1 Free Fields: A Recap

Before developing the perturbative machinery of Feynman diagrams, it is essential to have a firm grasp of the formalism for free quantum fields. This chapter provides a self-contained summary of the main results for scalar, Dirac, and photon fields: Lagrangians and Hamiltonians, field decompositions in terms of creation and annihilation operators, (anti)commutation relations, and propagators. We use natural units $\hbar = c = 1$ throughout, and work with continuum normalization (as opposed to box normalization, for instance in [1]).

1.1 The real scalar field

The Lagrangian density for a real scalar field $\phi(x) = \phi^\dagger(x)$ of mass m is:

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi\, \partial^\mu \phi - m^2 \phi^2\right) . \tag{1.1}$$

The canonical momentum conjugate to ϕ is:

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)} = \dot{\phi}(x) . \tag{1.2}$$

The Klein–Gordon equation

The equation of motion follows from the Euler–Lagrange equations:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0 . \tag{1.3}$$

Box 1.1 Klein–Gordon equation from the Lagrangian

Starting from $\mathcal{L} = \frac{1}{2}(\partial_\mu \phi\, \partial^\mu \phi - m^2 \phi^2)$, we compute:

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \partial^\mu \phi , \tag{1.4}$$

$$\frac{\partial \mathcal{L}}{\partial \phi} = -m^2 \phi . \tag{1.5}$$

Substituting into the Euler–Lagrange equation:

$$\partial_\mu(\partial^\mu \phi) + m^2\phi = 0\,. \tag{1.6}$$

This is the **Klein–Gordon equation**:

$$\boxed{(\Box + m^2)\phi = 0} \tag{1.7}$$

where $\Box = \partial_\mu\partial^\mu = \partial_t^2 - \nabla^2$ is the d'Alembertian operator.

The Klein–Gordon equation is a relativistic wave equation: it is second-order in both time and space derivatives, ensuring Lorentz covariance. Each mode of the field oscillates as $e^{\pm i(E_p t - \vec{p}\cdot\vec{x})}$ with the relativistic dispersion relation $E_p = \sqrt{|\vec{p}|^2 + m^2}$.

The Hamiltonian

The Hamiltonian density is obtained by the Legendre transformation:

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L} = \frac{1}{2}\left(\pi^2 + (\nabla\phi)^2 + m^2\phi^2\right)\,. \tag{1.8}$$

Box 1.2 Hamiltonian density via Legendre transformation

The Legendre transformation relates the Lagrangian formulation (in terms of ϕ and $\dot{\phi}$) to the Hamiltonian formulation (in terms of ϕ and π). The Hamiltonian density is defined as:

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L}\,, \tag{1.9}$$

where $\dot{\phi}$ must be expressed in terms of π.
Step 1: From the definition of the conjugate momentum, $\pi = \partial\mathcal{L}/\partial\dot{\phi} = \dot{\phi}$, we immediately have:

$$\dot{\phi} = \pi\,. \tag{1.10}$$

Step 2: Substitute into the Lagrangian density. Recall:

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\phi\,\partial^\mu\phi - m^2\phi^2\right) = \frac{1}{2}\left(\dot{\phi}^2 - (\nabla\phi)^2 - m^2\phi^2\right)\,. \tag{1.11}$$

Step 3: Compute the Hamiltonian density:

$$\begin{aligned}
\mathcal{H} &= \pi\dot{\phi} - \mathcal{L} \\
&= \pi \cdot \pi - \frac{1}{2}\left(\pi^2 - (\nabla\phi)^2 - m^2\phi^2\right) \\
&= \pi^2 - \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2 \\
&= \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2 .
\end{aligned} \tag{1.12}$$

The three terms have clear physical interpretations:

- $\frac{1}{2}\pi^2$: kinetic energy density (associated with time evolution)
- $\frac{1}{2}(\nabla\phi)^2$: gradient energy density (cost of spatial variations)
- $\frac{1}{2}m^2\phi^2$: mass (potential) energy density

The total Hamiltonian is $H = \int d^3x\,\mathcal{H}$.

Mode expansion

The field operator $\phi(x)$ and its conjugate momentum $\pi(x)$ can be decomposed as:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2E_p}} \left[a(\vec{p})\, e^{-ip\cdot x} + a^\dagger(\vec{p})\, e^{ip\cdot x}\right] \tag{1.13}$$

$$\pi(x) = \dot{\phi}(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2E_p}} (-iE_p) \left[a(\vec{p})\, e^{-ip\cdot x} - a^\dagger(\vec{p})\, e^{ip\cdot x}\right], \tag{1.14}$$

where $E_p = \sqrt{|\vec{p}|^2 + m^2}$ is the on-shell energy, and $p \cdot x = E_p t - \vec{p}\cdot\vec{x}$.

Normalization convention: We use the Bjorken–Drell (BD) normalization throughout these notes. A detailed comparison with other conventions appears below.

Commutation relations

Physical motivation

The transition from classical to quantum field theory requires promoting the classical fields to operators satisfying appropriate commutation relations. The choice of commutation relations is not arbitrary—it is dictated by fundamental physical requirements:

1. **Causality:** Field operators at spacelike-separated points must commute, ensuring that measurements at one point cannot instantaneously affect measurements

at another. This is automatically satisfied by the equal-time commutators below, which imply $[\phi(x), \phi(y)] = 0$ for $(x-y)^2 < 0$.

2. **Correspondence principle:** In the limit $\hbar \to 0$, the commutators reduce to Poisson brackets multiplied by $i\hbar$. The canonical commutation relations mirror the classical Poisson bracket $\{q, p\} = 1$.
3. **Positive-definite Hilbert space:** For bosonic fields (integer spin), commutation relations lead to the correct bosonic Fock-space structure with positive-definite norm. Using anticommutation relations would instead produce the wrong statistics and occupancy structure for bosons.
4. **Lorentz invariance:** The commutation relations must be compatible with Lorentz transformations, which is ensured by working with covariant fields.

Equal-time canonical commutation relations

The equal-time canonical commutation relations are:

$$[\phi(t,\vec{x}), \pi(t,\vec{x}')] = i\delta^{(3)}(\vec{x}-\vec{x}') \tag{1.15}$$
$$[\phi(t,\vec{x}), \phi(t,\vec{x}')] = [\pi(t,\vec{x}), \pi(t,\vec{x}')] = 0\,. \tag{1.16}$$

These are the field-theoretic analogue of the quantum mechanical commutator $[\hat{q}, \hat{p}] = i\hbar$ (with $\hbar = 1$), extended to a continuous infinity of degrees of freedom labeled by $\vec{x}$.

Let us verify that these commutation relations satisfy the four physical requirements:

1. **Causality:** The equal-time commutators

$$[\phi(t,\vec{x}), \phi(t,\vec{x}')] = 0 \quad \text{and} \quad [\pi(t,\vec{x}), \pi(t,\vec{x}')] = 0 \tag{1.17}$$

trivially vanish. More importantly, one can show that the full (unequal-time) commutator is

$$[\phi(x), \phi(y)] = i\Delta(x-y)\,, \tag{1.18}$$

where $\Delta(x-y)$ is the **Pauli–Jordan function** (see box below). This function vanishes for spacelike separations $(x-y)^2 < 0$, ensuring that field measurements at causally disconnected points do not interfere.

2. **Correspondence principle:** The classical Poisson bracket for fields is

$$\{\phi(t,\vec{x}), \pi(t,\vec{x}')\}_{\text{PB}} = \delta^{(3)}(\vec{x}-\vec{x}')\,. \tag{1.19}$$

The quantization rule $\{A, B\}_{\text{PB}} \to \frac{1}{i\hbar}[A, B]$ gives precisely $[\phi, \pi] = i\hbar\,\delta^{(3)}(\vec{x} - \vec{x}')$, which reduces to our expression with $\hbar = 1$.

3. **Positive-definite Hilbert space:** The commutation relation $[a, a^\dagger] = 1$ (for a single mode) implies that

$$\|a^\dagger |n\rangle\|^2 = \langle n| aa^\dagger |n\rangle = \langle n| (a^\dagger a + 1) |n\rangle = n + 1 > 0 \,. \tag{1.20}$$

All states $|n\rangle = (a^\dagger)^n |0\rangle / \sqrt{n!}$ have positive norm. Had we used anticommutators, we would find $(a^\dagger)^2 = 0$, and attempting to construct a Fock space leads to inconsistencies or negative norms.

4. **Lorentz invariance:** While the equal-time commutators are not manifestly covariant (they single out a time slice), the full theory is Lorentz invariant. The commutator $[\phi(x), \phi(y)]$ depends only on the invariant interval $(x-y)^2$, and the S-matrix constructed from these fields transforms correctly under the Poincaré group.

Box 1.3 The Pauli–Jordan function

The **Pauli–Jordan function** (also called the commutator function $\Delta(x-y)$) is defined as:

$$\begin{aligned} i\Delta(x) = [\phi(x), \phi(0)] &= \int \frac{d^3p}{(2\pi)^3 2E_p} \left(e^{-ip\cdot x} - e^{ip\cdot x}\right) \\ &= -2i \int \frac{d^3p}{(2\pi)^3 2E_p} \sin(p\cdot x) \,. \end{aligned} \tag{1.21}$$

Equivalently, in terms of the positive and negative frequency parts:

$$\Delta(x) = \Delta_+(x) - \Delta_-(x) \,, \tag{1.22}$$

where $i\Delta_+(x) = \langle 0| \phi(x)\phi(0) |0\rangle$ and $i\Delta_-(x) = \langle 0| \phi(0)\phi(x) |0\rangle$.
Key properties:

1. *Antisymmetry:* $\Delta(-x) = -\Delta(x)$, which follows from $[\phi(x), \phi(y)] = -[\phi(y), \phi(x)]$.
2. *Lorentz invariance:* $\Delta(x)$ depends only on $x^2 = t^2 - |\vec{x}|^2$ and $\text{sign}(t)$, making it a Lorentz scalar (up to the sign of t).
3. *Causality (microcausality):* $\Delta(x) = 0$ for $x^2 < 0$ (spacelike separations). *Proof:* For spacelike x, we can choose a Lorentz frame where $t = 0$. Then:

$$\Delta(0, \vec{x}) = -2 \int \frac{d^3p}{(2\pi)^3 2E_p} \sin(-\vec{p}\cdot\vec{x}) = 2 \int \frac{d^3p}{(2\pi)^3 2E_p} \sin(\vec{p}\cdot\vec{x}) \,. \tag{1.23}$$

Under $\vec{p} \to -\vec{p}$, the measure d^3p and $E_p = \sqrt{|\vec{p}|^2 + m^2}$ are invariant, but $\sin(\vec{p}\cdot\vec{x}) \to -\sin(\vec{p}\cdot\vec{x})$. Therefore the integral equals minus itself, hence $\Delta(0, \vec{x}) = 0$. By Lorentz invariance, $\Delta(x) = 0$ for all spacelike x.

4. *Equal-time limit:* $\Delta(0,\vec{x}) = 0$ and $\partial_t \Delta(t,\vec{x})|_{t=0} = -\delta^{(3)}(\vec{x})$.
   Proof of the derivative: From the mode expansion,

$$\partial_t \Delta(t,\vec{x})\Big|_{t=0} = -2\int \frac{d^3p}{(2\pi)^3 2E_p} E_p \cos(\vec{p}\cdot\vec{x})\Big|_{t=0} = -\int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} = -\delta^{(3)}(\vec{x})\,. \tag{1.24}$$

5. *Klein–Gordon equation:* $(\Box + m^2)\Delta(x) = 0$ for $x \neq 0$.
   Proof: Each plane-wave mode $e^{\pm ip\cdot x}$ satisfies the Klein–Gordon equation when $p^0 = E_p$, so the integral does as well.

The Pauli–Jordan function is the key object ensuring **microscopic causality**: observables at spacelike-separated points commute, so no signal can propagate faster than light.

Creation and annihilation operator commutators

These are equivalent to the commutation relations for the creation and annihilation operators:

$$[a(\vec{p}), a^\dagger(\vec{p}')] = \delta^{(3)}(\vec{p}-\vec{p}') \tag{1.25}$$

$$[a(\vec{p}), a(\vec{p}')] = [a^\dagger(\vec{p}), a^\dagger(\vec{p}')] = 0\,. \tag{1.26}$$

Box 1.4 Commutation relations for a and $a^\dagger$ from canonical commutators

We derive the operator commutation relations by inverting the mode expansion. From Eq. (1.13), we can express the creation and annihilation operators in terms of the fields:

$$a(\vec{p}) = \int \frac{d^3x}{(2\pi)^{3/2}\sqrt{2E_p}} e^{ip\cdot x}\left(E_p\phi(x) + i\pi(x)\right)\Big|_{t=0}, \tag{1.27}$$

where we evaluate at $t = 0$ (the result is independent of time for free fields). This can be verified by substituting the mode expansion back.
To compute $[a(\vec{p}), a^\dagger(\vec{p}')]$, we use the linearity of the commutator:

$$[a(\vec{p}), a^\dagger(\vec{p}')] = \int \frac{d^3x\, d^3x'}{(2\pi)^3\sqrt{4E_pE_{p'}}} e^{i\vec{p}\cdot\vec{x} - i\vec{p}'\cdot\vec{x}'} \times \left[(E_p\phi + i\pi)_{\vec{x}},\, (E_{p'}\phi - i\pi)_{\vec{x}'}\right]. \tag{1.28}$$

Expanding the commutator and using $[\phi,\phi]=[\pi,\pi]=0$:

$$= \int \frac{d^3x\, d^3x'}{(2\pi)^3\sqrt{4E_pE_{p'}}} e^{i\vec{p}\cdot\vec{x}-i\vec{p}'\cdot\vec{x}'} \times\left(-iE_{p'}[\phi_{\vec{x}},\pi_{\vec{x}'}]+iE_p[\pi_{\vec{x}},\phi_{\vec{x}'}]\right). \tag{1.29}$$

Using $[\phi(t,\vec{x}),\pi(t,\vec{x}')]=i\delta^{(3)}(\vec{x}-\vec{x}')$:

$$\begin{aligned} &= \int \frac{d^3x\, d^3x'}{(2\pi)^3\sqrt{4E_pE_{p'}}} e^{i\vec{p}\cdot\vec{x}-i\vec{p}'\cdot\vec{x}'}\,(E_{p'}+E_p)\,\delta^{(3)}(\vec{x}-\vec{x}') \\ &= \int \frac{d^3x}{(2\pi)^3}\,\frac{E_p+E_{p'}}{\sqrt{4E_pE_{p'}}}\, e^{i(\vec{p}-\vec{p}')\cdot\vec{x}}. \end{aligned} \tag{1.30}$$

The integral gives $(2\pi)^3\delta^{(3)}(\vec{p}-\vec{p}')$, and the delta function sets $E_{p'}=E_p$:

$$[a(\vec{p}),a^\dagger(\vec{p}')] = \frac{2E_p}{2E_p}\delta^{(3)}(\vec{p}-\vec{p}') = \delta^{(3)}(\vec{p}-\vec{p}'). \tag{1.31}$$

A similar calculation shows $[a(\vec{p}),a(\vec{p}')]=0$: the relevant terms give $(E_{p'}-E_p)\delta^{(3)}(\vec{p}-\vec{p}')=0$.

