\vec{x}), \phi^\dagger(t,\vec{x}')] &= [\pi(t,\vec{x}), \pi^\dagger(t,\vec{x}')] = 0\,, \end{aligned} \tag{4.312}$$

where $\pi = \dot{\phi}^\dagger$ *and* $\pi^\dagger = \dot{\phi}$.

Now, if one performs the same quantization procedure in a different frame K' related to K by a proper and orthochronous transformation Λ, then the question is whether or not the commutation relations in the two frames are consistent. Namely, does the set of quantization conditions in one frame leads to that in another frame and vice versa? We could prove the consistency for the commutation relations for ϕ and π or, equivalently, those for $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$'s. We choose the latter.

First, we need some principle to relate the operators in the two systems, which is provided by the scalar-field condition

$$\phi'(x') = \phi(x) \qquad (x' = \Lambda x)\,. \tag{4.313}$$

This can be converted to relations for a's and $a^\dagger$'s as follows: Momentum-expanding $\phi'(x')$,

$$\begin{aligned} \phi'(x') &= \int d^3p' \Big(a'_{\vec{p}'} e_{\vec{p}'}(x') + a'^\dagger_{\vec{p}'} e^*_{\vec{p}'}(x') \Big) \\ &= \int d^3p' \frac{1}{\sqrt{(2\pi)^3 2p^{0\prime}}} \Big(a'_{\vec{p}'} e^{-ip'\cdot x'} + a'^\dagger_{\vec{p}'} e^{ip'\cdot x'} \Big) \end{aligned} \tag{4.314}$$

where we labeled the integration variable as $\vec{p}'$ (could be anything since it is a dummy variable), and the prime on $a_{\vec{p}}$ indicates that it is defined in the K' frame. We will now change the integration variable to p (or more precisely, its space part $\vec{p}$) related to p' by $p' = \Lambda p$. Using the identity (4.286) and (4.296), we have

$$\begin{aligned} \int \frac{d^3p'}{2p^{0\prime}} \cdots &= \int d^4p' \delta(p'^2 - m^2)\theta(p^{0\prime}) \cdots \\ &= \int d^4p\, \delta(p^2-m^2)\theta(p^0) \cdots = \int \frac{d^3p}{2p^0} \cdots\,. \end{aligned} \tag{4.315}$$

Thus, replacing d^3p' by $d^3p\,(p^{0\prime}/p^0)$, and using $p'\cdot x' = p\cdot x$, $\phi'(x')$ is now written as

$$\begin{aligned}\phi'(x') &= \int d^3p \frac{1}{\sqrt{(2\pi)^3 2p^0}}\sqrt{\frac{p^{0\prime}}{p^0}}\Big(a'_{\vec{p}\prime}e^{-ip\cdot x} + a'^{\dagger}_{\vec{p}\prime}e^{ip\cdot x}\Big)\\ &= \int d^3p \sqrt{\frac{p^{0\prime}}{p^0}}\Big(a'_{\vec{p}\prime}e_{\vec{p}}(x) + a'^{\dagger}_{\vec{p}\prime}e^*_{\vec{p}}(x)\Big)\,,\end{aligned} \tag{4.316}$$

which should equal to $\phi(x)$; namely,

$$\int d^3p \sqrt{\frac{p^{0\prime}}{p^0}}\Big(a'_{\vec{p}\prime}e_{\vec{p}}(x) + a'^{\dagger}_{\vec{p}\prime}e^*_{\vec{p}}(x)\Big) = \phi(x) = \int d^3p\,\Big(a_{\vec{p}}\,e_{\vec{p}}(x) + a^{\dagger}_{\vec{p}}\,e^*_{\vec{p}}(x)\Big)\,. \tag{4.317}$$

Applying $e^*_{\vec{p}}(x) i\overset{\leftrightarrow}{\partial}$ from the left and integrating over $\vec{x}$, the orthonormality relation (4.302) gives

$$\sqrt{\frac{p^{0\prime}}{p^0}}a'_{\vec{p}\prime} = a_{\vec{p}} \quad\rightarrow\quad \sqrt{p^{0\prime}}a'_{\vec{p}\prime} = \sqrt{p^0}a_{\vec{p}}\,, \tag{4.318}$$

which means that a particle with 4-momentum $p' = \Lambda p$ in the frame K' and a particle with 4-momentum p in the frame K are represented by the same state in the Hilbert space (the over-all constants do not affect the physical meaning of the states).

We now assume that $a_{\vec{p}}$ and $a^{\dagger}_{\vec{q}}$ satisfy $[a_{\vec{p}}, a^{\dagger}_{\vec{q}}] = \delta^3(\vec{p}-\vec{q})$ in the frame K, and derive the corresponding relations in the frame K', $[a'_{\vec{p}\prime}, a'^{\dagger}_{\vec{q}\prime}] = \delta^3(\vec{p}\prime - \vec{q}\prime)$. Using the relation (4.318), the commutation relation between $a'_{\vec{p}\prime}$ and $a'^{\dagger}_{\vec{q}\prime}$ is then

$$[a'_{\vec{p}\prime}, a'^{\dagger}_{\vec{q}\prime}] = \sqrt{\frac{p^0q^0}{p^{0\prime}q^{0\prime}}}[a_{\vec{p}}\,, a^{\dagger}_{\vec{q}}] = \frac{p^0}{p^{0\prime}}\delta^3(\vec{p}-\vec{q})\,, \tag{4.319}$$

where $\vec{p}$ and $\vec{q}$ are given by $p' = \Lambda p$ and $q' = \Lambda q$ (or the space parts thereof), and when the delta function forces $\vec{p}$ and $\vec{q}$ to be equal it also forces $p_0 = q_0$ since $p_0 \equiv \sqrt{\vec{p}^{\,2} + m^2}$ and $q_0 \equiv \sqrt{\vec{q}^2 + m^2}$ by defintiion. Thus, the delta function forces $p = q$ which also leads to $p' = q'$. On the other hand, one can show that

$$p^{0\prime}\delta^3(\vec{p}\prime - \vec{q}\prime) = p^0\delta^3(\vec{p}-\vec{q}) \tag{4.320}$$

as follows: Using (4.315) and noting that $\vec{p}\prime = \vec{q}\prime$ corresponds to $\vec{p} = \vec{q}$, we have

$$\begin{aligned}\int \underbrace{\frac{d^3p}{2p^0}}_{d^3p'/(2p^{0\prime})}\Big[2p^{0\prime}\delta^3(\vec{p}\prime - \vec{q}\prime)\Big]f(\vec{p}) &= \int d^3p'\delta^3(\vec{p}\prime - \vec{q}\prime)f(\vec{p})\\ &= f(\vec{p})\Big|_{\vec{p}\prime=\vec{q}\prime} = f(\vec{q})\,;\end{aligned} \tag{4.321}$$

Namely, $2p^{0\prime}\delta^3(\vec{p}'-\vec{q}')$ behaves the same way as $2p^0\delta^3(\vec{p}-\vec{q})$:

$$\int \frac{d^3p}{2p^0}\Big[2p^0\delta^3(\vec{p}-\vec{q})\Big]f(\vec{p}) = \int d^3p\,\delta^3(\vec{p}-\vec{q})f(\vec{p}) = f(\vec{q})\,, \tag{4.322}$$

thus proving (4.320). Then, from (4.319) and (4.320), we have

$$[a'_{\vec{p}'}, a'^{\dagger}_{\vec{q}'}] = \delta^3(\vec{p}'-\vec{q}')\,, \tag{4.323}$$

which shows that the same quantization condition holds in the K' frame. Thus, we have shown that the quantization procedure is Lorentz-invariant.

4.6 Quantization of the Dirac field

Quantization of the Dirac field proceeds similarly to that of the Klein-Gordon field, except for one major difference: once the classical field is expanded into normal modes, the quantization condition should be imposed such that it is consistent with Pauli's exclusion principle; namely, each state can be occupied by at most one quantum. We have seen that such condition can be implemented if we use anticommutators instead of commutators among the creation and annihilation operators. Thus, we will first find the Lagrangian that gives the Dirac field, derive the conjugate field and the Hamiltonian, and upon momentum-expanding the field, we will introduce anticommutation relations among the expansion coefficients. Then, anticommutators among the field and its conjugate field will follow, time variations of operators will be given by Heisenberg's equations of motion as before, and the quantized system will be established. Our initial motivation for using the anticommutators is thus experimental - namely, nature demands it in the form of Pauli's exclusion principle. We will find, however, that it is also demanded by the theory if we require that both particle and antiparticle have positive energy, and that the Dirac field respects microscopic causality.

Lagrangian formulation of the Dirac field

As we will see shortly, the Lagrangian density for the Dirac equation is given by

$$\boxed{\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi\,.} \tag{4.324}$$

To show that this Lagrangian indeed gives the Dirac equation, we go back to the action principle itself:

$$\delta S = 0\,, \qquad S \equiv \int d^4x\mathcal{L} = \int d^4x\,\psi^\dagger\gamma^0(i\partial\!\!\!/ - m)\psi\,, \tag{4.325}$$

where the space integral is over the volume V (or the entire universe) and the time integral is from t_1 to t_2 which is the time window of interest. Here, we have 4-component field ψ with each component being complex. As before, ψ_n and ψ_n^* ($n =$

$1, 2, 3, 4$) can be regarded as independent variables. Thus, we first take all possible small variations of $\psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$ keeping ψ unchanged. We will vary ψ later. In fact, it does not matter which of the eight independent variables are changed; one could vary only one at a time if one wishes; the result will be the same. Then, the variation in S in this case is very simple:

$$\delta S = \int d^4x\, \delta\psi^\dagger\, \gamma^0 (i\partial\!\!\!/ - m)\psi = 0\,, \tag{4.326}$$

with

$$\delta\psi^\dagger \equiv (\delta\psi_1^*, \delta\psi_2^*, \delta\psi_3^*, \delta\psi_4^*)\,. \tag{4.327}$$

This should hold for all possible variations of ψ; thus, each of the four components of $\gamma^0(i\partial\!\!\!/ - m)\psi$ should be zero. Multiplying γ^0 from the left,

$$\gamma^0(i\partial\!\!\!/ - m)\psi = 0\,, \qquad \rightarrow \quad (i\partial\!\!\!/ - m)\psi = 0\,, \tag{4.328}$$

which is the Dirac equation as promised.