The Hamiltonian and normal ordering

Substituting the mode expansion into the Hamiltonian gives:

$$H = \frac{1}{2}\int d^3p\, E_p\left(a^\dagger(\vec{p})a(\vec{p})+a(\vec{p})a^\dagger(\vec{p})\right). \tag{1.32}$$

Using the commutation relation $[a(\vec{p}),a^\dagger(\vec{p}')]=\delta^{(3)}(\vec{p}-\vec{p}')$:

$$H = \int d^3p\, E_p\, a^\dagger(\vec{p})a(\vec{p}) + \frac{1}{2}\delta^{(3)}(\vec{0})\int d^3p\, E_p. \tag{1.33}$$

The second term is the (infinite) zero-point energy. By **normal ordering** (denoted by colons), we define:

$$:H: = \int d^3p\, E_p\, a^\dagger(\vec{p})a(\vec{p}). \tag{1.34}$$

Box 1.5 Hamiltonian in terms of creation and annihilation operators

Starting from the Hamiltonian density $\mathcal{H} = \frac{1}{2}(\pi^2 + (\nabla\phi)^2 + m^2\phi^2)$, we compute $H = \int d^3x\, \mathcal{H}$ by substituting the mode expansions.

Step 1: Mode expansions at $t = 0$. For convenience, we work at $t = 0$:

$$\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2E_p}} \left[a(\vec{p})\, e^{i\vec{p}\cdot\vec{x}} + a^\dagger(\vec{p})\, e^{-i\vec{p}\cdot\vec{x}} \right], \tag{1.35}$$

$$\pi(\vec{x}) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2E_p}} (-iE_p) \left[a(\vec{p})\, e^{i\vec{p}\cdot\vec{x}} - a^\dagger(\vec{p})\, e^{-i\vec{p}\cdot\vec{x}} \right]. \tag{1.36}$$

Step 2: Compute $\int d^3x\, \pi^2$. Squaring π and integrating:

$$\int d^3x\, \pi^2 = \int \frac{d^3p\, d^3q}{(2\pi)^3\sqrt{4E_pE_q}} (-E_pE_q) \times \left[a_p - a^\dagger_{-p}\right]\left[a_q - a^\dagger_{-q}\right] \int d^3x\, e^{i(\vec{p}+\vec{q})\cdot\vec{x}}, \tag{1.37}$$

where we use the shorthand $a_p \equiv a(\vec{p})$, $a_{-p} \equiv a(-\vec{p})$, etc.
The spatial integral gives $(2\pi)^3\delta^{(3)}(\vec{p}+\vec{q})$, which sets $\vec{q} = -\vec{p}$ (hence $E_q = E_p$):

$$\int d^3x\, \pi^2 = -\int d^3p\, \frac{E_p}{2} \left[a_p - a^\dagger_{-p}\right]\left[a_{-p} - a^\dagger_p\right]. \tag{1.38}$$

Expanding:

$$= \int d^3p\, \frac{E_p}{2} \left[-a_pa_{-p} + a_pa^\dagger_p + a^\dagger_{-p}a_{-p} - a^\dagger_{-p}a^\dagger_p\right]. \tag{1.39}$$

Step 3: Compute $\int d^3x\, [(\nabla\phi)^2 + m^2\phi^2]$. A similar calculation gives:

$$\int d^3x\, (\nabla\phi)^2 = \int d^3p\, \frac{|\vec{p}|^2}{2E_p} \left[a_pa_{-p} + a^\dagger_{-p}a^\dagger_p + a_pa^\dagger_p + a^\dagger_{-p}a_{-p}\right], \tag{1.40}$$

$$\int d^3x\, m^2\phi^2 = \int d^3p\, \frac{m^2}{2E_p} \left[a_pa_{-p} + a^\dagger_{-p}a^\dagger_p + a_pa^\dagger_p + a^\dagger_{-p}a_{-p}\right]. \tag{1.41}$$

Step 4: Combine all terms. Adding the gradient and mass terms:

$$\int d^3x \left[(\nabla\phi)^2 + m^2\phi^2\right] = \int d^3p\, \frac{m^2 + |\vec{p}|^2}{2E_p} \left[a_pa_{-p} + a^\dagger_{-p}a^\dagger_p\right] + \int d^3p\, \frac{m^2 + |\vec{p}|^2}{2E_p} \left[a_pa^\dagger_p + a^\dagger_{-p}a_{-p}\right]. \tag{1.42}$$

Using $E_p^2 = |\vec{p}|^2 + m^2$, both coefficients are equal to E_p.
Adding the π^2 term (which contributes $-E_p/2$ to the aa and $a^\dagger a^\dagger$ terms, and $+E_p/2$ to the $aa^\dagger$ and $a^\dagger a$ terms):

$$H = \frac{1}{2}\int d^3p \left[\left(-E_p + \frac{m^2 + |\vec{p}|^2}{E_p}\right)\left(a_p a_{-p} + a_{-p}^\dagger a_p^\dagger\right) + E_p\left(a_p a_p^\dagger + a_p^\dagger a_p\right)\right]. \tag{1.43}$$

Step 5: The aa and $a^\dagger a^\dagger$ terms cancel. Using $E_p^2 = |\vec{p}|^2 + m^2$, the coefficient of the pair-creation/annihilation terms is

$$\frac{1}{2}\left(-E_p + \frac{m^2 + |\vec{p}|^2}{E_p}\right) = \frac{-E_p^2 + m^2 + |\vec{p}|^2}{2E_p} = 0. \tag{1.44}$$

Therefore the terms proportional to $a_p a_{-p}$ and $a_{-p}^\dagger a_p^\dagger$ cancel *algebraically*, before normal ordering. Only the number-operator structure remains.
Final result:

$$H = \frac{1}{2}\int d^3p\, E_p \left(a^\dagger(\vec{p})a(\vec{p}) + a(\vec{p})a^\dagger(\vec{p})\right). \tag{1.45}$$

Using the commutation relation to normal order:

$$a(\vec{p})a^\dagger(\vec{p}) = a^\dagger(\vec{p})a(\vec{p}) + \delta^{(3)}(\vec{0}), \tag{1.46}$$

we obtain

$$H = \int d^3p\, E_p\, a^\dagger(\vec{p})a(\vec{p}) + \underbrace{\frac{1}{2}\int d^3p\, E_p\, \delta^{(3)}(\vec{0})}_{\text{vacuum energy } E_0}. \tag{1.47}$$

The vacuum energy E_0 is infinite (both from the $\delta^{(3)}(\vec{0})$ factor and the unbounded $\int d^3p\, E_p$), but is unobservable and removed by normal ordering.

Lorentz-invariant normalization: The simple choice $|p\rangle = a^\dagger(\vec{p})\,|0\rangle$ leads to $\langle p|q\rangle = \delta^{(3)}(\vec{p} - \vec{q})$, which is not Lorentz invariant. A Lorentz-invariant normalization is $|p\rangle \propto \sqrt{2E_p}\, a^\dagger(\vec{p})\,|0\rangle$.

Four-momentum operator

From Noether's theorem applied to spacetime translations, the conserved four-momentum is (see Section 1.6):

$$P^\mu = \int d^3x\, T^{0\mu}\,, \tag{1.48}$$

where $T^{\mu\nu}$ is the canonical energy-momentum tensor. For the real scalar field, the spatial components are:

$$P^i = \int d^3x\, \pi(\vec{x},t)\, \partial^i\phi(\vec{x},t) = -\int d^3x\, \pi(\vec{x},t)\, \partial_i\phi(\vec{x},t)\,, \tag{1.49}$$

where $\partial^i = -\partial_i$ in the mostly-minus signature $(+,-,-,-)$.

Substituting the mode expansion, one obtains the normal-ordered result:

$$\vec{P} = \int d^3p\, \vec{p}\, a^\dagger(\vec{p})\, a(\vec{p})\,, \tag{1.50}$$

confirming the physical interpretation: each quantum of momentum $\vec{p}$ contributes $\vec{p}$ to the total momentum.

P^μ as the generator of spacetime translations. The four-momentum operator generates translations:

$$[\phi(x), P^\mu] = +i\partial^\mu\phi(x)\,. \tag{1.51}$$

This means that under a finite translation $x \to x+a$, the field transforms as $\phi(x) \to e^{ia\cdot P}\phi(x)\, e^{-ia\cdot P} = \phi(x+a)$.[1]

Box 1.6 Proof that $[\phi(y), P^i] = +i\partial^i\phi(y)$

We prove the spatial component using only the equal-time canonical commutation relations (1.15).

[1]The exponentiation from the infinitesimal relation to the finite transformation follows from the **Hadamard lemma** (also known as the Baker–Campbell–Hausdorff formula for conjugation):

$$e^A B\, e^{-A} = B + [A,B] + \tfrac{1}{2!}[A,[A,B]] + \tfrac{1}{3!}[A,[A,[A,B]]] + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}\, \mathrm{ad}_A^n(B)\,,$$

where $\mathrm{ad}_A^n(B) \equiv [A, \mathrm{ad}_A^{n-1}(B)]$ with $\mathrm{ad}_A^0(B) = B$. This identity holds as a formal power series for *any* operators A and B; convergence is guaranteed whenever A is bounded or, more relevantly for QFT, when $A = ia_\mu P^\mu$ is anti-Hermitian (i.e. P^μ is Hermitian), so that e^A is unitary and the conjugation $e^A B\, e^{-A}$ is a well-defined automorphism of the operator algebra. Setting $A = ia_\mu P^\mu$ and using $[ia_\mu P^\mu, \phi] = -ia_\mu[\phi, P^\mu] = -ia_\mu(+i\partial^\mu\phi) = a_\mu\partial^\mu\phi$, the nested commutators reproduce the Taylor expansion: $e^{ia\cdot P}\phi(x)\, e^{-ia\cdot P} = \sum_{n=0}^{\infty}\frac{1}{n!}(a\cdot\partial)^n\phi(x) = \phi(x+a)$.

Starting from Eq. (1.49):

$$[\phi(\vec{y},t),\, P^i] = -\int d^3x\, [\phi(\vec{y},t),\, \pi(\vec{x},t)\, \partial_i\phi(\vec{x},t)]\,. \tag{1.52}$$

Step 1: Apply the identity $[A,BC] = [A,B]C + B[A,C]$:

$$[\phi(\vec{y},t),\, \pi(\vec{x},t)\, \partial_i\phi(\vec{x},t)] = \underbrace{[\phi(\vec{y},t),\, \pi(\vec{x},t)]}_{=i\delta^{(3)}(\vec{y}-\vec{x})}\, \partial_i\phi(\vec{x},t) + \pi(\vec{x},t)\, \underbrace{[\phi(\vec{y},t),\, \partial_i\phi(\vec{x},t)]}_{=0}\,. \tag{1.53}$$

The second term vanishes because $\partial_i^{(x)}$ acts on $\vec{x}$ and can be pulled out of the commutator: $[\phi(\vec{y},t), \partial_i\phi(\vec{x},t)] = \partial_i^{(x)}[\phi(\vec{y},t), \phi(\vec{x},t)] = 0$.
Step 2: Substituting:

$$[\phi(\vec{y},t),\, P^i] = -\int d^3x\, i\delta^{(3)}(\vec{y}-\vec{x})\, \partial_i\phi(\vec{x},t) = -i\,\partial_i\phi(\vec{y},t) = +i\partial^i\phi(\vec{y},t)\,, \tag{1.54}$$

where in the last step we used $\partial_i = -\partial^i$ for spatial indices with metric $(+,-,-,-)$.
□

Extension to composite operators. The relation (1.51) extends to *any* local operator built from ϕ and its derivatives. In particular, P^μ generates translations of $\partial_\nu\phi$ as well:

$$[\partial_\nu\phi(x),\, P^\mu] = +i\partial^\mu\partial_\nu\phi(x)\,. \tag{1.55}$$

This follows directly by taking $\partial_\nu^{(x)}$ of both sides of Eq. (1.51): since P^μ is an integral over all space and does not depend on x, it commutes with $\partial_\nu^{(x)}$, giving

$$\partial_\nu^{(x)}[\phi(x), P^\mu] = [\partial_\nu\phi(x), P^\mu] = +i\partial_\nu\partial^\mu\phi(x)\,. \tag{1.56}$$

More generally, for any local operator $\mathcal{O}(x)$ constructed from products of fields and their derivatives at the point x, one can show by repeated application of the identity $[AB,C] = A[B,C] + [A,C]B$ that

$$[\mathcal{O}(x),\, P^\mu] = +i\partial^\mu\mathcal{O}(x)\,. \tag{1.57}$$

This is the statement that P^μ is the generator of spacetime translations for the entire operator algebra, not just for the fundamental field ϕ.

Normalization conventions: a detailed comparison

Different textbooks use different normalization conventions for the field expansion. The choice affects the commutation relations, state normalization, and the form of the completeness relation. We compare three common conventions.

Convention I: Bjorken–Drell (used in these notes)

Field expansion:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2E_p}} \left[a(\vec{p})\, e^{-ip\cdot x} + a^\dagger(\vec{p})\, e^{ip\cdot x}\right] . \tag{1.58}$$

Commutation relations:

$$[a(\vec{p}), a^\dagger(\vec{p}')] = \delta^{(3)}(\vec{p}-\vec{p}') . \tag{1.59}$$

State normalization: $|\vec{p}\rangle = a^\dagger(\vec{p})\,|0\rangle$ gives

$$\langle\vec{p}|\vec{p}'\rangle = \delta^{(3)}(\vec{p}-\vec{p}') . \qquad \text{(not Lorentz invariant)} \tag{1.60}$$

Completeness relation:

$$\mathbb{1} = |0\rangle\langle 0| + \int d^3p\, |\vec{p}\rangle\langle\vec{p}| + \cdots . \tag{1.61}$$

Convention II: Covariant $(2E)$ normalization

Field expansion:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3 2E_p} \left[a(\vec{p})\, e^{-ip\cdot x} + a^\dagger(\vec{p})\, e^{ip\cdot x}\right] . \tag{1.62}$$

Commutation relations:

$$[a(\vec{p}), a^\dagger(\vec{p}')] = (2\pi)^3\, 2E_p\, \delta^{(3)}(\vec{p}-\vec{p}') . \tag{1.63}$$

State normalization: $|\vec{p}\rangle = a^\dagger(\vec{p})\,|0\rangle$ gives the **Lorentz-invariant** normalization

$$\langle\vec{p}|\vec{p}'\rangle = (2\pi)^3\, 2E_p\, \delta^{(3)}(\vec{p}-\vec{p}') . \tag{1.64}$$

This is Lorentz invariant because the one-particle phase-space measure $d^3p/(2E_p)$ is invariant under boosts, as we now show.

Box 1.7 **Lorentz invariance of $d^3p/(2E_p)$ and of $2E_p\,\delta^{(3)}$**

The measure $d^3p/(2E_p)$. Start from the manifestly Lorentz-invariant combination

$$\int \frac{d^3p}{2E_{\vec{p}}}\, f(\vec{p}) = \int d^4p\, \delta(p^2 - m^2)\, \theta(p^0)\, f(p)\,. \tag{1.65}$$

To verify the right-hand side, use

$$\delta(p^2 - m^2) = \delta\big((p^0)^2 - E_{\vec{p}}^2\big) = \frac{1}{2\,|p^0|}\Big[\delta(p^0 - E_{\vec{p}}) + \delta(p^0 + E_{\vec{p}})\Big], \tag{1.66}$$

and note that $\theta(p^0)$ selects only the positive-energy root, giving $\int dp^0\, \delta(p^2 - m^2)\, \theta(p^0) = 1/(2E_{\vec{p}})$.
Each factor on the right-hand side of (1.65) is Lorentz invariant:

- d^4p: because $|\det\Lambda| = 1$ for proper Lorentz transformations;
- $\delta(p^2 - m^2)$: because $p^2 - m^2$ is a Lorentz scalar, and the delta function of a scalar is a scalar;
- $\theta(p^0)$: because proper orthochronous transformations preserve the sign of p^0 for timelike vectors on the mass shell ($p^2 = m^2 > 0$).

Therefore $d^3p/(2E_{\vec{p}})$ is Lorentz invariant. □

Transformation of $\delta^{(3)}$ under boosts. By contrast, $\delta^{(3)}(\vec{p} - \vec{q})$ alone is *not* Lorentz invariant. Under a boost Λ, the three-momentum transforms with Jacobian

$$\left|\det \frac{\partial \vec{p}\,'}{\partial \vec{p}}\right| = \frac{E'}{E}\,, \tag{1.67}$$

so that $\delta^{(3)}(\vec{p} - \vec{q}) \to (E'/E)\, \delta^{(3)}(\vec{p}\,' - \vec{q}\,')$. The combination $2E_p\, \delta^{(3)}(\vec{p} - \vec{q})$ *is* invariant, because the factor E exactly compensates the Jacobian. This is why the covariant normalization $\langle \vec{p}|\vec{p}\,'\rangle = (2\pi)^3\, 2E_p\, \delta^{(3)}(\vec{p} - \vec{p}\,')$ transforms as a Lorentz scalar.

Completeness relation:

$$\mathbb{1} = |0\rangle\langle 0| + \int \frac{d^3p}{(2\pi)^3 2E_p}\, |\vec{p}\rangle\langle \vec{p}| + \cdots\,. \tag{1.68}$$

Convention III: Box normalization (Mandl–Shaw)

In a finite volume $V = L^3$ with periodic boundary conditions, momenta are discrete: $\vec{p} = (2\pi/L)(n_x, n_y, n_z)$.

Field expansion:

$$\phi(x) = \sum_{\vec{p}} \frac{1}{\sqrt{2VE_p}} \left[a(\vec{p})\, e^{-ip\cdot x} + a^\dagger(\vec{p})\, e^{ip\cdot x} \right] . \tag{1.69}$$

Commutation relations:

$$[a(\vec{p}), a^\dagger(\vec{p}\,')] = \delta_{\vec{p},\vec{p}\,'} . \tag{1.70}$$

State normalization:

$$\langle \vec{p}|\vec{p}\,'\rangle = \delta_{\vec{p},\vec{p}\,'} . \qquad \text{(dimensionless)} \tag{1.71}$$

The continuum limit is obtained via

$$\sum_{\vec{p}} \to V \int \frac{d^3p}{(2\pi)^3}, \qquad \delta_{\vec{p},\vec{p}\,'} \to \frac{(2\pi)^3}{V} \delta^{(3)}(\vec{p}-\vec{p}\,') .$$

The completeness relation: definition or consequence?

A natural question arises: is the completeness relation an independent axiom, or does it follow from the other assumptions? The answer is that **the completeness relation is a consequence** of the commutation relations and the definition of the Fock space, not an additional postulate.

Box 1.8 Derivation of the completeness relation

The completeness relation in the one-particle sector follows from the commutation relations and the definition of states.

Setup: We have operators satisfying $[a(\vec{p}), a^\dagger(\vec{p}\,')] = f(\vec{p})\,\delta^{(3)}(\vec{p}-\vec{p}\,')$ for some function $f(\vec{p})$ (which depends on the convention), and one-particle states $|\vec{p}\rangle = a^\dagger(\vec{p})\,|0\rangle$.

Goal: Show that $\int d\mu(\vec{p})\,|\vec{p}\rangle\langle\vec{p}| = \mathbb{1}_{\text{1-particle}}$ for an appropriate measure $d\mu(\vec{p})$.

Derivation: Consider the action of the proposed completeness operator on an arbitrary one-particle state $|\vec{q}\rangle$:

$$\left(\int d\mu(\vec{p})\,|\vec{p}\rangle\langle\vec{p}|\right)|\vec{q}\rangle = \int d\mu(\vec{p})\,|\vec{p}\rangle\langle\vec{p}|\vec{q}\rangle . \tag{1.72}$$

The inner product is $\langle\vec{p}|\vec{q}\rangle = \langle 0|\,a(\vec{p})a^\dagger(\vec{q})\,|0\rangle = f(\vec{p})\,\delta^{(3)}(\vec{p}-\vec{q})$, so:

$$= \int d\mu(\vec{p})\,|\vec{p}\rangle\, f(\vec{p})\,\delta^{(3)}(\vec{p}-\vec{q}) = f(\vec{q})\,d\mu(\vec{q})\,|\vec{q}\rangle . \tag{1.73}$$

For this to equal $|\vec{q}\rangle$, we need $f(\vec{q})\, d\mu(\vec{q}) = 1$, i.e., the measure must be:

$$d\mu(\vec{p}) = \frac{d^3p}{f(\vec{p})} . \tag{1.74}$$

Application to conventions:

- *BD convention:* $f(\vec{p}) = 1$, so $d\mu = d^3p$, giving $\mathbb{1} = \int d^3p\, |\vec{p}\rangle\langle\vec{p}|$.
- *Covariant convention:* $f(\vec{p}) = (2\pi)^3 2E_p$, so $d\mu = d^3p/[(2\pi)^3 2E_p]$, giving $\mathbb{1} = \int \frac{d^3p}{(2\pi)^3 2E_p} |\vec{p}\rangle\langle\vec{p}|$.

Conclusion: The completeness relation is uniquely determined by the commutation relations. Once you fix the normalization factor $\mathcal{N}(\vec{p})$ in the field expansion (which determines the commutator), both the state normalization $\langle\vec{p}|\vec{q}\rangle$ and the measure in the completeness relation are fixed. There is no freedom to choose them independently.

Conversion between conventions

Let a_{BD} denote the annihilation operator in Convention I, and let a_{PS} denote the Peskin–Schroeder continuum convention, where the same mode expansion is written with measure $d^3p/(2\pi)^3$ and coefficient $1/\sqrt{2E_p}$. The key relations between operators in different conventions are:

$$a_{\mathrm{PS}}(\vec{p}) = (2\pi)^{3/2} a_{\mathrm{BD}}(\vec{p}) , \tag{1.75}$$

$$a_{\mathrm{cov}}(\vec{p}) = (2\pi)^{3/2} \sqrt{2E_p}\, a_{\mathrm{BD}}(\vec{p}) , \tag{1.76}$$

$$a_{\mathrm{MS}}(\vec{p}) = \sqrt{\frac{V}{(2\pi)^3}}\, a_{\mathrm{PS}}(\vec{p}) = \sqrt{\frac{V}{(2\pi)^3\, 2E_p}}\, a_{\mathrm{cov}}(\vec{p}) . \tag{1.77}$$

Physical observables are convention-independent. When computing S-matrix elements and cross sections, different conventions lead to different intermediate expressions but the same final results for measurable quantities.

Advantages and drawbacks of each convention

While all conventions yield identical physical predictions, they differ in practical convenience depending on the context:

Convention I: Bjorken–Drell

+ Simple commutator: $[a, a^\dagger] = \delta^{(3)}(\vec{p} - \vec{p}\,')$ with no extra factors.
+ Direct analogy with non-relativistic quantum mechanics (harmonic oscillator algebra).

+ State counting is straightforward: $\langle\vec{p}|\vec{p}'\rangle = \delta^{(3)}(\vec{p}-\vec{p}')$.
– State normalization is *not* Lorentz invariant; must track transformation factors under boosts.
– The field expansion has an awkward $(2\pi)^{-3/2}$ factor.

Convention II: Covariant $(2E)$ normalization

+ Lorentz-invariant state normalization: $\langle\vec{p}|\vec{p}'\rangle = (2\pi)^3 2E_p\,\delta^{(3)}(\vec{p}-\vec{p}')$.
+ The completeness measure $d^3p/[(2\pi)^3 2E_p]$ is the Lorentz-invariant phase space.
+ Preferred in modern QFT textbooks (Peskin & Schroeder, Schwartz, Weinberg).
+ Cross-section formulas have manifestly covariant form.
– Commutator has energy-dependent factors: $[a, a^\dagger] = (2\pi)^3 2E_p\,\delta^{(3)}$.
– Less intuitive for beginners; the $2E_p$ factors can obscure the physics.

Convention III: Box normalization (Mandl–Shaw)

+ All quantities are dimensionless: $[a, a^\dagger] = \delta_{\vec{p},\vec{p}'}$, $\langle\vec{p}|\vec{p}'\rangle = \delta_{\vec{p},\vec{p}'}$.
+ No delta-function singularities; sums replace integrals.
+ Conceptually cleaner for counting states (e.g., in statistical mechanics).
+ Useful for lattice QFT and finite-volume calculations.
– Requires taking $V \to \infty$ limit at the end of calculations.
– Translation invariance is only approximate (recovered in the infinite-volume limit).
– Lorentz invariance is obscured by the finite box.

Recommendation: For learning QFT, the BD convention (used in these notes) offers the most direct connection to familiar quantum mechanics. For advanced calculations, especially those involving Lorentz transformations or crossing symmetry, the covariant convention is more natural. Box normalization is valuable for conceptual clarity and numerical work.

Consistency between field expansion, (anti)commutators, and state normalization

The normalization convention is not an arbitrary cosmetic choice: the factors in the field expansion, the (anti)commutation relations, and the normalization of one-particle states are tightly linked.

Box 1.9 Deriving the commutator from the field expansion

Consider a general mode expansion for a real scalar field:

$$\phi(x) = \int d^3p\, \mathcal{N}(\vec{p}) \left[a(\vec{p})\, e^{-ip\cdot x} + a^\dagger(\vec{p})\, e^{ip\cdot x} \right], \tag{1.78}$$

with normalization factor $\mathcal{N}(\vec{p})$ (assumed real for simplicity). The conjugate momentum is:

$$\pi(x) = \dot{\phi}(x) = \int d^3p\, \mathcal{N}(\vec{p})\, (-iE_p) \left[a(\vec{p})\, e^{-ip\cdot x} - a^\dagger(\vec{p})\, e^{ip\cdot x} \right]. \tag{1.79}$$

Goal: Determine $[a(\vec{p}), a^\dagger(\vec{q})]$ such that $[\phi(t,\vec{x}), \pi(t,\vec{y})] = i\delta^{(3)}(\vec{x}-\vec{y})$. Assume $[a(\vec{p}), a^\dagger(\vec{q})] = f(\vec{p})\,\delta^{(3)}(\vec{p}-\vec{q})$ for some function $f(\vec{p})$ to be determined. At equal times ($t = 0$):

$$\begin{aligned} [\phi(\vec{x}), \pi(\vec{y})] = \int & d^3p\, d^3q\, \mathcal{N}(\vec{p})\mathcal{N}(\vec{q})(-iE_q) \\ & \times \left[\left(a_p e^{i\vec{p}\cdot\vec{x}} + a_p^\dagger e^{-i\vec{p}\cdot\vec{x}}\right), \left(a_q e^{i\vec{q}\cdot\vec{y}} - a_q^\dagger e^{-i\vec{q}\cdot\vec{y}}\right) \right]. \end{aligned} \tag{1.80}$$

The only non-vanishing commutators are $[a_p, a_q^\dagger]$ and $[a_p^\dagger, a_q] = -[a_q, a_p^\dagger]$:

$$\begin{aligned} = \int & d^3p\, d^3q\, \mathcal{N}(\vec{p})\mathcal{N}(\vec{q})(-iE_q)\, f(\vec{p})\,\delta^{(3)}(\vec{p}-\vec{q}) \\ & \times \left[-e^{i\vec{p}\cdot\vec{x}} e^{-i\vec{q}\cdot\vec{y}} - e^{-i\vec{p}\cdot\vec{x}} e^{i\vec{q}\cdot\vec{y}} \right]. \end{aligned} \tag{1.81}$$

Integrating over $\vec{q}$ (setting $\vec{q} = \vec{p}$):

$$= \int d^3p\, \mathcal{N}(\vec{p})^2 E_p\, f(\vec{p})\, (i) \cdot 2\cos\big(\vec{p}\cdot(\vec{x}-\vec{y})\big). \tag{1.82}$$

Since the integrand is even in $\vec{p}$, we can write $2\cos(\vec{p}\cdot\vec{z}) = e^{i\vec{p}\cdot\vec{z}} + e^{-i\vec{p}\cdot\vec{z}} \to 2e^{i\vec{p}\cdot\vec{z}}$ under integration:

$$[\phi(\vec{x}), \pi(\vec{y})] = 2i \int d^3p\, \mathcal{N}(\vec{p})^2 E_p\, f(\vec{p})\, e^{i\vec{p}\cdot(\vec{x}-\vec{y})}. \tag{1.83}$$

Comparing with $i\delta^{(3)}(\vec{x}-\vec{y}) = i\int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{x}-\vec{y})}$, we require:

$$\boxed{f(\vec{p}) = \frac{1}{2(2\pi)^3 E_p\, \mathcal{N}(\vec{p})^2}} \tag{1.84}$$

The one-particle state $|\vec{p}\rangle \equiv a^\dagger(\vec{p})|0\rangle$ then satisfies

$$\langle \vec{p}\,'|\vec{p}\rangle = f(\vec{p})\,\delta^{(3)}(\vec{p}-\vec{p}\,') = \frac{1}{2(2\pi)^3 E_p \mathcal{N}(\vec{p})^2}\,\delta^{(3)}(\vec{p}-\vec{p}\,')\,. \tag{1.85}$$

Thus, once $\mathcal{N}(\vec{p})$ is chosen, the mode commutators and state normalization are uniquely determined.