You may find, however, something odd about the form of this Lagrangian: it is patently asymmetric between ψ and $\psi^\dagger$, and it is not real either. In fact, if we take complex conjugate of $\mathcal{L}$, we get

$$\begin{aligned} \mathcal{L}^* &= \left[\bar\psi(i\partial\!\!\!/ - m)\psi\right]^* = \left[\bar\psi\gamma^\mu i\partial_\mu\psi - m\bar\psi\psi\right]^* \\ &= (-i\partial_\mu\bar\psi)\underbrace{\overline{\gamma^\mu}}_{\gamma^\mu}\psi - m\bar\psi\psi \\ &= -\bar\psi(i\overleftarrow{\partial\!\!\!/} + m)\psi\,. \end{aligned} \tag{4.329}$$

which seems to differ from the original $\mathcal{L}$. It does, however, have the form from which we can extract the Dirac equation for $\psi^\dagger$ (or equivalently $\bar\psi$) by varying the four components of ψ:

$$\begin{aligned} \delta\int d^4x\,\mathcal{L}^* &= -\int d^4x\,\bar\psi(i\overleftarrow{\partial\!\!\!/} + m)\delta\psi = 0 \\ &\rightarrow\ \bar\psi(i\overleftarrow{\partial\!\!\!/} + m) = 0\,. \end{aligned} \tag{4.330}$$

which is the Dirac equation for $\bar\psi$ (3.70). In fact, $\mathcal{L}$ and $\mathcal{L}^*$ are related by partial integration:

$$\begin{aligned} S &= \int d^4x\,\mathcal{L} = \int d^4x\Big[i\underbrace{\bar\psi\gamma^\mu\partial_\mu\psi}_{\qquad\qquad} - m\bar\psi\psi\Big] \\ &\qquad\qquad\qquad\qquad \underbrace{\partial_\mu(\bar\psi\gamma^\mu\psi)}_{\rightarrow 0} - (\partial_\mu\bar\psi)\gamma^\mu\psi \\ &= -\int d^4x\,\bar\psi(i\overleftarrow{\partial\!\!\!/} + m)\psi \\ &= \int d^4x\,\mathcal{L}^*\,. \end{aligned} \tag{4.331}$$

The discarding of the integral $\int d^4x\,\partial_\mu(\bar\psi\gamma^\mu\psi)$ requires some clarification. First, we are not using the conservation of current $\partial_\mu j^\mu = 0$ ($j^\mu \equiv \bar\psi\gamma^\mu\psi$) since ψ does not necessarily satisfy the Dirac equation. It does not contribute, however, when δS is evaluated:

$$\begin{aligned}\delta\int d^4x\,\partial_\mu j^\mu &= \delta\int dt\int d^3x(\partial_0 j^0 + \underbrace{\vec\nabla\cdot\vec j}_{\to\ 0 \text{ upon } \int d^3x})\\ &= \int_{t_1}^{t_2} dt\,\partial_0\Big(\int d^3x\,\delta\, j^0\Big) = \Big[\int d^3x\,\delta\, j^0\Big]_{t_1}^{t_2} = 0\,,\end{aligned}\tag{4.332}$$

where the last equality is due to the constraint $\delta\psi(t_1) = \delta\psi(t_2) = 0$ (4.69) which leads to $\delta j^0(t_1) = \delta j^0(t_2) = 0$. Thus, the Lagrangian densities $\mathcal{L}$ and $\mathcal{L}^*$ are equivalent and results in the same equation of motion. In fact, $\bar\psi(i\overleftarrow{\not\partial} + m) = 0$ is just the spinor adjoint of $(i\not\partial - m)\psi = 0$. One could force the Lagrangian to be real by

$$\begin{aligned}\mathcal{L}^R \equiv \frac{1}{2}(\mathcal{L}+\mathcal{L}^*) &= \frac{1}{2}\bar\psi(i\not\partial - m - i\overleftarrow{\not\partial} - m)\psi\\ &= \bar\psi\Big(i\frac{\overleftrightarrow{\not\partial}}{2} - m\Big)\psi\,.\end{aligned}\tag{4.333}$$

We will use $\mathcal{L} = \bar\psi(i\not\partial - m)\psi$ most of the time.

The field conjugate to the n-th component of ψ is, by definition,

$$\pi_n \equiv \frac{\partial\mathcal{L}}{\partial\dot\psi_n} = \frac{\partial}{\partial\dot\psi_n}\big[\underbrace{\bar\psi\gamma^\mu i\partial_\mu\psi}_{\bar\psi\gamma^0 i\partial_0\psi + \cdots} - m\bar\psi\psi\big] = \underbrace{i\psi_n^*}_{}\dot\psi_n$$

$$\to \boxed{\pi = i\psi^\dagger}\,.\tag{4.334}$$

How about the field conjugate to ψ_n^*? Since there is no $\dot\psi_n^*$ appearing in $\mathcal{L}$, it does not have a conjugate field. In fact, the above expression for π tells us that ψ_n^* itself is conjugate to ψ_n up to a constant. The situation is similar if we start from $\mathcal{L}^*$ instead of $\mathcal{L}$; in this case, ψ_n's do not have conjugate field. In either case, we have four fields and their conjugate fields. If this is confusing, one can always take the real and imaginary parts of each field to be the independent fields, in which case we have eight fields to start with. Then, regardless of which Lagrangian one uses [$\mathcal{L}$, $\mathcal{L}^*$, or $\mathcal{L}^R = (\mathcal{L}+\mathcal{L}^*)/2$], one finds that four fields become conjugate to the other four fields or linear combinations thereof, ending up with four dynamical degrees of freedom in each case. In the case of the complex Klein-Gordon field, we saw that the field conjugate to ϕ was the time derivative of ϕ up to a constant. Why don't we

get time derivatives in this case when we extract conjugate fields? Apparently, that is because the Lagrangian is linear in the time derivative of fields which in turn is because the Dirac equation is linear in time derivative.

Using the Lagrangian $\mathcal{L}$ and the conjugate fields π given above, and regarding $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$ as independent fields, the Hamiltonian is

$$\begin{aligned}\mathcal{H} &\equiv \sum_{n=1}^{4} \underbrace{\pi_n}_{i\psi_n^*} \dot{\psi}_n - \underbrace{\mathcal{L}}_{\bar{\psi}\gamma^\mu i\partial_\mu\psi - m\bar{\psi}\psi} \\ &= \cancel{i\psi^\dagger\dot{\psi}} - \Big(\underbrace{i\bar{\psi}\gamma^0\partial_0\psi}_{\cancel{i\psi^\dagger\dot{\psi}}} + \underbrace{i\bar{\psi}\gamma^k\partial_k\psi}_{i\psi^\dagger \underbrace{\gamma^0\gamma^k}_{\alpha_k}\partial_k\psi} - m\psi^\dagger \underbrace{\gamma^0}_{\beta}\psi \Big) ,\end{aligned} \tag{4.335}$$

where k is the space index which is summed, or

$$\boxed{\mathcal{H} = \psi^\dagger(-i\vec{\alpha}\cdot\vec{\nabla} + m\beta)\psi} . \tag{4.336}$$

Note that the operator sandwiched by $\psi^\dagger$ and ψ is nothing but the 'Hamiltonian' we encountered when we introduced the Dirac equation: $i\partial_0\psi = (-i\vec{\alpha}\cdot\vec{\nabla} + m\beta)\psi$. When ψ satisfies the Dirac equation, the total Hamiltonian can be written as

$$H \equiv \int d^3x\, \mathcal{H} = \int d^3x\, \psi^\dagger i\partial_0\psi \,. \tag{4.337}$$

The total momentum is obtained by applying the general form (4.117):

$$\vec{P} \equiv -\int d^3x \underbrace{\pi_n}_{i\psi_n^*}\vec{\nabla}\psi_n = \int d^3x\, \psi^\dagger(-i\vec{\nabla})\psi \,. \tag{4.338}$$

As in the case of the Klein-Gordon field (4.201), we note the form for the total energy and momentum where the corresponding differential operator $\mathcal{O} = i\partial_0$ or $-i\vec{\nabla}$ is sandwiched by the appropriate inner product of field: $\frac{1}{2}\phi(i\overleftrightarrow{\partial_0})\mathcal{O}\phi$ for the real Klein-Gordon field, and $\psi^\dagger\mathcal{O}\psi$ for the Dirac field.

Momentum expansion of the Dirac field

We now proceed to quantize the Dirac field. The first step is to expand the general solution of the Dirac equation into normal modes. Suppose $\psi(t, \vec{x})$ is an arbitrary solution of the Dirac equation; then, at a given time, say $t = 0$, each component $\psi_n(0, \vec{x})$ can be uniquely Fourier-expanded:

$$\psi(0,\vec{x}) = \begin{pmatrix} \psi_1(0,\vec{x}) \\ \psi_2(0,\vec{x}) \\ \psi_3(0,\vec{x}) \\ \psi_4(0,\vec{x}) \end{pmatrix} = \begin{pmatrix} \sum_{\vec{p}} c_{1\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} \\ \sum_{\vec{p}} c_{2\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} \\ \sum_{\vec{p}} c_{3\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} \\ \sum_{\vec{p}} c_{4\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} \end{pmatrix} = \sum_{\vec{p}} \begin{pmatrix} c_{1\vec{p}} \\ c_{2\vec{p}} \\ c_{3\vec{p}} \\ c_{4\vec{p}} \end{pmatrix} e^{i\vec{p}\cdot\vec{x}} , \tag{4.339}$$

where $c_{n\vec{p}}$ $(n = 1, 2, 3, 4)$ are uniquely-determined complex coefficients. For each $\vec{p}$, we can use the orthonormal set of spinors $(u_{\vec{p},\pm\vec{s}}\,, v_{-\vec{p},\pm\vec{s}})$ (3.306) to uniquely expand $c_{n\vec{p}}$:

$$\begin{pmatrix} c_{1\vec{p}} \\ c_{2\vec{p}} \\ c_{3\vec{p}} \\ c_{4\vec{p}} \end{pmatrix} = \sum_{\vec{s}} (A_{\vec{p},\vec{s}}\, u_{\vec{p},\vec{s}} + B_{-\vec{p},\vec{s}}\, v_{-\vec{p},\vec{s}})\,, \tag{4.340}$$

where the sum over $\vec{s}$ is understood to be over $\pm\vec{s}$ where $\vec{s}$ is some fixed unit vector, which could be a function of $\vec{p}$ ($\vec{s} = \hat{p}$, for example) or fixed in space ($\vec{s} = \hat{z}$, for example). It is important to note that for any pair $(\vec{p}, \vec{s})$, we have an orthonormal set $(u_{\vec{p},\pm\vec{s}}\,, v_{-\vec{p},\pm\vec{s}})$ by which any complex 4-component vector can be uniquely expanded. In particular, the set of four spinors $(u_{\vec{p},\pm\vec{s}}\,, v_{-\vec{p},\pm\vec{s}})$ are different for different $\vec{p}$. We choose to expand the 4-component spinor $(c_{1\vec{p}}, c_{2\vec{p}}, c_{3\vec{p}}, c_{4\vec{p}})$ using the orthonormal set $(u_{\vec{p},\pm\vec{s}}\,, v_{-\vec{p},\pm\vec{s}})$ where $\vec{p}$ is the same one that appears in the indexes of $c_{n\vec{p}}$'s. The spin quantization axis $\vec{s}$ is arbitrary at this point; it could be in z-direction for all $\vec{p}$ or it could depend on $\vec{p}$. We now have a unique expansion of a general solution of the Dirac equation at $t = 0$:

$$\begin{aligned} \psi(0, \vec{x}) &= \sum_{\vec{p},\vec{s}} (A_{\vec{p},\vec{s}}\, u_{\vec{p},\vec{s}} + B_{-\vec{p},\vec{s}}\, v_{-\vec{p},\vec{s}}) e^{i\vec{p}\cdot\vec{x}} \\ &= \sum_{\vec{p},\vec{s}} (A_{\vec{p},\vec{s}}\, u_{\vec{p},\vec{s}}\, e^{i\vec{p}\cdot\vec{x}} + \underbrace{B_{\vec{p},\vec{s}}\, v_{\vec{p},\vec{s}}\, e^{-i\vec{p}\cdot\vec{x}}}_{\text{relabeled } \vec{p} \leftrightarrow -\vec{p}}) \,. \end{aligned} \tag{4.341}$$