BD choice. In the Bjorken–Drell convention,

$$\mathcal{N}_{\mathrm{BD}}(\vec{p}) = \frac{1}{(2\pi)^{3/2}\sqrt{2E_p}}\,, \qquad \mathcal{N}_{\mathrm{BD}}^2 = \frac{1}{(2\pi)^3 \cdot 2E_p}\,. \tag{1.86}$$

Substituting into (1.84):

$$f(\vec{p}) = \frac{1}{2(2\pi)^3 E_p \cdot \frac{1}{(2\pi)^3 \cdot 2E_p}} = \frac{(2\pi)^3 \cdot 2E_p}{2(2\pi)^3 E_p} = 1\,. \tag{1.87}$$

Therefore $[a(\vec{p}), a^\dagger(\vec{p}\,')] = \delta^{(3)}(\vec{p}-\vec{p}\,')$ and $\langle \vec{p}\,'|\vec{p}\rangle = \delta^{(3)}(\vec{p}-\vec{p}\,')$. This is the convention used throughout these notes.

Covariant choice. In the covariant convention,

$$\mathcal{N}_{\mathrm{cov}}(\vec{p}) = \frac{1}{(2\pi)^3 \cdot 2E_p}\,, \qquad \mathcal{N}_{\mathrm{cov}}^2 = \frac{1}{(2\pi)^6 \cdot 4E_p^2}\,. \tag{1.88}$$

Substituting into (1.84):

$$f(\vec{p}) = \frac{1}{2(2\pi)^3 E_p \cdot \frac{1}{(2\pi)^6 \cdot 4E_p^2}} = \frac{(2\pi)^6 \cdot 4E_p^2}{2(2\pi)^3 E_p} = (2\pi)^3 \cdot 2E_p\,. \tag{1.89}$$

Therefore $[a(\vec{p}), a^\dagger(\vec{p}\,')] = (2\pi)^3 2E_p\,\delta^{(3)}(\vec{p}-\vec{p}\,')$ and the state normalization is Lorentz-invariant:

$$\langle \vec{p}\,'|\vec{p}\rangle = (2\pi)^3 2E_p\,\delta^{(3)}(\vec{p}-\vec{p}\,')\,. \tag{1.90}$$

Fermions. The same logic applies to Dirac fields, with anticommutators replacing commutators and with a normalization choice for the spinors u, v consistent with the chosen mode normalization. In BD conventions, the factors $\sqrt{m/E_p}$ in the mode expansion, together with $\bar{u}u = \delta_{ss'}$, $\bar{v}v = -\delta_{ss'}$, ensure $\{b, b^\dagger\} = \delta^{(3)}$ and $\{d, d^\dagger\} = \delta^{(3)}$.

The key point is that the triple $(\mathcal{N}, [a, a^\dagger], \langle p'|p\rangle)$ is not independent: choosing any one of them fixes the others, and the same pattern holds for fermions with anticommutators.

1.2 The complex scalar field

A complex scalar field $\phi(x) = (\phi_1(x) + i\phi_2(x))/\sqrt{2}$ consists of two real scalar fields. The crucial difference from the real case is that it carries a conserved charge, implying the existence of distinct particles and antiparticles.

The Lagrangian for a complex field $\phi(x) \neq \phi^\dagger(x)$ is:

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi \,. \tag{1.91}$$

Note the absence of the factor $1/2$ compared to the real scalar case.

Equation of motion

Box 1.10 Klein–Gordon equation for the complex scalar field

For a complex field, we treat ϕ and $\phi^\dagger$ as independent variables. The Euler–Lagrange equation for $\phi^\dagger$ is:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\dagger)} \right) - \frac{\partial \mathcal{L}}{\partial \phi^\dagger} = 0 \,. \tag{1.92}$$

Computing the derivatives from $\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi$:

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\dagger)} = \partial^\mu \phi \,, \tag{1.93}$$

$$\frac{\partial \mathcal{L}}{\partial \phi^\dagger} = -m^2 \phi \,. \tag{1.94}$$

Substituting:

$$\partial_\mu(\partial^\mu \phi) + m^2 \phi = 0 \qquad \Longrightarrow \qquad \boxed{(\Box + m^2)\phi = 0} \tag{1.95}$$

Similarly, varying with respect to ϕ gives the equation for $\phi^\dagger$:

$$(\Box + m^2)\phi^\dagger = 0 \,. \tag{1.96}$$

Both ϕ and $\phi^\dagger$ satisfy the same Klein–Gordon equation, as expected for a free field theory.

Canonical momenta and Hamiltonian

The canonical momenta are:

$$\pi = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)} = \dot{\phi}^\dagger , \tag{1.97}$$

$$\pi^\dagger = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi^\dagger)} = \dot{\phi} . \tag{1.98}$$

Note that π is conjugate to ϕ (not to $\phi^\dagger$), and $\pi^\dagger$ is conjugate to $\phi^\dagger$.

Box 1.11 Hamiltonian density for the complex scalar field

The Hamiltonian density is obtained via the Legendre transformation:

$$\mathcal{H} = \pi \dot{\phi} + \pi^\dagger \dot{\phi}^\dagger - \mathcal{L} . \tag{1.99}$$

Step 1: Express time derivatives in terms of momenta. From the definitions:

$$\dot{\phi} = \pi^\dagger , \qquad \dot{\phi}^\dagger = \pi . \tag{1.100}$$

Step 2: Write the Lagrangian in terms of $\dot{\phi}$, $\dot{\phi}^\dagger$, and spatial derivatives:

$$\mathcal{L} = \dot{\phi}^\dagger \dot{\phi} - \nabla \phi^\dagger \cdot \nabla \phi - m^2 \phi^\dagger \phi . \tag{1.101}$$

Step 3: Compute the Hamiltonian density:

$$\begin{aligned}
\mathcal{H} &= \pi \dot{\phi} + \pi^\dagger \dot{\phi}^\dagger - \mathcal{L} \\
&= \dot{\phi}^\dagger \cdot \pi^\dagger + \dot{\phi} \cdot \pi - \left(\dot{\phi}^\dagger \dot{\phi} - \nabla \phi^\dagger \cdot \nabla \phi - m^2 \phi^\dagger \phi \right) \\
&= \pi \pi^\dagger + \pi^\dagger \pi - \pi \pi^\dagger + \nabla \phi^\dagger \cdot \nabla \phi + m^2 \phi^\dagger \phi \\
&= \pi^\dagger \pi + \nabla \phi^\dagger \cdot \nabla \phi + m^2 \phi^\dagger \phi .
\end{aligned} \tag{1.102}$$

The three terms represent:

- $\pi^\dagger \pi = \dot{\phi}^\dagger \dot{\phi}$: kinetic energy density
- $\nabla \phi^\dagger \cdot \nabla \phi$: gradient energy density
- $m^2 \phi^\dagger \phi$: mass energy density

The Hamiltonian density is:

$$\mathcal{H} = \pi^\dagger \pi + \nabla \phi^\dagger \cdot \nabla \phi + m^2 \phi^\dagger \phi . \tag{1.103}$$

Creation and annihilation operators

In terms of the operators for the two real fields:

$$a(\vec{k}) = \frac{1}{\sqrt{2}}\left(a_1(\vec{k}) + ia_2(\vec{k})\right), \tag{1.104}$$

$$b(\vec{k}) = \frac{1}{\sqrt{2}}\left(a_1(\vec{k}) - ia_2(\vec{k})\right). \tag{1.105}$$

Here $a(\vec{k})$ annihilates a particle (quantum of ϕ) and $b(\vec{k})$ annihilates an antiparticle (quantum of $\phi^\dagger$) when acting on kets.

The commutation relations are:

$$[a(\vec{k}), a^\dagger(\vec{k}')] = [b(\vec{k}), b^\dagger(\vec{k}')] = \delta^{(3)}(\vec{k} - \vec{k}'), \tag{1.106}$$

$$[a(\vec{k}), b(\vec{k}')] = [a(\vec{k}), b^\dagger(\vec{k}')] = 0. \tag{1.107}$$

All other commutators vanish.

Box 1.12 Normal-ordered Hamiltonian for the complex scalar field

The mode expansion for the complex scalar field is:

$$\phi(x) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2E_k}} \left[a(\vec{k})\, e^{-ik\cdot x} + b^\dagger(\vec{k})\, e^{ik\cdot x}\right], \tag{1.108}$$

$$\phi^\dagger(x) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2E_k}} \left[a^\dagger(\vec{k})\, e^{ik\cdot x} + b(\vec{k})\, e^{-ik\cdot x}\right]. \tag{1.109}$$

Note that ϕ contains a (annihilates particle) and $b^\dagger$ (creates antiparticle), while $\phi^\dagger$ contains $a^\dagger$ and b.

Substituting into $H = \int d^3x\,\mathcal{H}$ and following the same procedure as for the real scalar field, one finds:

$$H = \int d^3k\, E_k \left(a^\dagger(\vec{k})a(\vec{k}) + a(\vec{k})a^\dagger(\vec{k}) + b^\dagger(\vec{k})b(\vec{k}) + b(\vec{k})b^\dagger(\vec{k})\right)/2. \tag{1.110}$$

Using the commutation relations $[a, a^\dagger] = [b, b^\dagger] = \delta^{(3)}$:

$$H = \int d^3k\, E_k \left(a^\dagger(\vec{k})a(\vec{k}) + b^\dagger(\vec{k})b(\vec{k})\right) + \delta^{(3)}(\vec{0}) \int d^3k\, E_k. \tag{1.111}$$

The infinite vacuum energy is removed by normal ordering. The physical interpretation is clear:

- $a^\dagger a$ counts particles with energy E_k

- $b^\dagger b$ counts antiparticles with energy E_k
- Particles and antiparticles have the *same* mass and energy (as required by CPT symmetry)

The normal-ordered Hamiltonian is:

$$:H: = \int d^3k\, E_k \left(a^\dagger(\vec{k})a(\vec{k}) + b^\dagger(\vec{k})b(\vec{k}) \right). \tag{1.112}$$

The conserved charge

The complex scalar field has a global $U(1)$ symmetry: $\phi \to e^{-i\alpha}\phi$, $\phi^\dagger \to e^{+i\alpha}\phi^\dagger$ (the field ϕ carries charge -1, so $\phi^\dagger$ creates positive-charge states—this convention makes the charge of the a-particle in the mode expansion below come out positive). By Noether's theorem, this leads to a conserved current:

$$j^\mu = i\left(\phi^\dagger \partial^\mu \phi - (\partial^\mu \phi^\dagger)\phi\right). \tag{1.113}$$

Box 1.13 Derivation of the Noether current and its conservation

Step 1: Noether current from the symmetry. Under an infinitesimal $U(1)$ transformation $\phi \to \phi - i\alpha\phi$, $\phi^\dagger \to \phi^\dagger + i\alpha\phi^\dagger$ (with characteristic variations $\Delta\phi = -i\phi$, $\Delta\phi^\dagger = +i\phi^\dagger$), the Lagrangian $\mathcal{L} = \partial_\mu\phi^\dagger\partial^\mu\phi - m^2\phi^\dagger\phi$ is invariant. The Noether current is:

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\dagger)}\Delta\phi^\dagger = (\partial^\mu\phi^\dagger)(-i\phi) + (\partial^\mu\phi)(+i\phi^\dagger) = i\left(\phi^\dagger\partial^\mu\phi - (\partial^\mu\phi^\dagger)\phi\right). \tag{1.114}$$

Step 2: Proof of current conservation. Compute the divergence:

$$\begin{aligned} \partial_\mu j^\mu &= i\partial_\mu\left(\phi^\dagger\partial^\mu\phi - (\partial^\mu\phi^\dagger)\phi\right) \\ &= i\left((\partial_\mu\phi^\dagger)(\partial^\mu\phi) + \phi^\dagger(\partial_\mu\partial^\mu\phi) - (\partial_\mu\partial^\mu\phi^\dagger)\phi - (\partial^\mu\phi^\dagger)(\partial_\mu\phi)\right) \\ &= i\left(\phi^\dagger\Box\phi - (\Box\phi^\dagger)\phi\right). \end{aligned} \tag{1.115}$$

Using the Klein-Gordon equations $\Box\phi = -m^2\phi$ and $\Box\phi^\dagger = -m^2\phi^\dagger$:

$$\partial_\mu j^\mu = i\left(\phi^\dagger(-m^2\phi) - (-m^2\phi^\dagger)\phi\right) = i\left(-m^2\phi^\dagger\phi + m^2\phi^\dagger\phi\right) = 0. \tag{1.116}$$

The conserved charge is:

$$Q = \int d^3x\, j^0 = i\int d^3x\,(\phi^\dagger\dot{\phi} - \dot{\phi}^\dagger\phi) = i\int d^3x\,(\phi^\dagger\pi^\dagger - \pi\phi), \tag{1.117}$$

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

Box 1.14 Derivation of the charge operator

Substituting the mode expansions for ϕ and $\phi^\dagger$ given just above into the charge:

$$Q = i \int d^3x \,(\phi^\dagger \dot{\phi} - \dot{\phi}^\dagger \phi)\,. \tag{1.118}$$

Using $\dot{\phi} \sim -iE_k(a\,e^{-ikx} - b^\dagger e^{ikx})$ and the orthogonality of plane waves:

$$\begin{aligned} Q &= \int d^3k \left(a^\dagger(\vec{k})a(\vec{k}) - b(\vec{k})b^\dagger(\vec{k})\right) \\ &= \int d^3k \left(a^\dagger(\vec{k})a(\vec{k}) - b^\dagger(\vec{k})b(\vec{k})\right) - \delta^{(3)}(\vec{0}) \int d^3k\,. \end{aligned} \tag{1.119}$$

After normal ordering, the infinite constant is removed:

$$:Q: = \int d^3k \left(a^\dagger(\vec{k})a(\vec{k}) - b^\dagger(\vec{k})b(\vec{k})\right) = N_a - N_b\,. \tag{1.120}$$

The normal-ordered charge operator is:

$$:Q: = \int d^3k \left(a^\dagger(\vec{k})a(\vec{k}) - b^\dagger(\vec{k})b(\vec{k})\right) = N_a - N_b\,, \tag{1.121}$$

where N_a and N_b are the number operators for particles and antiparticles, respectively.

Physical interpretation:

- Particles (created by $a^\dagger$) carry charge $+1$
- Antiparticles (created by $b^\dagger$) carry charge -1
- The vacuum has zero charge: $Q\,|0\rangle = 0$
- Charge is *conserved*: $[Q, H] = 0$

Contrast with energy: While the Hamiltonian $H = \int d^3k\, E_k(a^\dagger a + b^\dagger b)$ counts particles and antiparticles with the *same sign* (both contribute positive energy), the charge $Q = N_a - N_b$ counts them with *opposite signs*. This reflects that energy is always positive, but charge can be positive or negative.

1.3 The Dirac field

The Lagrangian and Dirac equation

The Lagrangian for a free Dirac field ψ is:

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi\,, \tag{1.122}$$

where $\bar{\psi} = \psi^\dagger\gamma^0$ is the Dirac adjoint.