What is the time dependence of $\psi(t, \vec{x})$? Actually, we already know the answer: in order to be a solution of the Dirac equation, $u_{\vec{p},\vec{s}}\, e^{i\vec{p}\cdot\vec{x}}$ should be attached to the positive frequency $e^{-ip^0 t}$, and $v_{\vec{p},\vec{s}}\, e^{-i\vec{p}\cdot\vec{x}}$ should be attached to the negative frequency $e^{ip^0 t}$ with $p^0 \equiv \sqrt{\vec{p}^{\,2} + m^2}$:

$$u_{\vec{p},\vec{s}}\, e^{-ip\cdot x} = u_{\vec{p},\vec{s}}\, e^{-ip^0 t + i\vec{p}\cdot\vec{x}}\,, \qquad v_{\vec{p},\vec{s}}\, e^{ip\cdot x} = v_{\vec{p},\vec{s}}\, e^{ip^0 t - i\vec{p}\cdot\vec{x}}\,. \tag{4.342}$$

Thus, the general solution is now written as

$$\psi(t, \vec{x}) = \sum_{\vec{p},\vec{s}} (A_{\vec{p},\vec{s}}\, u_{\vec{p},\vec{s}}\, e^{-ip\cdot x} + B_{\vec{p},\vec{s}}\, v_{\vec{p},\vec{s}}\, e^{ip\cdot x}) \,. \tag{4.343}$$

This expansion is general and unique; namely, if you give me an arbitrary solution of the Dirac equation, I can uniquely expand it at a given time t in the Fourier series in the 3-dimensional space, and then the time dependence is uniquely determined for the function to be a solution of the Dirac equation. Note that, in addition to the appropriate time dependence, $u_{\vec{p},\vec{s}}$ had to be attached to $e^{i\vec{p}\cdot\vec{s}}$ and $v_{\vec{p},\vec{s}}$ to $e^{-i\vec{p}\cdot\vec{s}}$. The required opposite sign of $\vec{p}$ in the subscript of $v_{\vec{p},\vec{s}}$ and that in the corresponding exponent would not have resulted if we had used the other orthonormal set $(u_{\vec{p},\pm\vec{s}}\,, v_{\vec{p},\pm\vec{s}})$ which is defined by the inner product $\bar{a}b$ (try it).

In the case of the Klein-Gordon field, the normal mode $\phi(x) = e_{\vec{p}}(x)$ was normalized such that the classical probability density $\phi^* i \overset{\leftrightarrow}{\partial}_0 \phi$ gives unity when integrated over V as seen in (4.175). In the case of the Dirac field, the classical probability density is given by $j^0 = \psi^\dagger \psi$; thus, we take the normal-mode functions to be

$$\boxed{f_{\vec{p},\vec{s}}(x) \overset{\text{def}}{\equiv} \frac{u_{\vec{p},\vec{s}}}{\sqrt{2p^0 V}} e^{-ip\cdot x}\,, \quad g_{\vec{p},\vec{s}}(x) \overset{\text{def}}{\equiv} \frac{v_{\vec{p},\vec{s}}}{\sqrt{2p^0 V}} e^{ip\cdot x}\,,} \tag{4.344}$$

which is normalized properly:

$$\int d^3x f^\dagger_{\vec{p},\vec{s}}(x) f_{\vec{p},\vec{s}}(x) = \int_V d^3x \frac{1}{2p^0 V} \underbrace{u^\dagger_{\vec{p},\vec{s}} u_{\vec{p},\vec{s}}}_{2p^0 \text{ by (3.306)}} = 1\,, \text{ etc.} \tag{4.345}$$

Then, the expansion is now

$$\boxed{\psi(x) = \sum_{\vec{p},\vec{s}} \Big(a_{\vec{p},\vec{s}}\, f_{\vec{p},\vec{s}}(x) + b^\dagger_{\vec{p},\vec{s}}\, g_{\vec{p},\vec{s}}(x) \Big)}, \tag{4.346}$$

where

$$a_{\vec{p},\vec{s}} \equiv \sqrt{2p^0 V}\, A_{\vec{p},\vec{s}}\,, \qquad b^\dagger_{\vec{p},\vec{s}} \equiv \sqrt{2p^0 V}\, B_{\vec{p},\vec{s}}\,. \tag{4.347}$$

Note that we used $b^\dagger_{\vec{p},\vec{s}}$ instead of $b_{\vec{p},\vec{s}}$ in the second equation above. This anticipates that it will be a creation operator, rather than an annihilation operator, of antiparticle. We will discuss this choice later in the context of the energy sign of antiparticle. The spin sum is over $\pm\vec{s}$, where $\vec{s}$ is a unit vector which could in general be a function of $\vec{p}$. It is easily verified that the normal-mode functions satisfy the orthonormality relations given by

$$\boxed{\begin{aligned} \int d^3x f^\dagger_{\vec{p},\vec{s}}(x) f_{\vec{p}',\vec{s}'}(x) &= \int d^3x\, g^\dagger_{\vec{p},\vec{s}}(x) g_{\vec{p}',\vec{s}'}(x) = \delta_{\vec{p},\vec{p}'}\delta_{\vec{s},\vec{s}'} \\ \int d^3x f^\dagger_{\vec{p},\vec{s}}(x) g_{\vec{p}',\vec{s}'}(x) &= \int d^3x\, g^\dagger_{\vec{p},\vec{s}}(x) f_{\vec{p}',\vec{s}'}(x) = 0 \end{aligned}}, \tag{4.348}$$

where $\vec{s}$ and $\vec{s}'$ are plus or minus some direction which is in general a function of $\vec{p}$.

Exercise 4.9 *Verify (4.348).*

Then, we can use the orthonormality relations to write a's and $b^\dagger$'s as

$$\begin{aligned} a_{\vec{p},\vec{s}} &= \int d^3x\, f^\dagger_{\vec{p},\vec{s}}(x) \psi(x) \\ b^\dagger_{\vec{p},\vec{s}} &= \int d^3x\, g^\dagger_{\vec{p},\vec{s}}(x) \psi(x)\,. \end{aligned} \tag{4.349}$$

Up to this point, $\psi(x)$ is just a complex wave function. We will now regard a's and $b^\dagger$'s as operators in the Hilbert space and impose anticommutation relations given by

$$\boxed{\begin{array}{c}\{a_{\vec{p},\vec{s}}\,,a^\dagger_{\vec{p}',\vec{s}'}\}=\{b_{\vec{p},\vec{s}}\,,b^\dagger_{\vec{p}',\vec{s}'}\}=\delta_{\vec{p},\vec{p}'}\delta_{\vec{s},\vec{s}'}\\ \text{all others}=0\end{array}}\,. \tag{4.350}$$

This will assure that the oscillator associated with each normal mode will be occupied by at most one quantum; namely, the number operators $a^\dagger_{\vec{p},\vec{s}}a_{\vec{p},\vec{s}}$ and $b^\dagger_{\vec{p},\vec{s}}b_{\vec{p},\vec{s}}$ have eigenvalues 0 and 1 only. Note that the operator field $\psi(x)$ satisfies the Dirac equation since $f_{\vec{p},\vec{s}}(x)$ and $g_{\vec{p},\vec{s}}(x)$ in the expansion (4.346) do:

$$(i\partial\!\!\!/ - m)\psi(x)=0 \qquad (\psi:\text{operator})\,. \tag{4.351}$$

Using the momentum expansion, we can derive aniticommutation relations among fields. Let's define

$$\psi=\psi_a+\psi_{b^\dagger}\,,\quad \text{with}\quad \psi_a\stackrel{\text{def}}{\equiv}\sum_{\vec{p},\vec{s}}a_{\vec{p},\vec{s}}\,f_{\vec{p},\vec{s}}\,,\quad \psi_{b^\dagger}\stackrel{\text{def}}{\equiv}\sum_{\vec{p},\vec{s}}b^\dagger_{\vec{p},\vec{s}}\,g_{\vec{p},\vec{s}}\,. \tag{4.352}$$

Since only the combinations $\{a,a^\dagger\}$ or $\{b,b^\dagger\}$ survive,

$$\{\psi_n(t,\vec{x})\,,\psi_m(t,\vec{x}\,')\}=\{\psi^\dagger_n(t,\vec{x})\,,\psi^\dagger_m(t,\vec{x}\,')\}=0\,. \tag{4.353}$$

The anticommutator $\{\psi_n(t,\vec{x})\,,\psi^\dagger_m(t,\vec{x}\,')\}$ becomes

$$\begin{aligned}\{\psi_n(t,\vec{x})\,,\psi^\dagger_m(t,\vec{x}\,')\} &= \{\psi_{an}(t,\vec{x})+\psi_{b^\dagger n}(t,\vec{x})\,,\psi^\dagger_{am}(t,\vec{x}\,')+\psi^\dagger_{b^\dagger m}(t,\vec{x}\,')\}\\ &= \{\psi_{an}(t,\vec{x})\,,\psi^\dagger_{am}(t,\vec{x}\,')\}+\{\psi_{b^\dagger n}(t,\vec{x})\,,\psi^\dagger_{b^\dagger m}(t,\vec{x}\,')\}\end{aligned} \tag{4.354}$$

Using the anticommutation relation (4.350) and the explicit expression for $f_{\vec{p},\vec{s}}$ (4.344), we have

$$\begin{aligned}\{\psi_{an}(t,\vec{x})\,,\psi^\dagger_{am}(t,\vec{x}\,')\} &= \sum_{\vec{p},\vec{s},\vec{p}',\vec{s}'}\underbrace{\{a_{\vec{p},\vec{s}}\,,a^\dagger_{\vec{p}',\vec{s}'}\}}_{\delta_{\vec{p},\vec{p}'}\delta_{\vec{s},\vec{s}'}}f_{\vec{p},\vec{s}_n}(t,\vec{x})f^\dagger_{\vec{p}',\vec{s}'_m}(t,\vec{x}\,')\\ &= \sum_{\vec{p},\vec{s}}f_{\vec{p},\vec{s}_n}(t,\vec{x})f^\dagger_{\vec{p},\vec{s}_m}(t,\vec{x}\,')\\ &= \sum_{\vec{p}}\frac{e^{i\vec{p}\cdot(\vec{x}-\vec{x}\,')}}{2p^0V}\sum_{\vec{s}}(u_{\vec{p},\vec{s}}u^\dagger_{\vec{p},\vec{s}})_{nm}\,.\end{aligned} \tag{4.355}$$

Similarly,

$$\begin{aligned}\{\psi_{b^\dagger n}(t,\vec{x})\,,\psi^\dagger_{b^\dagger m}(t,\vec{x}\,')\} &= \sum_{\vec{p}}\frac{e^{-i\vec{p}\cdot(\vec{x}-\vec{x}\,')}}{2p^0V}\sum_{\vec{s}}(v_{\vec{p},\vec{s}}v^\dagger_{\vec{p},\vec{s}})_{nm}\\ (\text{relabel }\vec{p}\to-\vec{p}) &= \sum_{\vec{p}}\frac{e^{i\vec{p}\cdot(\vec{x}-\vec{x}\,')}}{2p^0V}\sum_{\vec{s}}(v_{-\vec{p},\vec{s}}v^\dagger_{-\vec{p},\vec{s}})_{nm}\,.\end{aligned} \tag{4.356}$$

Then, the anticommutator $\{\psi_n(t,\vec{x})\,,\psi_m^\dagger(t,\vec{x}')\}$ is, adding (4.355) and (4.356),

$$\begin{aligned}\{\psi_n(t,\vec{x})\,,\psi_m^\dagger(t,\vec{x}')\} &= \underbrace{\sum_{\vec{p}}\frac{1}{V}e^{i\vec{p}\cdot(\vec{x}-\vec{x}')}}_{\delta^3(\vec{x}-\vec{x}')\text{ by (4.151)}}\ \underbrace{\frac{1}{2p^0}\sum_{\vec{s}}(u_{\vec{p},\vec{s}}u^\dagger_{\vec{p},\vec{s}}+v_{-\vec{p},\vec{s}}v^\dagger_{-\vec{p},\vec{s}})_{nm}}_{\delta_{nm}\text{ by (3.319)}}\\ &= \delta_{nm}\delta^3(\vec{x}-\vec{x}')\,. \end{aligned} \tag{4.357}$$

Thus, using $\pi = i\psi^\dagger$ and together with (4.353),

$$\begin{aligned}\{\psi_n(t,\vec{x})\,,\pi_m(t,\vec{x}')\} &= i\delta_{nm}\delta^3(\vec{x}-\vec{x}')\\ \{\psi_n(t,\vec{x})\,,\psi_m(t,\vec{x}')\} &= \{\pi_n(t,\vec{x})\,,\pi_m(t,\vec{x}')\} = 0\end{aligned}\ . \tag{4.358}$$

Namely, the field ψ and its conjugate π satisfy similar quantization conditions as those of the Klein-Gordon fields, but with commutators replaced by anticommutators.