Box 1.15 Derivation of the Dirac equation from the Lagrangian

The Dirac Lagrangian can be written explicitly as:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi = i\bar{\psi}_\alpha(\gamma^\mu)_{\alpha\beta}\partial_\mu\psi_\beta - m\bar{\psi}_\alpha\psi_\alpha\,, \tag{1.123}$$

where $\alpha,\beta = 1,\ldots,4$ are spinor indices.
Variation with respect to $\bar{\psi}$: Treating ψ and $\bar{\psi}$ as independent variables (justified because $\bar{\psi} = \psi^\dagger\gamma^0$ involves complex conjugation), the Euler–Lagrange equation for $\bar{\psi}$ is:

$$\frac{\partial\mathcal{L}}{\partial\bar{\psi}} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{\psi})}\right) = 0\,. \tag{1.124}$$

Since $\mathcal{L}$ contains no derivatives of $\bar{\psi}$, the second term vanishes:

$$\frac{\partial\mathcal{L}}{\partial\bar{\psi}_\alpha} = i(\gamma^\mu)_{\alpha\beta}\partial_\mu\psi_\beta - m\psi_\alpha = 0\,. \tag{1.125}$$

This gives the **Dirac equation**:

$$\boxed{(i\partial\!\!\!/ - m)\psi = 0} \qquad \text{or} \qquad (i\gamma^\mu\partial_\mu - m)\psi = 0\,. \tag{1.126}$$

Variation with respect to ψ: The Euler–Lagrange equation for ψ is:

$$\frac{\partial\mathcal{L}}{\partial\psi_\beta} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi_\beta)}\right) = 0\,. \tag{1.127}$$

Computing:

$$\frac{\partial\mathcal{L}}{\partial\psi_\beta} = -m\bar{\psi}_\beta\,, \tag{1.128}$$

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi_\beta)} = i\bar{\psi}_\alpha(\gamma^\mu)_{\alpha\beta}\,. \tag{1.129}$$

Therefore:

$$-m\bar{\psi}_\beta - \partial_\mu\big(i\bar{\psi}_\alpha(\gamma^\mu)_{\alpha\beta}\big) = 0 \quad\Longrightarrow\quad -i\partial_\mu\bar{\psi}\gamma^\mu - m\bar{\psi} = 0\,. \tag{1.130}$$

This is the **adjoint Dirac equation**:

$$\bar{\psi}(i\overleftarrow{\partial\!\!\!/} + m) = 0 \qquad \text{i.e.} \qquad i(\partial_\mu\bar{\psi})\gamma^\mu + m\bar{\psi} = 0\,, \tag{1.131}$$

where $\overleftarrow{\partial}$ acts to the left.

Degrees of freedom

Now that we have the Dirac equation, we can understand the counting of degrees of freedom. The Dirac spinor ψ has four complex components, but not all represent independent physical degrees of freedom:

- **Off-shell:** The four complex components of ψ give 8 real degrees of freedom before imposing the Dirac equation.
- **On-shell:** The Dirac equation $(i\partial\!\!\!/ - m)\psi = 0$ is a *first-order* equation that relates components, effectively halving the degrees of freedom. For a plane-wave solution $\psi \sim u(\vec{p},s)\,e^{-ip\cdot x}$, the constraint $(p\!\!\!/ - m)u = 0$ reduces 4 complex components to 2 independent ones (the two spin polarizations $s = \pm 1/2$).
- **Physical count:** A massive spin-1/2 field has four physical one-particle states on shell: particle/antiparticle, each with two spin polarizations. This matches the reduction from 8 real off-shell components to 4 on-shell physical degrees of freedom.
- **Comparison with Klein–Gordon:** The scalar field ϕ is second-order in time, so both ϕ and $\dot{\phi}$ are needed as initial data (2 real DOF for real scalar, 4 for complex). The Dirac field is first-order, so only ψ is needed—but ψ has 4 complex components, giving 8 real DOF off-shell, reduced to 4 on-shell.

Canonical momenta and constraints

The momentum conjugate to ψ is:

$$\pi_\psi = \frac{\partial\mathcal{L}}{\partial(\partial_0\psi)} = i\bar{\psi}\gamma^0 = i\psi^\dagger\,. \tag{1.132}$$

The momentum conjugate to $\bar{\psi}$ is:

$$\pi_{\bar{\psi}} = \frac{\partial\mathcal{L}}{\partial(\partial_0\bar{\psi})} = 0\,. \tag{1.133}$$

Box 1.16 Physical meaning of $\pi_{\bar\psi} = 0$

The vanishing of $\pi_{\bar\psi}$ is not a pathology but reflects the **first-order nature** of the Dirac equation:

1. **Constraint, not dynamical equation:** In the Hamiltonian formalism, $\pi_{\bar\psi} = 0$ is a **primary constraint** (in Dirac's terminology). It tells us that $\bar\psi$ and ψ are not independent phase-space variables—they are related by $\bar\psi = \psi^\dagger\gamma^0$.
2. **First-order vs. second-order:** The Klein–Gordon equation is second-order in time, so ϕ and $\dot\phi = \pi$ are independent. The Dirac equation is first-order, so ψ and $\dot\psi$ are *not* independent—$\dot\psi$ is determined by spatial derivatives of ψ via the equation of motion.
3. **Counting degrees of freedom:** For the complex scalar field, we have (ϕ, π) and $(\phi^\dagger, \pi^\dagger)$—four phase-space variables. For the Dirac field, the constraint $\pi_{\bar\psi} = 0$ and the relation $\pi_\psi = i\psi^\dagger$ mean we effectively have only ψ (4 complex components = 8 real) as independent, matching the scalar field count.
4. **Anticommutators:** The canonical quantization proceeds with $\{\psi_\alpha, \pi_{\psi,\beta}\} = \{\psi_\alpha, i\psi^\dagger_\beta\} = i\delta_{\alpha\beta}\delta^{(3)}(\vec{x}-\vec{y})$, which gives the standard anti-commutation relation $\{\psi_\alpha, \psi^\dagger_\beta\} = \delta_{\alpha\beta}\delta^{(3)}$.

The Hamiltonian

The Hamiltonian density can be written in two equivalent forms:

$$\mathcal{H} = \psi^\dagger(-i\vec{\alpha}\cdot\nabla + \beta m)\psi\,, \tag{1.134}$$

$$\mathcal{H} = \bar\psi(-i\vec{\gamma}\cdot\nabla + m)\psi\,, \tag{1.135}$$

where $\beta = \gamma^0$, $\vec{\alpha} = \gamma^0\vec{\gamma}$, and $\vec{\gamma} = (\gamma^1, \gamma^2, \gamma^3)$.

The matrices $\vec{\alpha}$ and β satisfy the anticommutation relations $\{\alpha^i, \alpha^j\} = 2\delta^{ij}$, $\{\alpha^i, \beta\} = 0$, and $\beta^2 = 1$, which are the defining relations of the Dirac algebra in the Hamiltonian formalism.

Box 1.17 Derivation of the Dirac Hamiltonian

The Hamiltonian density is obtained via the Legendre transformation. Since $\pi_{\bar\psi} = 0$:

$$\mathcal{H} = \pi_\psi\dot\psi - \mathcal{L} = i\psi^\dagger\dot\psi - \mathcal{L}\,. \tag{1.136}$$

Step 1: Write the Lagrangian explicitly:

$$\mathcal{L} = i\bar\psi\gamma^0\partial_0\psi + i\bar\psi\gamma^k\partial_k\psi - m\bar\psi\psi = i\psi^\dagger\dot\psi + i\bar\psi\gamma^k\partial_k\psi - m\bar\psi\psi\,, \tag{1.137}$$

where we used $\bar{\psi}\gamma^0 = \psi^\dagger\gamma^0\gamma^0 = \psi^\dagger$.
Step 2: Compute the Hamiltonian density:

$$\begin{aligned}\mathcal{H} &= i\psi^\dagger\dot{\psi} - \left(i\psi^\dagger\dot{\psi} + i\bar{\psi}\gamma^k\partial_k\psi - m\bar{\psi}\psi\right) \\ &= -i\bar{\psi}\gamma^k\partial_k\psi + m\bar{\psi}\psi \\ &= \bar{\psi}(-i\gamma^k\partial_k + m)\psi\,. \end{aligned} \tag{1.138}$$

Step 3: Express in terms of $\psi^\dagger$. Using $\bar{\psi} = \psi^\dagger\gamma^0$:

$$\mathcal{H} = \psi^\dagger\gamma^0(-i\gamma^k\partial_k + m)\psi = \psi^\dagger(-i\gamma^0\gamma^k\partial_k + m\gamma^0)\psi\,. \tag{1.139}$$

Defining $\alpha^k = \gamma^0\gamma^k$ and $\beta = \gamma^0$:

$$\mathcal{H} = \psi^\dagger(-i\alpha^k\partial_k + \beta m)\psi = \psi^\dagger(-i\vec{\alpha}\cdot\nabla + \beta m)\psi\,. \tag{1.140}$$

Alternative form using γ matrices:

$$\mathcal{H} = \psi^\dagger\gamma^0(-i\gamma^k\partial_k + m)\psi = \bar{\psi}(-i\vec{\gamma}\cdot\nabla + m)\psi\,. \tag{1.141}$$

Mode expansion

The field operators are expanded as:

$$\psi(x) = \sum_{s=\pm}\int \frac{d^3p}{(2\pi)^{3/2}}\sqrt{\frac{m}{E_p}}\left[b(\vec{p},s)\,u(\vec{p},s)\,e^{-ip\cdot x} + d^\dagger(\vec{p},s)\,v(\vec{p},s)\,e^{ip\cdot x}\right] \tag{1.142}$$

$$\bar{\psi}(x) = \sum_{s=\pm}\int \frac{d^3p}{(2\pi)^{3/2}}\sqrt{\frac{m}{E_p}}\left[b^\dagger(\vec{p},s)\,\bar{u}(\vec{p},s)\,e^{ip\cdot x} + d(\vec{p},s)\,\bar{v}(\vec{p},s)\,e^{-ip\cdot x}\right]\,. \tag{1.143}$$

Here:

- $b(\vec{p},s)$ annihilates an electron with momentum $\vec{p}$ and spin s when acting on kets
- $d^\dagger(\vec{p},s)$ creates a positron with momentum $\vec{p}$ and spin s
- $u(\vec{p},s)$ and $v(\vec{p},s)$ are the positive and negative energy spinors

When acting on bras, the roles of creation/annihilation are reversed.

Spinor properties

The spinors satisfy the Dirac equation in momentum space:

$$(\not{p}-m)u(\vec{p},s)=0\,,\qquad \bar{u}(\vec{p},s)(\not{p}-m)=0\,, \tag{1.144}$$
$$(\not{p}+m)v(\vec{p},s)=0\,,\qquad \bar{v}(\vec{p},s)(\not{p}+m)=0\,. \tag{1.145}$$

The orthogonality relations are:

$$\bar{u}(\vec{p},s)u(\vec{p},s')=\delta_{ss'}\,,\qquad \bar{v}(\vec{p},s)v(\vec{p},s')=-\delta_{ss'}\,, \tag{1.146}$$
$$u^\dagger(\vec{p},s)u(\vec{p},s')=\frac{E_p}{m}\delta_{ss'}\,,\qquad v^\dagger(\vec{p},s)v(\vec{p},s')=\frac{E_p}{m}\delta_{ss'}\,, \tag{1.147}$$
$$\bar{v}(\vec{p},s)u(\vec{p},s')=0\,. \tag{1.148}$$

The completeness relations (spin sums) are:

$$\sum_s u_\alpha(\vec{p},s)\bar{u}_\beta(\vec{p},s)=\left(\frac{\not{p}+m}{2m}\right)_{\alpha\beta}\,, \tag{1.149}$$
$$\sum_s v_\alpha(\vec{p},s)\bar{v}_\beta(\vec{p},s)=\left(\frac{\not{p}-m}{2m}\right)_{\alpha\beta}\,. \tag{1.150}$$

Box 1.18 Proofs of spinor properties

1. Orthogonality $\bar{u}(\vec{p},s)u(\vec{p},s')=\delta_{ss'}$:
In the rest frame $\vec{p}=0$, the spinors take a simple form. Let $u^{(s)}$ denote the rest-frame spinor for spin s. In the Dirac representation with BD normalization:

$$u^{(1)}=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad u^{(2)}=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\quad v^{(1)}=\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\quad v^{(2)}=\begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{1.151}$$

With $\gamma^0=\mathrm{diag}(1,1,-1,-1)$ in the Dirac representation, we have $\bar{u}^{(s)}=u^{(s)\dagger}\gamma^0$:

$$\bar{u}^{(1)}u^{(1)}=(1,0,0,0)\begin{pmatrix}1\\0\\0\\0\end{pmatrix}=1\,,\qquad \bar{u}^{(1)}u^{(2)}=0\,. \tag{1.152}$$

For a general frame, boost the spinors: the combination $\bar{u}u$ is Lorentz invariant (a scalar), so $\bar{u}(\vec{p},s)u(\vec{p},s')=\delta_{ss'}$ in all frames.
2. Orthogonality $\bar{v}(\vec{p},s)v(\vec{p},s')=-\delta_{ss'}$:

Similarly, in the rest frame:

$$\bar{v}^{(1)}v^{(1)} = (0,0,-1,0)\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = -1\,. \tag{1.153}$$

The minus sign arises from γ^0 having -1 entries in the lower block.
3. Mixed orthogonality $\bar{v}(\vec{p},s)u(\vec{p},s')=0$:
From the Dirac equations $(\not{p}-m)u=0$ and $\bar{v}(\not{p}+m)=0$:

$$\bar{v}(\not{p}+m)u = \bar{v}\not{p}u + m\bar{v}u = 0\,. \tag{1.154}$$

But also $\not{p}u = mu$, so $\bar{v}mu + m\bar{v}u = 2m\bar{v}u = 0$, hence $\bar{v}u = 0$.
4. Relation between $u^\dagger u$ and $\bar{u}u$:
Note first that $u^\dagger u$ is *not* a Lorentz scalar — it is the timelike component of the conserved current $\bar{u}\gamma^\mu u$, since $\bar{u}\gamma^0 u = u^\dagger\gamma^0\gamma^0 u = u^\dagger u$. We therefore cannot just "invoke Lorentz covariance" to deduce its value from the rest frame; we must derive it directly. The cleanest route uses the well-known identity

$$\bar{u}(p,s)\,\gamma^\mu\,u(p,s') = \frac{p^\mu}{m}\,\bar{u}(p,s)\,u(p,s') = \frac{p^\mu}{m}\,\delta_{ss'}\,, \tag{1.155}$$

which is itself a consequence of the Gordon decomposition together with the spin-summed Dirac equation $(\not{p}-m)u=0$ (a self-contained proof is given below). Taking $\mu=0$ in (1.155):

$$u^\dagger(\vec{p},s)\,u(\vec{p},s') \;=\; \bar{u}(p,s)\,\gamma^0\,u(p,s') \;=\; \frac{p^0}{m}\,\delta_{ss'} \;=\; \frac{E_p}{m}\,\delta_{ss'}\,, \tag{1.156}$$

which is the desired result.
Sketch of (1.155). Using $\not{p}\,u = mu$ and $\bar{u}\not{p} = m\bar{u}$, write $2m\,\bar{u}\gamma^\mu u = \bar{u}(\not{p}\gamma^\mu + \gamma^\mu\not{p})u = \bar{u}\{\not{p},\gamma^\mu\}u = 2p^\mu\,\bar{u}u$. Hence $\bar{u}\gamma^\mu u = (p^\mu/m)\bar{u}u$, which on the diagonal $s=s'$ gives (1.155) after using $\bar{u}u = 1$.
5. Completeness relation $\sum_s u(\vec{p},s)\bar{u}(\vec{p},s) = (\not{p}+m)/(2m)$:
The sum $\sum_s u\bar{u}$ is a 4×4 matrix. It must be built from available objects: $\mathbb{1}$, γ^μ, and p^μ. The most general Lorentz-covariant form is:

$$\sum_s u(\vec{p},s)\bar{u}(\vec{p},s) = A\not{p} + B\,\mathbb{1}\,. \tag{1.157}$$

To find A and B, use two constraints:

- *Dirac equation:* $(\not{p}-m)u=0$ implies $(\not{p}-m)\sum_s u\bar{u}=0$, so $(A\not{p}^2+B\not{p})-m(A\not{p}+B)=0$. Using $\not{p}^2=p^2=m^2$: $(Am^2-mB)+(B-mA)\not{p}=0$. This gives $B=mA$.
- *Trace:* $\mathrm{Tr}(\sum_s u\bar{u})=\sum_s \bar{u}u=2$ (summing over 2 spin states). Also, $\mathrm{Tr}(A\not{p}+B)=4B$ (since $\mathrm{Tr}(\not{p})=0$). So $4B=2$, giving $B=1/2$ and $A=1/(2m)$.