We have seen that one complex field represents a charged spin-0 particle and a set of four complex fields represents a spin-1/2 particle. For a charged spin-0 particle with a given moemtnum, there are two degrees of freedom: a particle and its antiparticle. On the other hand, for a spin-1/2 particle with a given momentum, there are four degrees of freedom: electron, positron and spin up, down. The question is then why a spinor field requires factor of two more complex fields per degree of freedom compared to a scalar field. One way to understand this is to count the number of independent *canonical pairs* of fields. For a charged spin-0 field, there are two such pairs (ϕ,π) and $(\phi^\dagger,\pi^\dagger)$. For a spin-1/2 field, the conjugate field is given by $\pi = i\psi^\dagger$, and thus there are only four independent canonical pairs. Another way is to simply count the number of independent complex coefficients required for a given $\vec{p}$ in the Fourier expansion of field. For a complex scalar field, there are two time dependences that can be assigned for a given $\vec{p}$; namely, $e^{-ip\cdot x}$ and $e^{ip\cdot x}$. Then, after quantization, the corresponding coefficients $a_{\vec{p}}$ and $b^\dagger_{\vec{p}}$ are interpreted as the annihilation operator of particle and the creation opearator of antiparticle, respectively. For a spinor field, when the spatial dependence is Fourier transformed and expanded into u and v spinors there are four coefficients for a given $\vec{p}$. However, the time dependences corresponding to $u_{\vec{p}\vec{s}}$ or $v_{\vec{p}\vec{s}}$ are already uniquely defined, and thus one has only four independent coefficients, leading to four degrees of freedom for a given $\vec{p}$. The difference between the spin-0 and spin-1/2 cases arises essentially from the fact that the Klein-Gordon equation is second-order in time derivative while the Dirac equation is first-order.

Total energy and momentum

We will now express the total energy $H = \int d^3x\,\psi^\dagger i\partial_0\psi$ (4.337) in terms of annihilation and creation operators. Using the expansion (4.346) and the orthonormality relations (4.348), we obtain

$$H = \int d^3x\psi^\dagger i\partial_0\psi$$

$$
\begin{aligned}
&= \int d^3x \sum_{\vec{p},\vec{s}} (a^\dagger_{\vec{p},\vec{s}} f^\dagger_{\vec{p},\vec{s}} + b_{\vec{p},\vec{s}} g^\dagger_{\vec{p},\vec{s}}) \sum_{\vec{p}',\vec{s}'} p^{0\prime} (a_{\vec{p}',\vec{s}'} f_{\vec{p}',\vec{s}'} - b^\dagger_{\vec{p}',\vec{s}'} g_{\vec{p}',\vec{s}'}) \\
&= \sum_{\vec{p},\vec{s},\vec{p}',\vec{s}'} p^{0\prime} (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p}',\vec{s}'} \delta_{\vec{p},\vec{p}'} \delta_{\vec{s},\vec{s}'} - b_{\vec{p},\vec{s}} b^\dagger_{\vec{p}',\vec{s}'} \delta_{\vec{p},\vec{p}'} \delta_{\vec{s},\vec{s}'}) \\
&= \sum_{\vec{p},\vec{s}} p^0 (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} - b_{\vec{p},\vec{s}} b^\dagger_{\vec{p},\vec{s}}) \qquad (4.359)
\end{aligned}
$$

So far, we have used only the orthonormality relations of normal-mode functions, and no operator relations have been used. In fact, the above is valid for the non-quantized fields also if hermitian conjugation is understood to be complex conjugation. Now, we can change the second term to a number operator by $\{b_{\vec{p},\vec{s}}\,, b^\dagger_{\vec{p},\vec{s}}\} = 1$ or $b_{\vec{p},\vec{s}} b^\dagger_{\vec{p},\vec{s}} = 1 - b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}}$:

$$
H = \sum_{\vec{p},\vec{s}} p^0 (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} + b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}} - 1)\,, \qquad (4.360)
$$

which shows that both a-particles and b-particles contribute positively to the total energy. Since, in the expansion of ψ, a-particles are associated with electron solutions $u_{\vec{p},\vec{s}}\, e^{-ip\cdot x}$ and b-particle with positron solutions $v_{\vec{p},\vec{s}}\, e^{ip\cdot x}$, one expects that a-particles are electrons and b-particles positrons (or fermions and anti-fermions). Such an interpretation will be justified later when we find 'charge' carried by those particles. Thus, we see that both electrons and positrons carry positive energy. Suppose we had quantized by commutators instead of anticommutators; namely, suppose the braces in (4.350) were square brackets. Then, (4.359) would still be valid, and using the commutator $[b_{\vec{p},\vec{s}}\,, b^\dagger_{\vec{p},\vec{s}}] = 1$, the total energy would then be

$$
H = \sum_{\vec{p},\vec{s}} p^0 (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} - b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}} - 1) \quad \text{(commutators used)}\,, \qquad (4.361)
$$

which indicates that b-particles carry negative energy. Since the commutation relation indicates that any number of quanta can occupy a given normal mode, we see that there is no lowest energy state which would have defined the vacuum. Thus, it was critical that we quantize the Dirac field by anticommutation relations in order to avoid the negative energy problem.

It is instructive to examine the origin of the negative sign on the second term $b_{\vec{p},\vec{s}} b^\dagger_{\vec{p},\vec{s}}$ in (4.359). This came about due to the first-order time derivative $\dot{\psi}$ that appear in $\mathcal{H}$ which picked up the factor $(-ip^0)$ from $f_{\vec{p},\vec{s}}$ and ip^0 from $g_{\vec{p},\vec{s}}$. The corresponding sign in the case of the Klein-Gordon field was positive as seen in (4.192), and this is because there are two time derivatives per term for the Klein-Gordon Hamiltonian $\mathcal{H} = (\dot{\phi}^2 - \phi\ddot{\phi})/2$.

One subtle point is left: when we expanded ϕ in (4.346), we could have labeled the coefficient of the second term as $b_{\vec{p},\vec{s}}$ instead of $b^\dagger_{\vec{p},\vec{s}}$ with all else identical including the anticommutation relations which are symmetric between $b_{\vec{p},\vec{s}}$ and $b^\dagger_{\vec{p},\vec{s}}$. Then, the

total energy (4.359) before any quantization condition is used would be

$$H = \sum_{\vec{p},\vec{s}} p^0 (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} - b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}}) \,, \tag{4.362}$$

which is already expressed in terms of number operators. Then, does this mean that the energy of the b-particle can be negative? Not so. We recall that the formulation of fermionic oscillator was symmetric under the exchange of $a^\dagger \leftrightarrow a$ and $|0\rangle \leftrightarrow |1\rangle$. Which to use was to be determined by defining the lower-energy state to be the empty state $|0\rangle$. The above Hamiltonian shows that the state $|1\rangle$ for a given $\vec{p}$ of b-particle contributes $-p^0 < 0$ to the total energy while the $|0\rangle$ contributes zero. This indicates that we have mislabeled $|0\rangle$ and $|1\rangle$ for the b-particles. Thus, we have to redefine as $|0\rangle \leftrightarrow |1\rangle$ and also $b^\dagger_{\vec{p},\vec{s}} \leftrightarrow b_{\vec{p},\vec{s}}$, which recovers the original derivation (4.360).

Just as in the case of the Klein-Gordon field, the total Hamiltonian (4.360) contains an apparent infinity $\sum_{\vec{p},\vec{s}}(-p^0)$. Again, we regard it as a harmless constant offset and choose to discard it by normal ordering. This time, however, we have to change the sign of the term when the creation and annihilation operators are swapped:

$$: a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} - b_{\vec{p},\vec{s}} b^\dagger_{\vec{p},\vec{s}} : = a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} + b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}} \,. \tag{4.363}$$

sign flip

Thus, the normal ordering procedure is now extended to include fermion operators: it simply reorders the annihilation and creation operators such that all creation operators are to the left of all annihilation operators, and add a minus sign if odd number of swaps of fermion operators are needed for the reordering. If both boson and fermion operators are present, it is assumed that fermion operators *commute* with boson operators. For example, if all are fermion operators,

$$: a_1 a_2^\dagger : = -a_2^\dagger a_1 \,, \quad : b_1 a_1 a_2^\dagger : = a_2^\dagger \underbrace{b_1 a_1}_{-a_1 b_1} = -a_2^\dagger a_1 b_1 \,, \text{ etc.} \tag{4.364}$$

Note that, when fermions are involved, one does have to keep track of the ordering among creation operators or that among annihilation operators.

Using this extended definition of normal ordering, the total Hamiltonian can be written as

$$H \equiv : \int d^3x \, \psi^\dagger i\partial_0 \psi : = \sum_{\vec{p},\vec{s}} p^0 (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} + b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}}) \,. \tag{4.365}$$

Similarly, the total momentum (4.338) can be expressed in terms of a's and b's as

$$\vec{P} \equiv : \int d^3x \, \psi^\dagger (-i\vec{\nabla}) \psi : = \sum_{\vec{p},\vec{s}} \vec{p} \, (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} + b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}}) \,. \tag{4.366}$$

Using the above expressions of $P^\mu = (H, \vec{P})$, it is straightforward to show

$$[P^\mu, \psi(x)] = -i\partial^\mu \psi(x)\,, \quad [P^\mu, \pi(x)] = -i\partial^\mu \pi(x)\,, \tag{4.367}$$

which can be extended to any polynomial function $F(\psi, \pi)$:

$$[P^\mu, F(\psi, \pi)] = -i\partial^\mu F(\psi, \pi)\,. \tag{4.368}$$

Thus, the total energy-momentum operators act as space-time translation operators just as in the case of the Klein-Gordon field.

Exercise 4.10 *Heisenberg's equation of motion for the Dirac field.*
Use the momentum expansion of Dirac field and the total Hamiltonian (expressed in terms of the number operators) to show that the Dirac field obeys Heisenberg's equation of motion (note the commutator, not anticommutator):

$$-i\dot{\psi}(x) = [H, \psi(x)]\,. \tag{4.369}$$

(*hint: You may find the following identity handy:* $[A, BC] = \{A, B\}C - B\{A, C\}$.)