Therefore:

$$\sum_s u(\vec{p},s)\bar{u}(\vec{p},s)=\frac{\not{p}+m}{2m}. \tag{1.158}$$

6. Completeness for v:
The same argument with $(\not{p}+m)v=0$ and $\sum_s \bar{v}v=-2$ gives:

$$\sum_s v(\vec{p},s)\bar{v}(\vec{p},s)=\frac{\not{p}-m}{2m}. \tag{1.159}$$

Explicit construction of spinors in the Dirac representation

We now construct the explicit form of the spinors $u(\vec{p},s)$ and $v(\vec{p},s)$ in the **Dirac (standard) representation** of the gamma matrices, following the **Bjorken–Drell (BD) normalization** conventions.

The Dirac representation of gamma matrices

In the Dirac (or standard) representation, the gamma matrices are:

$$\gamma^0=\begin{pmatrix}\mathbb{1}_2 & 0\\ 0 & -\mathbb{1}_2\end{pmatrix},\qquad \gamma^i=\begin{pmatrix}0 & \sigma^i\\ -\sigma^i & 0\end{pmatrix}, \tag{1.160}$$

where σ^i are the Pauli matrices:

$$\sigma^1=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad \sigma^2=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\quad \sigma^3=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{1.161}$$

The matrices $\alpha^i=\gamma^0\gamma^i$ and $\beta=\gamma^0$ appearing in the Hamiltonian are:

$$\alpha^i=\begin{pmatrix}0 & \sigma^i\\ \sigma^i & 0\end{pmatrix},\qquad \beta=\begin{pmatrix}\mathbb{1}_2 & 0\\ 0 & -\mathbb{1}_2\end{pmatrix}. \tag{1.162}$$

The fifth gamma matrix is:

$$\gamma^5\equiv i\gamma^0\gamma^1\gamma^2\gamma^3=\begin{pmatrix}0 & \mathbb{1}_2\\ \mathbb{1}_2 & 0\end{pmatrix}. \tag{1.163}$$

Rest-frame spinors

In the rest frame ($\vec{p} = 0$, $p^\mu = (m, \vec{0})$), the Dirac equation $(\not{p} - m)u = 0$ becomes:

$$(\gamma^0 m - m)u = m(\gamma^0 - 1)u = 0. \tag{1.164}$$

In the Dirac representation, $\gamma^0 - 1 = \mathrm{diag}(0,0,-2,-2)$, so the solutions have vanishing lower components.

With the BD normalization $\bar{u}u = 1$, the rest-frame spinors are:

$$u^{(1)}(0) = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \qquad u^{(2)}(0) = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}. \tag{1.165}$$

For the negative-energy spinors, $(\not{p} + m)v = 0$ gives $(\gamma^0 + 1)v = 0$, so the solutions have vanishing upper components:

$$v^{(1)}(0) = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \qquad v^{(2)}(0) = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{1.166}$$

Spin interpretation: The labels (1) and (2) correspond to spin-up and spin-down along the z-axis. In the rest frame, the spin operator is $\vec{S} = \frac{1}{2}\vec{\Sigma}$ with:

$$\Sigma^i = \begin{pmatrix}\sigma^i & 0\\ 0 & \sigma^i\end{pmatrix}, \qquad S^3 u^{(1)}(0) = +\tfrac{1}{2}u^{(1)}(0), \quad S^3 u^{(2)}(0) = -\tfrac{1}{2}u^{(2)}(0). \tag{1.167}$$

Boosted spinors: general momentum

To obtain spinors for arbitrary momentum $\vec{p}$, we apply a Lorentz boost to the rest-frame spinors. The boost is implemented by the spinor transformation:

$$u(\vec{p}, s) = \frac{\not{p} + m}{\sqrt{2m(E_p + m)}} u^{(s)}(0). \tag{1.168}$$

Box 1.19 Derivation of the boost formula

A Lorentz boost along direction $\hat{n}$ with rapidity η is represented on spinors by:

$$S(\Lambda) = \exp\left(\frac{\eta}{2}\vec{\alpha}\cdot\hat{n}\right) = \cosh\frac{\eta}{2} + \vec{\alpha}\cdot\hat{n}\,\sinh\frac{\eta}{2}, \tag{1.169}$$

where $\tanh\eta = |\vec{p}|/E_p$, so $\cosh\eta = E_p/m$ and $\sinh\eta = |\vec{p}|/m$.

Using the identities $\cosh(\eta/2) = \sqrt{(E_p+m)/(2m)}$ and $\sinh(\eta/2) = |\vec{p}|/\sqrt{2m(E_p+m)}$:

$$S(\Lambda) = \frac{E_p + m + \vec{\alpha}\cdot\vec{p}}{\sqrt{2m(E_p+m)}}. \tag{1.170}$$

Since $\vec{\alpha}\cdot\vec{p} = \gamma^0\vec{\gamma}\cdot\vec{p}$, and using $\{\gamma^0,\gamma^i\} = 0$ together with $\gamma^0 u^{(s)}(0) = +u^{(s)}(0)$ in the rest frame, one finds

$$\vec{\alpha}\cdot\vec{p}\,u^{(s)}(0) = \gamma^0\vec{\gamma}\cdot\vec{p}\,u^{(s)}(0) = -\vec{\gamma}\cdot\vec{p}\,\gamma^0\,u^{(s)}(0) = -\vec{\gamma}\cdot\vec{p}\,u^{(s)}(0)\,.$$

Therefore

$$S(\Lambda)\,u^{(s)}(0) = \frac{E_p + m - \vec{\gamma}\cdot\vec{p}}{\sqrt{2m(E_p+m)}}\,u^{(s)}(0) = \frac{\gamma^0 E_p + m - \vec{\gamma}\cdot\vec{p}}{\sqrt{2m(E_p+m)}}\,u^{(s)}(0), \tag{1.171}$$

where in the last step we used $E_p u^{(s)}(0) = \gamma^0 E_p u^{(s)}(0)$ (since $\gamma^0 u(0) = u(0)$). Recognizing $\gamma^0 E_p - \vec{\gamma}\cdot\vec{p} = \gamma^\mu p_\mu = \not{p}$ (with $p_\mu = (E_p, -\vec{p})$ in the BD metric):

$$u(\vec{p},s) = \frac{\not{p} + m}{\sqrt{2m(E_p+m)}}\,u^{(s)}(0). \tag{1.172}$$

Explicit form. Writing $u = \begin{pmatrix}\xi\\ \eta\end{pmatrix}$ with two-component spinors ξ and η, and using:

$$\not{p} + m = \begin{pmatrix} E_p + m & -\vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -E_p + m \end{pmatrix}, \tag{1.173}$$

we obtain for $u^{(s)}(0) = \begin{pmatrix}\xi^{(s)}\\ 0\end{pmatrix}$:

$$\boxed{u(\vec{p},s) = \sqrt{\frac{E_p+m}{2m}}\begin{pmatrix} \xi^{(s)} \\ \dfrac{\vec{\sigma}\cdot\vec{p}}{E_p+m}\,\xi^{(s)} \end{pmatrix}} \tag{1.174}$$

where $\xi^{(1)} = \begin{pmatrix}1\\0\end{pmatrix}$ and $\xi^{(2)} = \begin{pmatrix}0\\1\end{pmatrix}$ are the two-component Pauli spinors.

For the v spinors, starting from $v^{(s)}(0) = \begin{pmatrix} 0 \\ \eta^{(s)} \end{pmatrix}$ with $\eta^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\eta^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$:

$$\boxed{v(\vec{p},s) = \sqrt{\frac{E_p+m}{2m}} \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{E_p+m}\eta^{(s)} \\ \eta^{(s)} \end{pmatrix}} \tag{1.175}$$

Explicit spinors for motion along z-axis

For $\vec{p} = (0,0,p_z)$ with $p_z > 0$, we have $\vec{\sigma}\cdot\vec{p} = p_z\sigma^3 = p_z\,\mathrm{diag}(1,-1)$.

Particle spinors:

$$u(\vec{p},\uparrow) = \sqrt{\frac{E_p+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E_p+m} \\ 0 \end{pmatrix}, \qquad u(\vec{p},\downarrow) = \sqrt{\frac{E_p+m}{2m}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{-p_z}{E_p+m} \end{pmatrix}. \tag{1.176}$$

Antiparticle spinors:

$$v(\vec{p},\uparrow) = \sqrt{\frac{E_p+m}{2m}} \begin{pmatrix} \frac{p_z}{E_p+m} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad v(\vec{p},\downarrow) = \sqrt{\frac{E_p+m}{2m}} \begin{pmatrix} 0 \\ \frac{-p_z}{E_p+m} \\ 0 \\ 1 \end{pmatrix}. \tag{1.177}$$

Verification of normalization

Let us verify that these spinors satisfy the BD normalization $\bar{u}u = 1$.

Direct calculation:

$$\begin{aligned} \bar{u}(\vec{p},s)u(\vec{p},s) = u^\dagger\gamma^0 u &= \frac{E_p+m}{2m}\begin{pmatrix} \xi^{(s)\dagger}, & \xi^{(s)\dagger}\frac{\vec{\sigma}\cdot\vec{p}}{E_p+m} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \xi^{(s)} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E_p+m}\xi^{(s)} \end{pmatrix} \\ &= \frac{E_p+m}{2m}\left[\xi^{(s)\dagger}\xi^{(s)} - \frac{(\vec{\sigma}\cdot\vec{p})^2}{(E_p+m)^2}\xi^{(s)\dagger}\xi^{(s)}\right]. \end{aligned} \tag{1.178}$$

Using $(\vec{\sigma}\cdot\vec{p})^2 = |\vec{p}|^2 = E_p^2 - m^2 = (E_p-m)(E_p+m)$ and $\xi^{(s)\dagger}\xi^{(s)} = 1$:

$$\bar{u}u = \frac{E_p+m}{2m}\left[1 - \frac{E_p-m}{E_p+m}\right] = \frac{E_p+m}{2m}\cdot\frac{2m}{E_p+m} = 1. \quad \checkmark \tag{1.179}$$

Similarly, one can verify $\bar{v}v = -1$, where the minus sign arises from the -1 entries in the lower block of γ^0.

Relation between u and v via charge conjugation

The antiparticle spinor v can be obtained from u through the *charge conjugation matrix*

$$C = i\gamma^0\gamma^2\,, \tag{1.180}$$

which is the unique (up to phase) 4×4 matrix satisfying $C^{-1}\gamma^\mu C = -(\gamma^\mu)^\top$, or equivalently the defining property

$$C(\gamma^\mu)^\top = -\gamma^\mu C\,, \qquad C^\dagger = C^{-1} = -C\,. \tag{1.181}$$

A more detailed discussion of C and its properties is given in Chapter 9. We prove that $C\bar{u}(p,s)^\top$ satisfies the antiparticle Dirac equation and is therefore a v spinor.

Start from the particle equation $(\not{p}-m)\,u=0$. Taking the Dirac adjoint gives $\bar{u}\,(\not{p}-m)=0$. Transposing:

$$(\not{p}^\top - m)\,\bar{u}^\top = 0\,. \tag{1.182}$$

Multiplying from the left by C and using the defining property $C(\gamma^\mu)^\top = -\gamma^\mu C$ above, which implies $C\not{p}^\top = -\not{p}\,C$:

$$C(\not{p}^\top - m)\,\bar{u}^\top = (-\not{p}\,C - mC)\,\bar{u}^\top = -(\not{p}+m)\,C\bar{u}^\top = 0\,. \tag{1.183}$$

Therefore $C\bar{u}(p,s)^\top$ satisfies $(\not{p}+m)\psi=0$, which is precisely the Dirac equation for antiparticle spinors (1.145). With an appropriate phase convention for the two-component spinors $\eta^{(s)}$ relative to $\xi^{(s)}$, one can identify:

$$\boxed{v(p,s) = C\bar{u}(p,s)^\top} \tag{1.184}$$

Non-relativistic and ultra-relativistic limits

Non-relativistic limit ($|\vec{p}| \ll m$, so $E_p \approx m$):

$$u(\vec{p},s) \approx \begin{pmatrix} \xi^{(s)} \\ \frac{\vec{\sigma}\cdot\vec{p}}{2m}\,\xi^{(s)} \end{pmatrix}. \tag{1.185}$$

The lower ("small") components are suppressed by $|\vec{p}|/m \ll 1$. This is the regime where the Dirac equation reduces to the Pauli equation.

Ultra-relativistic limit ($|\vec{p}| \gg m$, so $E_p \approx |\vec{p}|$):

$$u(\vec{p},s) \approx \sqrt{\frac{E_p}{2m}} \begin{pmatrix} \xi^{(s)} \\ (\hat{p}\cdot\vec{\sigma})\,\xi^{(s)} \end{pmatrix}, \tag{1.186}$$

where $\hat{p} = \vec{p}/|\vec{p}|$. The upper and lower components become comparable in magnitude. In this limit, chirality (eigenvalue of γ^5) becomes approximately equal to helicity (projection of spin on momentum).

Helicity spinors

For many applications, especially in high-energy physics, it is convenient to use **helicity spinors**—eigenstates of the helicity operator:

$$h = \frac{\vec{\Sigma}\cdot\vec{p}}{2|\vec{p}|} = \frac{1}{2|\vec{p}|}\begin{pmatrix} \vec{\sigma}\cdot\vec{p} & 0 \\ 0 & \vec{\sigma}\cdot\vec{p} \end{pmatrix}. \tag{1.187}$$

The helicity eigenvalues are $\lambda = \pm 1/2$, corresponding to spin parallel (+) or antiparallel (−) to the momentum direction.

For motion along the z-axis, the helicity eigenstates coincide with the spin eigenstates constructed above. For general $\vec{p}$, one must rotate the two-component spinors $\xi^{(s)}$ to the direction of $\vec{p}$.

If $\vec{p}$ has polar angles (θ, ϕ), the helicity eigenspinors are:

$$\xi^{(+)}(\hat{p}) = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix}, \qquad \xi^{(-)}(\hat{p}) = \begin{pmatrix} -e^{-i\phi}\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{pmatrix}. \tag{1.188}$$

These satisfy $(\vec{\sigma}\cdot\hat{p})\,\xi^{(\pm)} = \pm\xi^{(\pm)}$.

Anticommutation relations

The Dirac field obeys **anticommutation** relations (required by the spin-statistics theorem):

$$\{b(\vec{p},s), b^\dagger(\vec{p}\,',s')\} = \{d(\vec{p},s), d^\dagger(\vec{p}\,',s')\} = \delta_{ss'}\delta^{(3)}(\vec{p}-\vec{p}\,'), \tag{1.189}$$

$$\{b(\vec{p},s), b(\vec{p}\,',s')\} = \{d(\vec{p},s), d(\vec{p}\,',s')\} = 0, \tag{1.190}$$

$$\{b(\vec{p},s), d(\vec{p}\,',s')\} = \{b(\vec{p},s), d^\dagger(\vec{p}\,',s')\} = 0. \tag{1.191}$$

These imply the equal-time anticommutation relations for the fields:

$$\{\psi_\alpha(t,\vec{x}), \psi^\dagger_\beta(t,\vec{x}\,')\} = \delta_{\alpha\beta}\delta^{(3)}(\vec{x}-\vec{x}\,'), \tag{1.192}$$

$$\{\psi_\alpha(t,\vec{x}), \psi_\beta(t,\vec{x}\,')\} = \{\psi^\dagger_\alpha(t,\vec{x}), \psi^\dagger_\beta(t,\vec{x}\,')\} = 0. \tag{1.193}$$

Box 1.20 Field anticommutators from operator anticommutators

We derive $\{\psi_\alpha(t,\vec{x}), \psi^\dagger_\beta(t,\vec{x}\,')\} = \delta_{\alpha\beta}\delta^{(3)}(\vec{x}-\vec{x}\,')$ from the mode expansion.

At equal times ($t = 0$ for simplicity), the mode expansion gives:

$$\psi_\alpha(\vec{x}) = \sum_s \int \frac{d^3p}{(2\pi)^{3/2}} \sqrt{\frac{m}{E_p}} \left[b_{\vec{p},s}\, u_\alpha(\vec{p},s)\, e^{i\vec{p}\cdot\vec{x}} + d^\dagger_{\vec{p},s}\, v_\alpha(\vec{p},s)\, e^{-i\vec{p}\cdot\vec{x}} \right], \tag{1.194}$$

$$\psi^\dagger_\beta(\vec{x}') = \sum_{s'} \int \frac{d^3p'}{(2\pi)^{3/2}} \sqrt{\frac{m}{E_{p'}}} \Big[b^\dagger_{\vec{p}',s'}\, u^*_\beta(\vec{p}',s')\, e^{-i\vec{p}'\cdot\vec{x}'} + d_{\vec{p}',s'}\, v^*_\beta(\vec{p}',s')\, e^{i\vec{p}'\cdot\vec{x}'} \Big]. \tag{1.195}$$

The anticommutator $\{\psi_\alpha, \psi^\dagger_\beta\}$ has four terms. Using $\{b, b^\dagger\} = \{d, d^\dagger\} = \delta_{ss'}\delta^{(3)}$ and $\{b, d\} = \{b, d^\dagger\} = 0$:

$$\{\psi_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{x}')\} = \sum_s \int \frac{d^3p}{(2\pi)^3} \frac{m}{E_p} \Big[u_\alpha(\vec{p},s) u^*_\beta(\vec{p},s)\, e^{i\vec{p}\cdot(\vec{x}-\vec{x}')} + v_\alpha(\vec{p},s) v^*_\beta(\vec{p},s)\, e^{-i\vec{p}\cdot(\vec{x}-\vec{x}')} \Big]. \tag{1.196}$$

Now use the relation $u^* = \gamma^0 \bar{u}^\top$ (in standard representations), which gives:

$$\sum_s u_\alpha u^*_\beta = \sum_s u_\alpha (\gamma^0)_{\beta\gamma} \bar{u}_\gamma = \left[(\not{p} + m)\gamma^0/(2m) \right]_{\alpha\beta}. \tag{1.197}$$

Similarly for v. Using the identity (valid for on-shell p):

$$\frac{(\not{p} + m)\gamma^0 + (\not{p} - m)\gamma^0}{2m} = \frac{2p^0\gamma^0\gamma^0}{2m} = \frac{E_p}{m}, \tag{1.198}$$

where we used $\not{p}\gamma^0 + \gamma^0\not{p} = 2p^0$.
After changing $\vec{p} \to -\vec{p}$ in the second term (which is symmetric under this):

$$\{\psi_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{x}')\} = \delta_{\alpha\beta} \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{x}-\vec{x}')} = \delta_{\alpha\beta}\delta^{(3)}(\vec{x} - \vec{x}'). \tag{1.199}$$

The proof that $\{\psi, \psi\} = 0$ follows similarly: the bb and $d^\dagger d^\dagger$ terms vanish by the operator anticommutators, and the cross terms $bd^\dagger$ and $d^\dagger b$ also vanish.

Four-momentum operator

The four-momentum operator is:

$$P^\mu = \sum_s \int d^3p\, p^\mu \left[b^\dagger(\vec{p},s)b(\vec{p},s) + d^\dagger(\vec{p},s)d(\vec{p},s)\right] . \tag{1.200}$$

Physical interpretation:

- $P^0 = H$ is the Hamiltonian (total energy). The normal-ordered form ensures $\langle 0|H|0\rangle = 0$.
- $\vec{P} = (P^1, P^2, P^3)$ is the total three-momentum.
- $b^\dagger b$ counts electrons: $N_e = \sum_s \int d^3p\, b^\dagger(\vec{p},s)b(\vec{p},s)$.
- $d^\dagger d$ counts positrons: $N_{\bar{e}} = \sum_s \int d^3p\, d^\dagger(\vec{p},s)d(\vec{p},s)$.
- Both particles and antiparticles contribute *positively* to the energy and momentum. This is a consequence of normal ordering and the anticommutation relations.

P^μ as the generator of spacetime translations. The four-momentum operator generates infinitesimal spacetime translations via the commutator:

$$[\psi(x), P^\mu] = +i\partial^\mu\psi(x) . \tag{1.201}$$

This means that under a translation $x^\mu \to x^\mu + a^\mu$, the field transforms as:

$$\psi(x) \to e^{ia\cdot P}\psi(x)e^{-ia\cdot P} = \psi(x+a) . \tag{1.202}$$

This is the quantum field theory analogue of the classical mechanics relation $\{f, p\} = \partial f/\partial q$, where the momentum generates translations in position space.

Box 1.21 Proof that $[\psi(x), P^\mu] = +i\partial^\mu\psi(x)$

We prove this for $P^0 = H$ (time translations); the spatial components follow analogously.

Method 1: Direct calculation from the mode expansion. Recall the Heisenberg-picture mode expansion

$$\psi(x) = \sum_s \int \frac{d^3p}{(2\pi)^{3/2}} \sqrt{\frac{m}{E_p}} \left[b_{\vec{p},s}\, u(p,s)\, e^{-ip\cdot x} + d^\dagger_{\vec{p},s}\, v(p,s)\, e^{+ip\cdot x}\right],$$

together with the normal-ordered Hamiltonian $H = \sum_s \int d^3k\, E_k \left[b^\dagger_{\vec{k},s} b_{\vec{k},s} + d^\dagger_{\vec{k},s} d_{\vec{k},s}\right]$. Using the canonical anticommutator identity for fermionic operators,

$$[A, BC] = \{A, B\}C - B\{A, C\} , \tag{1.203}$$

we evaluate the relevant brackets:

$$
\begin{aligned}
[b_{\vec{p},s}, b^\dagger_{\vec{k},s'} b_{\vec{k},s'}] &= \{b_{\vec{p},s}, b^\dagger_{\vec{k},s'}\}\, b_{\vec{k},s'} - b^\dagger_{\vec{k},s'}\, \{b_{\vec{p},s}, b_{\vec{k},s'}\} \\
&= \delta_{ss'}\delta^{(3)}(\vec{p}-\vec{k})\, b_{\vec{k},s'} - 0 = +\delta_{ss'}\delta^{(3)}(\vec{p}-\vec{k})\, b_{\vec{p},s}\,,
\end{aligned}
\tag{1.204}
$$

$$
\begin{aligned}
[d^\dagger_{\vec{p},s}, d^\dagger_{\vec{k},s'} d_{\vec{k},s'}] &= \{d^\dagger_{\vec{p},s}, d^\dagger_{\vec{k},s'}\}\, d_{\vec{k},s'} - d^\dagger_{\vec{k},s'}\, \{d^\dagger_{\vec{p},s}, d_{\vec{k},s'}\} \\
&= 0 - \delta_{ss'}\delta^{(3)}(\vec{p}-\vec{k})\, d^\dagger_{\vec{k},s'} = -\delta_{ss'}\delta^{(3)}(\vec{p}-\vec{k})\, d^\dagger_{\vec{p},s}\,.
\end{aligned}
\tag{1.205}
$$

Substituting and collapsing the delta functions:

$$
[\psi(x), H] = \sum_s \int \frac{d^3p}{(2\pi)^{3/2}} \sqrt{\tfrac{m}{E_p}}\, E_p \left[+b_{\vec{p},s}\, u\, e^{-ip\cdot x} - d^\dagger_{\vec{p},s}\, v\, e^{+ip\cdot x} \right].
\tag{1.206}
$$

On the other hand, differentiating the mode expansion w.r.t. x^0 brings down $(-iE_p)$ on the b term and $(+iE_p)$ on the $d^\dagger$ term, so

$$
i\partial_0\psi(x) = \sum_s \int \frac{d^3p}{(2\pi)^{3/2}} \sqrt{\tfrac{m}{E_p}}\, E_p \left[+b_{\vec{p},s}\, u\, e^{-ip\cdot x} - d^\dagger_{\vec{p},s}\, v\, e^{+ip\cdot x} \right].
\tag{1.207}
$$

Comparing (1.206) and (1.207): $[\psi, H] = +i\partial_0\psi = +i\partial^0\psi$, since $\partial_0 = \partial^0$ in Bjorken–Drell metric.

Method 2: From the Heisenberg equation. In the Heisenberg picture, operators evolve as $\frac{dO}{dt} = i[H, O]$, equivalently

$$
[O, H] \;=\; -[H, O] \;=\; -\frac{1}{i}\frac{dO}{dt} \;=\; i\,\frac{dO}{dt}\,.
\tag{1.208}
$$

Applied to $\psi(x)$, for which $dO/dt = \partial_0\psi$ (no explicit Schrödinger-picture time dependence), this gives directly $[\psi, H] = +i\partial_0\psi = +i\partial^0\psi$.

For spatial translations, the same logic (now using $P^i = -\int d^3x\, \pi\, \partial_i\psi$ and the canonical equal-time anticommutator) gives $[\psi, P^i] = +i\partial^i\psi$.

The conserved charge

The Dirac Lagrangian is invariant under the global $U(1)$ transformation $\psi \to e^{-i\alpha}\psi$, $\bar{\psi} \to e^{+i\alpha}\bar{\psi}$ (the field ψ carries charge -1: it annihilates positive-charge particles, and $\bar{\psi}$ creates them; this reproduces the standard $j^\mu = \bar{\psi}\gamma^\mu\psi$ and $Q = N_b - N_d$ below). The associated Noether current is:

$$j^\mu = \bar{\psi}\gamma^\mu\psi\,. \tag{1.209}$$

Box 1.22 Derivation of the Noether current and its conservation

Step 1: Noether current from the symmetry. Under the infinitesimal transformation $\psi \to \psi - i\alpha\psi$, $\bar{\psi} \to \bar{\psi} + i\alpha\bar{\psi}$ (characteristic variations $\Delta\psi = -i\psi$, $\Delta\bar{\psi} = +i\bar{\psi}$), the Lagrangian $\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi$ is invariant. The canonical momentum conjugate to ψ is $\pi_\psi = i\bar{\psi}\gamma^0 = i\psi^\dagger$. The Noether current is:

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\Delta\psi + \Delta\bar{\psi}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{\psi})} = (i\bar{\psi}\gamma^\mu)(-i\psi) + 0 = -i^2\bar{\psi}\gamma^\mu\psi = \bar{\psi}\gamma^\mu\psi\,, \tag{1.210}$$

where we used that the Dirac Lagrangian has no $\partial_\mu\bar{\psi}$ term.
Step 2: Proof of current conservation. Compute the divergence:

$$\partial_\mu j^\mu = \partial_\mu(\bar{\psi}\gamma^\mu\psi) = (\partial_\mu\bar{\psi})\gamma^\mu\psi + \bar{\psi}\gamma^\mu(\partial_\mu\psi)\,. \tag{1.211}$$

The Dirac equation gives $i\gamma^\mu\partial_\mu\psi = m\psi$, hence $\gamma^\mu\partial_\mu\psi = -im\psi$. The adjoint Dirac equation gives $i\partial_\mu\bar{\psi}\gamma^\mu = -m\bar{\psi}$, hence $\partial_\mu\bar{\psi}\gamma^\mu = im\bar{\psi}$. Substituting:

$$\partial_\mu j^\mu = (im\bar{\psi})\psi + \bar{\psi}(-im\psi) = im\bar{\psi}\psi - im\bar{\psi}\psi = 0\,. \tag{1.212}$$