Exercise 4.11 *Use the momentum expansion of the Dirac field to verify (4.366).*

The Noether current for the phase transformation
The Lagrangian $\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi$ is invariant under the phase rotation

$$\psi' = e^{i\theta}\psi \quad \rightarrow \quad \bar{\psi}' = e^{-i\theta}\bar{\psi}\,, \tag{4.370}$$

where θ is a real parameter and all four components of ψ are rotated by the same angle simultaneously. If each component of ψ is phase-rotated individually, then the Lagrangian is not invariant because of the off-diagonal terms such as $\psi_1^*\partial_\mu\psi_2$ (due to the off diagonal elements of the γ matrices). The derivation of the Noether current corresponding to this phase rotation is identical to the case of charged Klein-Gordon field (4.253) except that this time there are four fields ψ_n $(n = 1, 2, 3, 4)$ whose contribution to the variation of $\mathcal{L}$ should be added up:

$$\partial_\mu j^\mu = 0\,, \quad j^\mu \equiv i\Big(\underbrace{\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_n^*)}}_{0}\, \psi_n^* - \underbrace{\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_n)}}_{i(\bar{\psi}\gamma^\mu)_n}\, \psi_n \Big)\,, \tag{4.371}$$

namely,

$$\boxed{j^\mu = \bar{\psi}\gamma^\mu\psi}\,. \tag{4.372}$$

This is exactly the same current as the 'probability current' (3.66) we have derived directly from the Dirac equation in the non-quantized theory; we noted that the time component $j^0 = \psi^\dagger\psi$ was always positive and consistent with the interpretation that it is the probability density. We will now see that the same quantity in the framework of the quantum field theory is interpreted as the charge current, and that the conserved quantity - the total charge - can take both positive and negative values.

The conserved quantify is the space integral of the time component of a conserved current:

$$Q \equiv \int d^3x\, j^0 = \int d^3x\, \psi^\dagger\psi\,. \tag{4.373}$$

We use the momentum expansion of ψ and the orthonormality relations (4.348) to obtain

$$\begin{aligned} Q &= \int d^3x \sum_{\vec{p},\vec{s},\vec{p}',\vec{s}'} (a^\dagger_{\vec{p},\vec{s}} f^\dagger_{\vec{p},\vec{s}} + b_{\vec{p},\vec{s}} g^\dagger_{\vec{p},\vec{s}})(a_{\vec{p},\vec{s}} f_{\vec{p},\vec{s}} + b^\dagger_{\vec{p},\vec{s}} g_{\vec{p},\vec{s}}) \\ &= \sum_{\vec{p},\vec{s},\vec{p}',\vec{s}'} (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p}',\vec{s}'}\,\delta_{\vec{p},\vec{p}'}\delta_{\vec{s},\vec{s}'} + b_{\vec{p},\vec{s}} b^\dagger_{\vec{p}',\vec{s}'}\,\delta_{\vec{p},\vec{p}'}\delta_{\vec{s},\vec{s}'}) \\ &= \sum_{\vec{p},\vec{s}} (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} + \underbrace{b_{\vec{p},\vec{s}} b^\dagger_{\vec{p},\vec{s}}}) \\ &\qquad\qquad 1 - b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}} \quad \leftarrow\ \{b_{\vec{p},\vec{s}}\,, b^\dagger_{\vec{p},\vec{s}}\} = 1 \\ &= \sum_{\vec{p},\vec{s}} (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} - b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}} + 1)\,. \end{aligned} \tag{4.374}$$

We see that a-particles contribute to the quantity Q by $+1$ each regardless of momentum and spin, and b-particles contribute by -1 each. Thus, one can interpret that a-particles are electrons carrying $+1$ 'electron number' each and b-particles are positrons carrying -1 electron number each. The extra term $\sum_{\vec{p},\vec{s}} 1$ is again infinite, which reminds us of the *electron sea* in the hole theory. Again, we discard the constant offset by normal ordering:

$$:\! a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} + b_{\vec{p},\vec{s}} b^\dagger_{\vec{p},\vec{s}} \!: = a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} - b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}}\,; \tag{4.375}$$

namely,

$$Q \equiv :\!\int d^3x\, \psi^\dagger\psi \!: = \sum_{\vec{p},\vec{s}} (a^\dagger_{\vec{p},\vec{s}} a_{\vec{p},\vec{s}} - b^\dagger_{\vec{p},\vec{s}} b_{\vec{p},\vec{s}})\,. \tag{4.376}$$

Note that $\int d^3x\, \psi^\dagger\psi$ is a linear combination of terms which are product of creation and annihilation operators, and that the normal ordering applies directly to each term.

Noether current of rotational invariance (general)

In (3.244), the hole theory guided us to assign spin $+1/2$ to the spinor $v_{\vec{0},+\vec{s}}$ even though its egenvalue of the spin operator $\Sigma\cdot\vec{s}/2$ was $-1/2$. It is now time to show that the creation operator $b^\dagger_{\vec{0},\vec{s}}$ which is associated with $v_{\vec{0},+\vec{s}}$ indeed creates a state with spin $+1/2$ in the $\vec{s}$ direction. To do so, we have to find the total angular momentum operator which should contain the spin term. As we will see below, the orbital angular momentum $\vec{x}\times\vec{p}$ shows up as a part of the conserved quantity corresponding to the invariance of Lagrangian under rotation. The rest of the conserved quantity is then identified as the spin of the particle. We will first derive a general expression for the total angular momentum.

Our starting point is that the Lagrangian density is a Lorentz scalar:

$$\mathcal{L}'(x') = \mathcal{L}(x) \qquad (x' = \Lambda x)\,, \tag{4.377}$$

where Λ is a proper and orthochronous Lorentz transformation. Let us be specific about the meaning of this relation. In general, a Lagrangian density is a function of a set of fields $\tilde{\phi} = (\phi_1, \ldots, \phi_n)$ and its derivative $\partial_\mu\tilde{\phi}$:

$$\mathcal{L}(x) \stackrel{\text{def}}{\equiv} \mathcal{L}(\tilde{\phi}(x), \partial_\mu\tilde{\phi}(x)) \tag{4.378}$$

Under a Lorentz transformation, the field $\tilde{\phi}$ transforms by a certain $n\times n$ matrix S:

$$\phi'_a(x') = S_{ab}\phi_b(x) \qquad (a, b = 1, \ldots, n)\,. \tag{4.379}$$

The Lagrangian in the transformed frame $\mathcal{L}'(x')$ is then defined by *the same functional form* as $\mathcal{L}(\tilde{\phi}, \partial_\mu\tilde{\phi})$ where $\tilde{\phi}(x)$ is replaced by $\tilde{\phi}'(x')$ and x by x':

$$\mathcal{L}'(x') \stackrel{\text{def}}{\equiv} \mathcal{L}(\tilde{\phi}'(x'), \partial'_\mu\tilde{\phi}'(x'))\,. \tag{4.380}$$

We have already seen in (4.125) that, with these definitions, the spin-0 Lagrangian indeed satisfies the relation $\mathcal{L}'(x') = \mathcal{L}(x)$. It is also straightforward to show that the Dirac field Lagrangian $\bar{\psi}(i\partial\!\!\!/ - m)\psi$ is a Lorentz scalar.

Exercise 4.12 *Show that the Lagrangian density of free Dirac field $\bar{\psi}(i\partial\!\!\!/ - m)\psi$ is indeed a Lorentz scalar as defined above.*

As can be seen in the definitions (4.378) and (4.380), $\mathcal{L}(x)$ as a function of $(\tilde{\phi}(x), \partial_\mu\tilde{\phi}(x))$ and $\mathcal{L}'(x')$ as a function of $(\tilde{\phi}'(x'), \partial'_\mu\tilde{\phi}'(x'))$ have the same functional form; thus, we can express $\mathcal{L}'(x') - \mathcal{L}(x) = 0$ in terms of the changes in the arguments

$$\delta\phi_a \stackrel{\text{def}}{\equiv} \phi'_a(x') - \phi_a(x) \quad \text{and} \quad \delta(\partial_\mu\phi_a) \stackrel{\text{def}}{\equiv} \partial'_\mu\phi'_a(x') - \partial_\mu\phi_a(x)\,. \tag{4.381}$$

Namely,

$$\begin{aligned} 0 &= \mathcal{L}(\tilde{\phi}'(x'), \partial'_\mu \tilde{\phi}(x')) - \mathcal{L}(\tilde{\phi}(x), \partial_\mu \tilde{\phi}(x)) \\ &= \underbrace{\frac{\delta \mathcal{L}}{\delta \phi_a}}_{\partial_\mu \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)} \text{ (by equation of motion)}} \delta\phi_a + \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)} \delta(\partial_\mu \phi_a) \,. \end{aligned} \tag{4.382}$$

There are six independent generators $M^{\alpha\beta}$ for proper and orthochronous Lorentz transformations Λ, and correspondingly there should be six independent generators in the space of n fields $\tilde{\phi}$. Namely, corresponding to the infinitesimal Lorentz transformation (1.94)

$$\Lambda^\mu{}_\nu = g^\mu{}_\nu + \omega^\mu{}_\nu = g^\mu{}_\nu + \frac{1}{2}\omega_{\alpha\beta}(M^{\alpha\beta})^\mu{}_\nu \tag{4.383}$$

the matrix S can be written as

$$S_{ab} = \delta_{ab} + \frac{1}{2}\omega_{\alpha\beta}(T^{\alpha\beta})_{ab} \,, \tag{4.384}$$

where $T^{\alpha\beta}$ are the generators in the space of $\tilde{\phi}$ and are antisymmetric under the exchange $\alpha \leftrightarrow \beta$:

$$T^{\alpha\beta} = -T^{\beta\alpha} \tag{4.385}$$

making the number of independent generators to be six. Then $\delta\phi_a$ and $\delta(\partial_\mu \phi_a)$ are

$$\begin{aligned} \delta\phi_a &= \phi'_a(x') - \phi_a(x) \\ &= S_{ab}\phi_b(x) - \phi_a(x) \\ &= \frac{1}{2}\omega_{\alpha\beta}(T^{\alpha\beta})_{ab}\phi_b(x) \end{aligned} \tag{4.386}$$

$$\begin{aligned} \delta(\partial_\mu \phi_a) &= \partial'_\mu \phi'_a(x') - \partial_\mu \phi_a(x) \\ &= (\Lambda_\mu{}^\nu \partial_\nu) S_{ab}\phi_b(x) - \partial_\mu \phi_a(x) \\ &= (g_\mu{}^\nu + \omega_\mu{}^\nu)\partial_\nu \Big(\delta_{ab} + \frac{1}{2}\omega_{\alpha\beta}(T^{\alpha\beta})_{ab}\phi_b(x)\Big) - \partial_\mu \phi_a(x) \\ &= (\partial_\mu + \omega_\mu{}^\nu \partial_\nu)\Big(\phi_a(x) + \frac{1}{2}\omega_{\alpha\beta}(T^{\alpha\beta})_{ab}\phi_b(x)\Big) - \partial_\mu \phi_a(x) \\ &= \omega_\mu{}^\nu \partial_\nu \phi_a(x) + \frac{1}{2}\omega_{\alpha\beta}(T^{\alpha\beta})_{ab}\partial_\mu \phi_b(x) \,, \end{aligned} \tag{4.387}$$

where in the last step we discarded the term that is second order in ω. Using these in (4.382),

$$0 = \left(\partial_\mu \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)}\right)\Big(\frac{1}{2}\omega_{\alpha\beta}(T^{\alpha\beta})_{ab}\phi_b\Big) + \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)}\Big(\omega_\mu{}^\nu \partial_\nu \phi_a + \frac{1}{2}\omega_{\alpha\beta}(T^{\alpha\beta})_{ab}\partial_\mu \phi_b\Big)$$

$$
\begin{aligned}
&= \partial_\mu \left(\frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)} \frac{1}{2} \omega_{\alpha\beta} (T^{\alpha\beta})_{ab} \phi_b \right) + \underbrace{\frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)} \omega_{\mu\nu} \partial^\nu \phi_a}_{\displaystyle \frac{\delta \mathcal{L}}{\delta(\partial_\alpha \phi_a)} \omega_{\alpha\beta} \partial^\beta \phi_a} \\
&= \omega_{\alpha\beta} \left[\partial_\mu \left(\frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)} \frac{(T^{\alpha\beta})_{ab}}{2} \phi_b \right) + \frac{\delta \mathcal{L}}{\delta(\partial_\alpha \phi_a)} \partial^\beta \phi_a \right] . \qquad (4.388)
\end{aligned}
$$

This holds for any (small) $\omega_{\alpha\beta}$. Then, can we set the quantity inside the square bracket to be zero? No, since $\omega^{\alpha\beta}$'s are antisymmetric with respect to (α, β) and are not independent. We can, however, pick a specific pair (α, β) and set

$$
\omega_{\mu\nu} = 0 \quad \text{for all } \mu \text{ and } \nu, \text{ except} \quad \omega_{\alpha\beta} = -\omega_{\beta\alpha} \, . \qquad (4.389)
$$

Then using $T^{\alpha\beta} = -T^{\beta\alpha}$, (4.388) becomes (no sum over α, β)

$$
\begin{aligned}
0 &= \omega_{\alpha\beta} \left[\partial_\mu \left(\frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)} \frac{(T^{\alpha\beta})_{ab}}{2} \phi_b \right) + \frac{\delta \mathcal{L}}{\delta(\partial_\alpha \phi_a)} \partial^\beta \phi_a \quad - \quad (\alpha \leftrightarrow \beta) \right] \\
&= \omega_{\alpha\beta} \Bigg[\partial_\mu \left(\frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)} \frac{(T^{\alpha\beta})_{ab} - (T^{\beta\alpha})_{ab}}{2} \phi_b \right) + \frac{\delta \mathcal{L}}{\delta(\partial_\alpha \phi_a)} \partial^\beta \phi_a - \frac{\delta \mathcal{L}}{\delta(\partial_\beta \phi_a)} \partial^\alpha \phi_a \\
&\qquad\qquad \underbrace{- g^{\alpha\beta} \mathcal{L} \qquad\qquad + g^{\alpha\beta} \mathcal{L}}_{0} \Bigg] \\
&= \omega_{\alpha\beta} \left[\partial_\mu \left(\frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)} (T^{\alpha\beta})_{ab} \phi_b \right) + J^{\alpha\beta} - J^{\beta\alpha} \right] \qquad (4.390)
\end{aligned}
$$

where $J^{\alpha\beta}$ is the energy momentum tensor defined in (4.116). Using the conservation of $J^{\alpha\beta}$, $J^{\alpha\beta} - J^{\beta\alpha}$ can be written as

$$
\begin{aligned}
\partial_\mu (x^\alpha J^{\mu\beta} - x^\beta J^{\mu\alpha}) &= \delta_{\mu\alpha} J^{\mu\alpha} + x^\alpha \underbrace{\partial_\mu J^{\mu\alpha}}_{0} + \delta_{\mu\beta} J^{\mu\beta} + x^\beta \underbrace{\partial_\mu J^{\mu\beta}}_{0} \\
&= J^{\alpha\beta} - J^{\beta\alpha} \, . \qquad (4.391)
\end{aligned}
$$

Then (4.390) becomes

$$
0 = \omega_{\alpha\beta} \partial_\mu M^{\mu\alpha\beta} \qquad (4.392)
$$

with

$$
\boxed{M^{\mu\alpha\beta} \equiv \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_a)} (T^{\alpha\beta})_{ab} \phi_b + (x^\alpha J^{\mu\beta} - x^\beta J^{\mu\alpha})} \, . \qquad (4.393)
$$

Since there is no sum over α and β, we immeidately obtain the conservation relation

$$
\boxed{\partial_\mu M^{\mu\alpha\beta} = 0} \, , \qquad (4.394)
$$

which is valid for any (α, β) since the choice of (α, β) was arbitrary in (4.389). Note that $M^{\mu\alpha\beta}$ is also antisymmetric under the exchange $\alpha \leftrightarrow \beta$.

The conserved quantity corresponding to the rotation around the k-th axis is then $(i, j, k$: cyclic)

$$\begin{aligned} J^k &= \int d^3x\, M^{0ij} \\ &= \int d^3x \left[\frac{\delta \mathcal{L}}{\delta \dot{\phi}_a} (T^{ij})_{ab} \phi_b + (x^i J^{0j} - x^j J^{oi}) \right] \\ &= \int d^3x\, \pi_a (T^{ij})_{ab} \phi_b + \int d^3x\, (\vec{x} \times \vec{J}^0)_k \end{aligned} \tag{4.395}$$

where $\vec{J}^0 \equiv (J^{01}, J^{02}, J^{03})$ is the momentum density and used the definition of conjugate field $\pi_a \equiv \delta \mathcal{L}/\delta \dot{\phi}_a$. The second term has the form of orbital angular momentum around the origin, and the first term is then interpreted as the spin angular momentum carried by the particle. The quantity J^k is the total angular momentum which is a constant of motion only after spin of the particle is added to the orbital angular momentum.

You may be wondering what are the 'conserved quantities' of (4.394) for $\alpha = 0$ or $\beta = 0$ which should correspond to the generators of boost. The time derivative of such quantities would indeed be zero; because of the explicit time dependence that appear in the definition (4.393), however, such quantities are not genuine physical conserved quantities.

Spin of electron

We can now apply (4.395) to the Dirac field Lagrangian to obtain the corresponding total angular momentum operator. Using the definition (4.116), the momentum density J^{0i} is (using $\pi = i\psi^\dagger$)

$$J^{0i} = \frac{\partial \mathcal{L}}{\partial \dot{\psi}_a} \partial^i \psi_a - g^{0i} \mathcal{L} = \pi_a \partial^i \psi_a = \psi^\dagger (-i \nabla_i) \psi \,. \tag{4.396}$$

The generators in the spinor space are $T^{ij} = B^{ij} = -i\Sigma_k/2$ (3.180). The total angular momentum (4.395) is then

$$J^k = \int d^3x\, \psi^\dagger \left(\frac{\Sigma_k}{2} + (\vec{x} \times \vec{p})_k \right) \psi \qquad (\vec{p} \equiv -i\vec{\nabla}) \,, \tag{4.397}$$

and the component along some fixed direction $\vec{s}$ is given by

$$\vec{s} \cdot \vec{J} = \int d^3x\, \psi^\dagger \left(\frac{\vec{s} \cdot \vec{\Sigma}}{2} + \vec{s} \cdot (\vec{x} \times \vec{p}) \right) \psi \,. \tag{4.398}$$

We will now evaluate the expectation value of this operator for a state that represents a positron at rest: $b^\dagger_{\vec{0},\vec{s}}|0\rangle$ where $\vec{s}$ is taken to be the same $\vec{s}$ that is used in $\vec{s}\cdot\vec{J}$. Momentum-expanding the fields and recalling the implicit normal ordering, we get

$$\begin{aligned}
&\langle 0|b_{\vec{0},\vec{s}}(\vec{s}\cdot\vec{J})b^\dagger_{\vec{0},\vec{s}}|0\rangle \\
&\quad= \langle 0|b_{\vec{0},\vec{s}}\sum_{\vec{p},\vec{s}_1,\vec{p}',\vec{s}_2}\int d^3x\ :(\cancel{a^\dagger_{\vec{p},\vec{s}_1}f^\dagger_{\vec{p},\vec{s}_1}}+b_{\vec{p},\vec{s}_1}g^\dagger_{\vec{p},\vec{s}_1}) \\
&\qquad\times\left(\frac{\vec{s}\cdot\vec{\Sigma}}{2}+\vec{s}\cdot(\vec{x}\times\vec{p})\right)(\cancel{a_{\vec{p}',\vec{s}_2}f_{\vec{p}',\vec{s}_2}}+b^\dagger_{\vec{p}',\vec{s}_2}g_{\vec{p}',\vec{s}_2}):b^\dagger_{\vec{0},\vec{s}}|0\rangle \\
&\quad= \sum_{\vec{p},\vec{s}_1,\vec{p}',\vec{s}_2}\langle 0|b_{\vec{0},\vec{s}}\underbrace{:b_{\vec{p},\vec{s}_1}b^\dagger_{\vec{p}',\vec{s}_2}:}_{\underbrace{-b^\dagger_{\vec{p}',\vec{s}_2}b_{\vec{p},\vec{s}_1}}_{-\delta_{\vec{0},\vec{p}'}\delta_{\vec{s},\vec{s}_2}\delta_{\vec{p},\vec{0}}\delta_{\vec{s}_1,\vec{s}}}}b^\dagger_{\vec{0},\vec{s}}|0\rangle\int d^3x\,g^\dagger_{\vec{p},\vec{s}_1}\left(\frac{\vec{s}\cdot\vec{\Sigma}}{2}+\vec{s}\cdot(\vec{x}\times\vec{p})\right)g_{\vec{p}',\vec{s}_2} \\
&\quad= -\int d^3x\,g^\dagger_{\vec{0},\vec{s}}\left(\frac{\vec{s}\cdot\vec{\Sigma}}{2}+\vec{s}\cdot(\vec{x}\times\vec{p})\right)g_{\vec{0},\vec{s}}\,. \qquad (4.399)
\end{aligned}$$

Since applying $\vec{p}=-i\vec{\nabla}$ on $g_{\vec{0},\vec{s}}$ annihilates it and

$$(\vec{s}\cdot\vec{\Sigma})g_{\vec{0},\vec{s}}=-g_{\vec{0},\vec{s}} \qquad (4.400)$$

by construction [see (3.261)], we have

$$\langle 0|b_{\vec{0},\vec{s}}(\vec{s}\cdot\vec{J})b^\dagger_{\vec{0},\vec{s}}|0\rangle=\frac{1}{2}\int d^3x\,g^\dagger_{\vec{0},\vec{s}}g_{\vec{0},\vec{s}}=\frac{1}{2}\,. \qquad (4.401)$$

Thus, the angular momentum of the positron state $b^\dagger_{\vec{0},\vec{s}}|0\rangle$ is indeed $+1/2$ in the $\vec{s}$ direction. Note the crucial minus sign when the normal ordering was applied to $b_{\vec{p},\vec{s}_1}b^\dagger_{\vec{p}',\vec{s}_2}$ which cancelled the minus sign in (4.400). When the state is changed to an electron at rest, then all needed is to change $b^{(\dagger)}_{\vec{0},\vec{s}}$ to $a^{(\dagger)}_{\vec{0},\vec{s}}$ in the procedure above. One sees that the relevant product $a^\dagger_{\vec{p},\vec{s}_1}a_{\vec{p}',\vec{s}_2}$ in the angular momentum operator is already normal ordered and there is no minus sign in (4.400) when g is replaced by f, and one obtains

$$\langle 0|a_{\vec{0},\vec{s}}(\vec{s}\cdot\vec{J})a^\dagger_{\vec{0},\vec{s}}|0\rangle=\frac{1}{2}\,. \qquad (4.402)$$

Thus, the spin indexes of $a^\dagger_{\vec{p},\vec{s}}$ and $b^\dagger_{\vec{p},\vec{s}}$ both correctly represent the corresponding physical spin.

Microscopic causality of the Dirac field

We have seen that the Klein-Gordon field satisfied microscopic causality; namely, field

operators $\phi(x)$ and $\phi(y)$ commuted if x and y are separated by a space-like distance:

$$[\phi(x), \phi(y)] = i\Delta(x-y) = 0\,, \quad \text{if } (x-y)^2 < 0\,, \tag{4.403}$$

where $i\Delta(x)$ is given by (4.284). In the case of the Dirac field also, in order for two measurements to be independent, the corresponding operators A and B should *commute* and not *anticommute*. In fact, independence of two meansurements means that for any eigenvalue a of the operator A and any eigenvalue b of the operator B, there exists a simultaneous eigenstate $|a,b\rangle$ that allows measurements of the two observables with infinite accuracy. Then, the operators A and B should commute:

$$\begin{aligned} (AB-BA)|a,b\rangle &= (ab-ba)|a,b\rangle = 0 \quad (\text{for all } a,b) \\ &\rightarrow \quad [A,B] = 0\,. \end{aligned} \tag{4.404}$$

On the other hand, the operator relations for the Dirac fields are given as anticommutators, and commutators such as $[\psi_n(x), \psi_m^\dagger(y)]$ do not vanish for $(x-y)^2 < 0$; in fact, it is a messy expression of a's and b's and not even a c-number. Physically relevant quantities, however, always appear as bilinear covariants which have the form

$$\bar\psi\,\Gamma\psi = \Gamma_{nm}\bar\psi_n\psi_m\,, \tag{4.405}$$

where Γ is a 4×4 matrix. Thus, if

$$[\bar\psi_n(x)\psi_m(x), \bar\psi_k(y)\psi_l(y)] = 0 \quad (x-y)^2 < 0 \quad (\text{for any } n,m,k,l), \tag{4.406}$$

then we conclude that microscopic causality is satisfied for the Dirac field. In proving this, we first note that *if* field components at x, $\bar\psi_n(x)$ and $\psi_m(x)$, *anticommute* with those at y, $\bar\psi_k(y)$ and $\psi_l(y)$, then $\bar\psi_n(x)\psi_m(x)$ commutes with $\bar\psi_k(y)\psi_l(y)$:

$$\begin{aligned} \overleftarrow{\bar\psi_n(x)\psi_m(x)}\,\boxed{\bar\psi_k(y)}\psi_l(y) &= \bar\psi_k(y)\,\overleftarrow{\bar\psi_n(x)\psi_m(x)}\boxed{\psi_l(y)} \\ &= \bar\psi_k(y)\psi_l(y)\;\bar\psi_n(x)\psi_m(x)\,. \end{aligned} \tag{4.407}$$

Thus, we need to show that, for any $n,m = 1,2,3,4$,

$$\begin{cases} \{\psi_n(x), \psi_m(y)\} = 0 \\ \{\bar\psi_n(x), \bar\psi_m(y)\} = 0 \\ \{\psi_n(x), \bar\psi_m(y)\} = 0 \end{cases} \quad \text{for } (x-y)^2 < 0\,. \tag{4.408}$$

First two are trivial since the only non-zero anticommutators are of the type $\{a, a^\dagger\}$ or $\{b, b^\dagger\}$:

$$\begin{aligned} \psi \text{ contains } a \text{ and } b^\dagger \text{ only} \quad &\rightarrow \quad \{\psi_n(x), \psi_m(y)\} = 0\,, \\ \bar\psi \text{ contains } a^\dagger \text{ and } b \text{ only} \quad &\rightarrow \quad \{\bar\psi_n(x), \bar\psi_m(y)\} = 0\,. \end{aligned} \tag{4.409}$$

The evaluation of the anticommutator $\{\psi_n(x), \bar{\psi}_m(y)\}$ requires some care. For convenience, let's extend the definition of matrix of the form ab^T, where a and b are column vectors, to the anticommutators (and similarly to the commutators): $\{a, b^T\}$ is a matrix whose components are defined by

$$\{a, b^T\}_{nm} \stackrel{\text{def}}{\equiv} \{a_n\,, b_m\}\,. \tag{4.410}$$

Then, using the division of ψ into creation and annihilation parts (4.352),

$$\begin{aligned}\{\psi(x), \bar{\psi}(y)\} &= \{\psi_a(x) + \psi_{b^\dagger}(x)\,, \bar{\psi}_a(y) + \bar{\psi}_{b^\dagger}(y)\} \\ &= \{\psi_a(x)\,, \bar{\psi}_a(y)\} + \{\psi_{b^\dagger}(x)\,, \bar{\psi}_{b^\dagger}(y)\}\,.\end{aligned} \tag{4.411}$$

Following the procedure similar to (4.355),

$$\begin{aligned}\{\psi_a(x)\,, \bar{\psi}_a(y)\} &= \sum_{\vec{p},\vec{s},\vec{p}',\vec{s}'} \underbrace{\{a_{\vec{p},\vec{s}}\,, a^\dagger_{\vec{p}',\vec{s}'}\}}_{\delta_{\vec{p},\vec{p}'}\delta_{\vec{s},\vec{s}'}} f_{\vec{p},\vec{s}}(x)\overline{f}_{\vec{p}',\vec{s}'}(y) = \sum_{\vec{p},\vec{s}} f_{\vec{p},\vec{s}}(x)\overline{f}_{\vec{p},\vec{s}}(y) \\ &= \underbrace{\sum_{\vec{p}} \frac{1}{2p^0 V}}_{\frac{1}{(2\pi)^3}\int \frac{d^3p}{2p^0}} \underbrace{\sum_{\vec{s}} u_{\vec{p},\vec{s}}\,\bar{u}_{\vec{p},\vec{s}}\, e^{-ip\cdot(x-y)}}_{(\underbrace{\not{p}}_{i\not{\partial}_x} + m) \text{ by (3.316)}} \\ &= (i\not{\partial}_x + m)\frac{1}{(2\pi)^3}\int \frac{d^3p}{2p^0} e^{-ip\cdot(x-y)} \\ &= (i\not{\partial}_x + m)\Delta_+(x-y)\,,\end{aligned} \tag{4.412}$$

where $\partial^\mu_x \equiv \partial/\partial x_\mu$ (namely, operates on x and not on y). Similarly,

$$\begin{aligned}\{\psi_{b^\dagger}(x)\,, \bar{\psi}_{b^\dagger}(y)\} &= \sum_{\vec{p},\vec{s},\vec{p}',\vec{s}'} \underbrace{\{b^\dagger_{\vec{p},\vec{s}}\,, b_{\vec{p}',\vec{s}'}\}}_{\delta_{\vec{p},\vec{p}'}\delta_{\vec{s},\vec{s}'}} g_{\vec{p},\vec{s}}(x)\overline{g}_{\vec{p}',\vec{s}'}(y) = \sum_{\vec{p},\vec{s}} g_{\vec{p},\vec{s}}(x)\overline{g}_{\vec{p},\vec{s}}(y) \\ &= \underbrace{\sum_{\vec{p}} \frac{1}{2p^0 V}}_{\frac{1}{(2\pi)^3}\int \frac{d^3p}{2p^0}} \underbrace{\sum_{\vec{s}} v_{\vec{p},\vec{s}}\,\bar{v}_{\vec{p},\vec{s}}\, e^{ip\cdot(x-y)}}_{(\underbrace{\not{p}}_{-i\not{\partial}_x} - m) \text{ by (3.316)}} \\ &= -(i\not{\partial}_x + m)\frac{1}{(2\pi)^3}\int \frac{d^3p}{2p^0} e^{ip\cdot(x-y)} \\ &= -(i\not{\partial}_x + m)\Delta_+(y-x)\,.\end{aligned} \tag{4.413}$$

Then, we have

$$\begin{aligned}\{\psi(x), \bar{\psi}(y)\} &= (i\not{\partial}_x + m)\left[\Delta_+(x-y) - \Delta_+(y-x)\right] \\ &= (i\not{\partial}_x + m) i\Delta(x-y) \\ &\stackrel{\text{def}}{\equiv} -iS(x-y)\end{aligned} \tag{4.414}$$

Now, the function $i\Delta(x-y)$ is already known to vanish for $(x-y)^2 < 0$; thus, $\{\psi(x), \bar{\psi}(y)\} = 0$ for $(x-y)^2 < 0$, and all bilinear covariants commute at space-like distances. Thus, the quantized Dirac field satisfies microscopic causality.

What would have happened to the microscopic causality if we had quantized the Dirac field with commutators instead of anticommutators? Then, we would have

$$[a_{\vec{p},\vec{s}}\,, a^\dagger_{\vec{p}',\vec{s}'}] = [b_{\vec{p},\vec{s}}\,, b^\dagger_{\vec{p}',\vec{s}'}] = \delta_{\vec{p},\vec{p}'}\delta_{\vec{s},\vec{s}'}\,, \tag{4.415}$$

and all other *commutators* would be zero. The momentum expansion in terms of normal modes would be the same as before, and we would evaluate $[\psi_n(x), \bar{\psi}_m(y)]$ instead of the corresponding anticommutator (which would be a mess in this case). The critical difference occurs when we use $[b^\dagger_{\vec{p},\vec{s}}\,, b_{\vec{p}',\vec{s}'}] = -[b_{\vec{p}',\vec{s}'}\,, b^\dagger_{\vec{p},\vec{s}}] = -\delta_{\vec{p},\vec{p}'}\delta_{\vec{s},\vec{s}'}$ which will change the sign in the first line of (4.413), and as a result we would obtain

$$[\psi(x), \bar{\psi}(y)] = (i\partial\!\!\!/_x + m)\,[\Delta_+(x-y) + \Delta_+(y-x)]\,, \tag{4.416}$$

which does not vanish for $(x-y)^2 < 0$ since $\Delta_+(x-y) = \Delta^*_+(y-x)$ and $\Delta_+(x-y)$ is real and positive in the space-like region. Thus, in addition to the positive definiteness of the energy, the microscopic causality also requires that the Dirac field be quantized by anticommutators. How did it work out for the case of the Klein-Gordon field? There, we did not have the minus sign in the commutator corresponding to (4.413) and the necessary relative minus sign between $\Delta_+(x-y)$ and $\Delta_+(y-x)$ came from the property of commutator $[\phi^\dagger_a(x), \phi_a(y)] = -[\phi_a(y), \phi^\dagger_a(x)]$ used in (4.281). In 1940, Pauli extended the above argument of microscopic causality to particles with general spins, and established the fundamental connection between spin and statistics:

$$\begin{aligned} \text{Integer spin} &\;\leftrightarrow\; \text{Bose-Einstein statistics,} \\ \text{Half-integer spin} &\;\leftrightarrow\; \text{Fermi-Dirac statistics,} \end{aligned} \tag{4.417}$$

where 'Bose-Einstein statistics' means that the field must be quantized by commutators and thus a given state can be occupied by any number of particles, and 'Fermi-Dirac statistics' means that the field must be quantized by anticommutators and thus a given state can be occupied by at most one particle.

Lorentz invariance of the Dirac-field quantization

Let's now turn to the question of whether systems quantized in different Lorentz frames are equivalent or not. As in the case of the Klein-Gordon field, we could prove the equivalence of the anticommutation relations of fields in different frames, or those of creation and annihilation operators. Let's take the fields this time. First, we will see below that the general-time anticommutation relation (4.409) and (4.414) reduces to the equal-time anticommutation relation (4.358) when we set $x^0 = y^0$. In fact, setting $x^0 = y^0 = t$ in (4.409), we obtain

$$\{\psi_n(t,\vec{x}), \psi_m(t,\vec{y})\} = 0\,, \quad \{\pi_n(t,\vec{x}), \pi_m(t,\vec{y})\} = 0\,. \tag{4.418}$$

Setting $x^0 = y^0$ in $\{\psi(t,\vec{x}), \bar{\psi}(t,\vec{y})\}$ is equivalent to repeating the derivation (4.414) with $x^0 = y^0$ from the beginning. We start from just before $\not{p}$ is replaced by the differential operator in (4.412) and (4.413):

$$\begin{aligned}
\{\psi(x)\,, \bar{\psi}(y)\}_{x^0=y^0} &= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2p^0}\big[\underbrace{(\not{p}}_{p^0\gamma^0 - \vec{p}\cdot\vec{\gamma}} + m)e^{i\vec{p}\cdot(\vec{x}-\vec{y})} + \underbrace{(\not{p}}_{\text{relabel } \vec{p}\to -\vec{p}} - m)e^{-i\vec{p}\cdot(\vec{x}-\vec{y})}\big] \\
&= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2p^0}\big[\underbrace{(p^0\gamma^0 - \vec{p}\cdot\vec{\gamma} + m) + (p^0\gamma^0 + \vec{p}\cdot\vec{\gamma} - m)}_{2p^0\gamma^0}\big]e^{i\vec{p}\cdot(\vec{x}-\vec{y})} \\
&= \gamma^0\frac{1}{(2\pi)^3}\int d^3p\, e^{i\vec{p}\cdot(\vec{x}-\vec{y})} = \gamma^0\delta^3(\vec{x}-\vec{y})\,, \qquad (4.419)
\end{aligned}$$

where $\vec{\gamma} \stackrel{\text{def}}{\equiv} (\gamma^1, \gamma^2, \gamma^3)$. Now, the definition (4.410) leads to

$$\Big(\{a, b^T\}M\Big)_{nm} = \underbrace{\{a, b^T\}_{nk}}_{\{a_n, b_k\}} M_{km} = \{a_n,\ \underbrace{b_k M_{km}}_{(b^T M)_m}\} = \{a, b^T M\}_{nm}\,; \qquad (4.420)$$

where a_n and b_m are operators while M_{nm} is a c-number; namely,

$$\{a, b^T\}M = \{a, b^T M\}\,, \quad \text{and similarly,} \quad M\{a, b^T\} = \{Ma, b^T\}\,. \qquad (4.421)$$

Thus, we have

$$\{\psi(t,\vec{x}), \underbrace{\psi^\dagger(t,\vec{y})}_{\bar{\psi}\gamma^0}\} = \underbrace{\{\psi(t,\vec{x}), \bar{\psi}(t,\vec{y})\}}_{\gamma^0\delta^3(\vec{x}-\vec{y})}\gamma^0 = I\delta^3(\vec{x}-\vec{y})\,. \qquad (4.422)$$

Since $\pi = i\psi^\dagger$, this is equivalent to $\{\psi(t,\vec{x}), \pi(t,\vec{y})\} = iI\delta^3(\vec{x}-\vec{y})$ which is the equal-time quantization condition.

Thus, the Lorentz invariance of the quantization procedure is proven if we can show that

$$\begin{cases} \{\psi_n(x), \psi_m(y)\} = 0 \\ \{\bar{\psi}_n(x), \bar{\psi}_m(y)\} = 0 \\ \{\psi(x), \bar{\psi}(y)\} = -iS(x-y) \end{cases} \qquad (4.423)$$

in one frame leads to the same relations in another frame where the field in the new frame are related to the original field by the spinor representation of Lorentz transformation:

$$\psi'(x') = S(\Lambda)\psi(x)\,, \qquad \psi'(y') = S(\Lambda)\psi(y)\,, \qquad (4.424)$$

where $x' = \Lambda x$ and $y' = \Lambda y$ (Λ: proper and orthochronous). The first two relations are trivial:

$$\{\psi'_n(x'), \psi'_m(x')\} = \{S_{nk}\psi_k(x), S_{ml}\psi_l(x)\} = S_{nk}S_{ml}\{\psi_k(x), \psi_l(y)\} = 0\,, \qquad (4.425)$$

and similarly $\{\bar\psi'_n(x'), \bar\psi'_m(x')\} = 0$. Using (4.421),

$$\begin{aligned}
\{\psi'(x'), \bar\psi'(y')\} &= \{S\psi(x), \bar\psi(y)\bar S\} = S\underbrace{\{\psi(x), \bar\psi(y)\}}_{(i\partial\!\!\!/_x + m)i\Delta(x-y)}\bar S \\
&= (i\partial_{x\mu}\underbrace{S\gamma^\mu\bar S}_{\Lambda_\nu{}^\mu\gamma^\nu} + m)\underbrace{i\Delta(x-y)}_{i\Delta(x'-y')} \\
&= (i\partial_{x'\nu}\gamma^\nu + m)i\Delta(x'-y') \\
&= -iS(x'-y')\,, \qquad (4.426)
\end{aligned}$$

where we have used the Lorentz invariance of the function $i\Delta$: $i\Delta(x') = i\Delta(x)$, the property $S\gamma^\mu\bar S = \Lambda_\nu{}^\mu\gamma^\nu$ (3.272), and $\partial_{x'\nu} = \Lambda_\nu{}^\mu\partial_{x\mu}$. Thus, the same set of anticommutation relations are satisfied in the new frame, and the Lorentz invariance of the quantization procedure is proven.

Incidentally, $\Delta_+(x)$, and thus $i\Delta(x) = \Delta_+(x) - \Delta_+(-x)$, is a solution of the Klein-Gordon equation:

$$(\partial^2 + m^2)\Delta_+(x) = \frac{1}{(2\pi)^3}\int\frac{d^3p}{2p^0}\underbrace{(\partial^2 + m^2)}_{(-p^2+m^2)=0}e^{-ip\cdot x} = 0\,, \qquad (4.427)$$

and $-iS(x)$ is a solution of the Dirac equation:

$$\begin{aligned}
(i\partial\!\!\!/ - m)(-iS(x)) &= \underbrace{(i\partial\!\!\!/ - m)(i\partial\!\!\!/ + m)}_{(-\underbrace{\partial\!\!\!/\partial\!\!\!/}_{\partial^2} - m^2)}i\Delta(x) \\
&= -(\partial^2 + m^2)i\Delta(x) = 0\,. \qquad (4.428)
\end{aligned}$$

Problems

4.1 *Momentum expansion.*
In the momentum expansion of a hermitian Klein-Gordon field

$$\phi(x) = \sum_{\vec{p}} \Big(a_{\vec{p}}\, e_{\vec{p}}(x) + a^\dagger_{\vec{p}}\, e^*_{\vec{p}}(x) \Big),$$

suppose one takes the normalization of the normal mode functions as

$$e_{\vec{p}}(x) \stackrel{\text{def}}{\equiv} \frac{e^{-ip\cdot x}}{\sqrt{2c_{\vec{p}}V}}$$

where $c_{\vec{p}}$ is some real function of $\vec{p}$. This defines the normalization of the operators $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$.
(a) Obtain the orthonormality relations of the normal mode functions; namely, calculate $\int d^3x\, e^*_{\vec{p}}(x) e_{\vec{p}'}(x)$ and $\int d^3x\, e_{\vec{p}}(x) e_{\vec{p}'}(x)$.
(b) Express $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$ in terms of ϕ and π. Note that the relation $\pi = \dot{\phi}$ is a result of the Heisenberg's equation of motion for ϕ which results from the commutation relations among ϕ and π and not affected by the normalization of the normal modes.
(c) Require that $[a_{\vec{p}}, a^\dagger_{\vec{p}'}] = \delta_{\vec{p},\vec{p}'}$, and find the correct normalization factor $c_{\vec{p}}$.

4.2 *Lagrangian density for the Dirac field.*
The Lagrangian density for the Dirac field ψ is given by

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi\,.$$

In the text, we have derived the Dirac equation directly from the action principle. This time, apply the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial \phi_k} = \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_k)}$$

to each component of ψ; namely take ψ_a $(a = 1, \ldots, 4)$ as ϕ_k, to derive the Dirac's equation the Dirac equation for $\bar{\psi}$:

$$\bar{\psi}(i \overleftarrow{\partial\!\!\!/} + m) = 0\,,$$

and apply it for $\psi^\dagger_a$ $(a = 1, \ldots, 4)$ to obtain

$$(i\partial\!\!\!/ - m)\psi = 0\,.$$

Use the same Lagrangian density given above for both cases, and also do not resort to the Dirac or other explicit representation.

4.3 *Continous-$\vec{p}$ formalisms for Dirac field.*
For continuous momentum, the normal mode functions of the Dirac field are given by

$$f_{\vec{p},\vec{s}}(x) \equiv \frac{u_{\vec{p},\vec{s}} e^{-ip\cdot x}}{\sqrt{(2\pi)^3 2p^0}}\,, \quad g_{\vec{p},\vec{s}}(x) \equiv \frac{v_{\vec{p},\vec{s}} e^{ip\cdot x}}{\sqrt{(2\pi)^3 2p^0}}\,,$$

and the momentum expansion is

$$\psi(x) = \sum_{\vec{s}} \int d^3p (a_{\vec{p},\vec{s}} f_{\vec{p},\vec{s}}(x) + b^\dagger_{\vec{p},\vec{s}} g_{\vec{p},\vec{s}}(x))\,.$$

First, show that the following normalizations hold:

$$\begin{aligned} \int d^3x f^\dagger_{\vec{p},\vec{s}}(x) f_{\vec{p}',\vec{s}'}(x) &= \int d^3x g^\dagger_{\vec{p},\vec{s}}(x) g_{\vec{p}',\vec{s}'}(x) = \delta^3(\vec{p}-\vec{p}')\delta_{\vec{s},\vec{s}'} \\ \int d^3x f^\dagger_{\vec{p},\vec{s}}(x) g_{\vec{p}',\vec{s}'}(x) &= \int d^3x g^\dagger_{\vec{p},\vec{s}}(x) f_{\vec{p}',\vec{s}'}(x) = 0 \end{aligned}\,.$$

Then, repeat the derivation of

$$\{\psi(x), \bar{\psi}(y)\} = -iS(x-y)\,, \quad \textit{with} \quad -iS(x) \stackrel{\text{def}}{\equiv} (i\partial\!\!\!/ + m) i\Delta(x)\,.$$

Use the anticommutation relations

$$\begin{gathered} \{a_{\vec{p},\vec{s}}\,, a^\dagger_{\vec{p}',\vec{s}'}\} = \{b_{\vec{p},\vec{s}}\,, b^\dagger_{\vec{p}',\vec{s}'}\} = \delta^3(\vec{p}-\vec{p}')\delta_{\vec{s},\vec{s}'} \\ \textit{all others} = 0\ . \end{gathered}$$