The conserved charge is:

$$Q = \int d^3x\, j^0 = \int d^3x\, \psi^\dagger\psi\,. \tag{1.213}$$

Box 1.23 Derivation of the Dirac charge operator

Using the mode expansion (1.142) at $t = 0$ (so that $\psi \sim b\,u\,e^{+i\vec{p}\cdot\vec{x}} + d^\dagger\,v\,e^{-i\vec{p}\cdot\vec{x}}$):

$$\begin{aligned}\psi^\dagger\psi = \sum_{s,s'}\int \frac{d^3p\,d^3p'}{(2\pi)^3}\frac{m}{\sqrt{E_pE_{p'}}}\Big\{&b^\dagger_{\vec{p},s}b_{\vec{p}',s'}\,u^\dagger_s(\vec{p})u_{s'}(\vec{p}')\,e^{-i(\vec{p}-\vec{p}')\cdot\vec{x}}\\ &+ b^\dagger_{\vec{p},s}d^\dagger_{\vec{p}',s'}\,u^\dagger_s(\vec{p})v_{s'}(\vec{p}')\,e^{-i(\vec{p}+\vec{p}')\cdot\vec{x}}\\ &+ d_{\vec{p},s}b_{\vec{p}',s'}\,v^\dagger_s(\vec{p})u_{s'}(\vec{p}')\,e^{+i(\vec{p}+\vec{p}')\cdot\vec{x}}\\ &+ d_{\vec{p},s}d^\dagger_{\vec{p}',s'}\,v^\dagger_s(\vec{p})v_{s'}(\vec{p}')\,e^{+i(\vec{p}-\vec{p}')\cdot\vec{x}}\Big\}\,.\end{aligned} \tag{1.214}$$

Integrating over $\vec{x}$: the *diagonal* pieces ($b^\dagger b$ and $dd^\dagger$) yield $\delta^{(3)}(\vec{p}-\vec{p}')$, while the *mixed* pieces ($b^\dagger d^\dagger$ and db) yield $\delta^{(3)}(\vec{p}+\vec{p}')$ — this is the crucial point that determines which spinor orthogonality must be used:

- For the diagonal pieces we need spinor scalar products at the *same* 3-momentum:

$$u_s^\dagger(\vec{p})\, u_{s'}(\vec{p}) = v_s^\dagger(\vec{p})\, v_{s'}(\vec{p}) = \frac{E_p}{m}\,\delta_{ss'}\,. \tag{1.215}$$

- For the mixed pieces the delta function forces $\vec{p}' = -\vec{p}$, so we need spinor scalar products at *opposite* 3-momenta, and the relevant relation is

$$u_s^\dagger(\vec{p})\, v_{s'}(-\vec{p}) = v_s^\dagger(\vec{p})\, u_{s'}(-\vec{p}) = 0\,. \tag{1.216}$$

The orthogonality (1.216) (not the naive "$u^\dagger v = 0$" at the same momentum, which is not the relation needed here) is precisely what kills the mixed terms. Both (1.215) and (1.216) follow directly from the explicit BD-normalised spinors.
With these relations,

$$\begin{aligned} Q &= \sum_s \int d^3p\,\big(b^\dagger_{\vec{p},s} b_{\vec{p},s} + d_{\vec{p},s} d^\dagger_{\vec{p},s}\big) \\ &= \sum_s \int d^3p\,\big(b^\dagger_{\vec{p},s} b_{\vec{p},s} - d^\dagger_{\vec{p},s} d_{\vec{p},s}\big) + (\text{infinite constant})\,, \end{aligned} \tag{1.217}$$

where we used the anticommutator $\{d, d^\dagger\} = \delta_{ss'}\delta^{(3)}(\vec{p}-\vec{p}')$.

The normal-ordered charge operator is:

$$:Q: = \sum_s \int d^3p\,\big(b^\dagger(\vec{p},s)b(\vec{p},s) - d^\dagger(\vec{p},s)d(\vec{p},s)\big) = N_e - N_{\bar{e}}\,, \tag{1.218}$$

where N_e and $N_{\bar{e}}$ are the number operators for electrons and positrons.
Physical interpretation:

- Electrons (created by $b^\dagger$) carry charge $+1$ (in units where the electron charge is $+1$; in SI, multiply by e)
- Positrons (created by $d^\dagger$) carry charge -1
- The vacuum has zero charge: $Q\,|0\rangle = 0$
- Charge is conserved: $[Q, H] = 0$
- The electric charge operator is $Q_{\text{em}} = -e \cdot Q = e(N_{\bar{e}} - N_e)$, where $e > 0$ is the elementary charge magnitude

Contrast with four-momentum: While P^μ sums over particles and antiparticles with the same sign (both contribute positive energy and momentum in their direction of motion), the charge $Q = N_e - N_{\bar{e}}$ counts them with opposite signs. This is the key difference between "gravitational charge" (mass-energy) and electric charge.

1.4 The photon field

The electromagnetic field is the most familiar gauge field, yet the most subtle to quantize. The difficulty arises because the four-potential A^μ contains redundant degrees of freedom: a massless spin-1 particle has only **two physical polarizations**, but A^μ has four components. This redundancy is encoded in the *gauge invariance* of electromagnetism.

The Lagrangian and equations of motion

The free electromagnetic Lagrangian is:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,, \tag{1.219}$$

where the **field strength tensor** is defined as:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,. \tag{1.220}$$

The four-potential $A^\mu = (A^0, \vec{A}) = (\Phi, \vec{A})$ contains the scalar potential Φ and the vector potential $\vec{A}$. The electric and magnetic fields are:

$$\vec{E} = -\nabla\Phi - \frac{\partial\vec{A}}{\partial t}\,, \qquad \vec{B} = \nabla\times\vec{A}\,. \tag{1.221}$$

In components, with metric $(+,-,-,-)$ and $A^\mu = (\Phi, \vec{A})$, one finds $F^{0i} = \partial^0 A^i - \partial^i A^0 = \partial_t A^i + \partial_i\Phi = -E^i$ and $B^i = -\frac{1}{2}\epsilon^{ijk}F_{jk} = \frac{1}{2}\epsilon^{ijk}F^{jk}$. Hence the identifications are

$$F^{0i} = -E^i\,, \qquad F^{ij} = -\epsilon^{ijk}B^k\,. \tag{1.222}$$

The field strength tensor in matrix form is:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{1.223}$$

The Lagrangian can be written in terms of the electric and magnetic fields:

$$\mathcal{L} = \frac{1}{2}(\vec{E}^2 - \vec{B}^2)\,. \tag{1.224}$$

Box 1.24 Derivation of Maxwell's equations from the Lagrangian

The Euler–Lagrange equations for the field A_ν are:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} - \frac{\partial \mathcal{L}}{\partial A_\nu} = 0\,. \tag{1.225}$$

Step 1: Expand the Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu)\,. \tag{1.226}$$

Step 2: Compute $\partial\mathcal{L}/\partial(\partial_\mu A_\nu)$:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} &= -\frac{1}{4}\cdot 2(\partial^\mu A^\nu - \partial^\nu A^\mu) - \frac{1}{4}\cdot 2(-1)(\partial^\nu A^\mu - \partial^\mu A^\nu) \\ &= -(\partial^\mu A^\nu - \partial^\nu A^\mu) = -F^{\mu\nu}\,.\end{aligned} \tag{1.227}$$

Step 3: Since $\partial\mathcal{L}/\partial A_\nu = 0$ (no mass term), the equation of motion is:

$$\partial_\mu F^{\mu\nu} = 0\,. \tag{1.228}$$

These are the **source-free Maxwell equations**:

- $\nu = 0$: $\partial_i F^{i0} = \nabla\cdot\vec{E} = 0$ (Gauss's law, no charges)
- $\nu = j$: $\partial_0 F^{0j} + \partial_i F^{ij} = -\dot{E}^j + (\nabla\times\vec{B})^j = 0$ (Ampère's law, no currents)

The other two Maxwell equations ($\nabla\cdot\vec{B} = 0$ and Faraday's law) follow identically from the definition $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ (the Bianchi identity $\partial_{[\lambda}F_{\mu\nu]} = 0$).

Gauge invariance: classical analysis

The electromagnetic Lagrangian possesses a fundamental symmetry: it is invariant under **gauge transformations**:

$$A_\mu(x) \longrightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu\chi(x)\,, \tag{1.229}$$

for any scalar function $\chi(x)$.

Box 1.25 Proof of gauge invariance

Under the transformation (1.229), the field strength transforms as:

$$\begin{aligned} F'_{\mu\nu} &= \partial_\mu A'_\nu - \partial_\nu A'_\mu \\ &= \partial_\mu (A_\nu + \partial_\nu \chi) - \partial_\nu (A_\mu + \partial_\mu \chi) \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu + \partial_\mu \partial_\nu \chi - \partial_\nu \partial_\mu \chi \\ &= F_{\mu\nu}, \end{aligned} \tag{1.230}$$

since partial derivatives commute. Therefore the Lagrangian $\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ is gauge invariant.

Physical meaning of gauge invariance:

- The electric and magnetic fields $\vec{E}$ and $\vec{B}$ are *physical observables* and are gauge invariant.
- The four-potential A_μ is *not* directly observable; different gauge choices give the same physics.
- Gauge invariance implies that A_μ contains *redundant* degrees of freedom.
- A massless vector field in 4D has $4 - 2 = 2$ physical degrees of freedom (two transverse polarizations).

Counting degrees of freedom: The four-potential A_μ has 4 components. Gauge invariance allows us to impose one condition (gauge fixing), reducing to 3. The constraint $\partial_\mu F^{\mu 0} = 0$ (Gauss's law) is not a dynamical equation but a constraint, removing one more degree of freedom. This leaves **2 physical degrees of freedom**, corresponding to the two transverse polarizations of the photon.

Gauge fixing: summary of choices

To quantize the theory, we must *fix the gauge*, i.e., choose a specific representative from each equivalence class of gauge-related potentials. Different gauge choices are useful in different contexts.

Gauge	Condition	Covariant?	Ghosts?
Coulomb (radiation)	$\nabla \cdot \vec{A} = 0$	No	No
Lorenz	$\partial_\mu A^\mu = 0$	Yes	No in QED
Feynman ($\xi = 1$)	$\partial_\mu A^\mu = 0$ + gauge-fixing term	Yes	No in QED
Landau ($\xi = 0$)	$\partial_\mu A^\mu = 0$ (strict)	Yes	No in QED
Axial	$n_\mu A^\mu = 0$	Partially	No
Temporal	$A^0 = 0$	No	No

Coulomb (radiation) gauge

The Coulomb gauge imposes:

$$\nabla \cdot \vec{A} = 0 \qquad \text{(transversality condition)}. \tag{1.231}$$

This condition can always be achieved: given any $\vec{A}$, we can find χ such that $\nabla \cdot (\vec{A} + \nabla\chi) = 0$ by solving the Poisson equation $\nabla^2\chi = -\nabla \cdot \vec{A}$.

In this gauge, Gauss's law $\nabla \cdot \vec{E} = 0$ becomes:

$$\nabla^2 A^0 + \frac{\partial}{\partial t}(\nabla \cdot \vec{A}) = \nabla^2 A^0 = 0. \tag{1.232}$$

For the free field with appropriate boundary conditions, this gives $A^0 = 0$.

Key features:

- Only the two transverse components of $\vec{A}$ are dynamical.
- Manifest Lorentz covariance is lost, but physical results are Lorentz invariant.
- No unphysical (ghost) states appear in the quantum theory.
- Most transparent for understanding the physical degrees of freedom.

Lorenz gauge and covariant gauges

The Lorenz gauge[2] imposes:

$$\partial_\mu A^\mu = 0 \qquad \text{(Lorenz condition)}. \tag{1.233}$$

This condition is manifestly Lorentz covariant. However, it does not completely fix the gauge: if A_μ satisfies (1.233), so does $A_\mu + \partial_\mu\chi$ for any χ satisfying $\Box\chi = 0$ (residual gauge freedom).

The equation of motion $\partial_\mu F^{\mu\nu} = 0$ can be rewritten as:

$$\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = \Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = 0. \tag{1.234}$$

[2] Named after Ludvig Lorenz, not Hendrik Lorentz.

In the Lorenz gauge, $\partial_\mu A^\mu = 0$, so the second term vanishes and the equation of motion simplifies to:

$$\Box A^\mu = 0\,, \tag{1.235}$$

i.e., each component of A^μ satisfies the massless wave equation.

For practical calculations, one adds a **gauge-fixing term** to the Lagrangian:

$$\mathcal{L}_{\mathrm{gf}} = -\frac{1}{2\xi}(\partial_\mu A^\mu)^2\,, \tag{1.236}$$

where ξ is the *gauge parameter*. Common choices:

- $\xi = 1$: **Feynman gauge** — simplest propagator: $D_{\mu\nu}(k) = -g_{\mu\nu}/(k^2 + i\epsilon)$
- $\xi = 0$: **Landau gauge** — propagator is transverse: $D_{\mu\nu}(k) = (-g_{\mu\nu} + k_\mu k_\nu/k^2)/(k^2 + i\epsilon)$
- General ξ: $D_{\mu\nu}(k) = \big(-g_{\mu\nu} + (1-\xi)k_\mu k_\nu/k^2\big)/(k^2 + i\epsilon)$

Box 1.26 Why does $\xi \to 0$ make sense?

At first sight, the limit $\xi \to 0$ seems problematic: the gauge-fixing term $-\frac{1}{2\xi}(\partial_\mu A^\mu)^2$ appears to diverge!
The resolution is that this term acts as a *penalty function* that suppresses field configurations violating the Lorenz condition. Consider the contribution to the action:

$$S_{\mathrm{gf}} = -\frac{1}{2\xi}\int d^4x\,(\partial_\mu A^\mu)^2\,. \tag{1.237}$$

In the path integral formulation: The path integral weight is e^{iS}. For any configuration with $\partial_\mu A^\mu \neq 0$, the action $S_{\mathrm{gf}} \to -\infty$ as $\xi \to 0$, making the weight $e^{iS_{\mathrm{gf}}}$ oscillate infinitely fast and average to zero. Only configurations with $\partial_\mu A^\mu = 0$ *exactly* contribute.
Analogy with constrained mechanics: This is analogous to enforcing a constraint $f(q) = 0$ by adding a potential $V = \frac{1}{2\xi}f(q)^2$ and taking $\xi \to 0$. In the limit, the system is forced onto the constraint surface.
In practice: The propagator for general ξ is:

$$D_{\mu\nu}(k) = \frac{1}{k^2 + i\epsilon}\left(-g_{\mu\nu} + (1-\xi)\frac{k_\mu k_\nu}{k^2}\right). \tag{1.238}$$

Taking $\xi \to 0$ gives the transverse (Landau) propagator, which automatically satisfies $k^\mu D_{\mu\nu}(k) = 0$. The limit is smooth and well-defined at the level of the propagator and Feynman rules.

Physical observables are gauge-parameter independent: The Ward-Takahashi identities[3] ensure that S-matrix elements do not depend on ξ.

Axial and temporal gauges

The **axial gauge** fixes $n_\mu A^\mu = 0$ for a fixed vector n^μ. The **temporal gauge** is the special case $A^0 = 0$.

These gauges avoid ghost fields but have more complicated propagators with spurious singularities that require careful treatment.

Plane wave expansion and polarization vectors

In the Coulomb gauge ($\nabla \cdot \vec{A} = 0$, $A^0 = 0$), the equation of motion for $\vec{A}$ is:

$$\Box \vec{A} = \left(\frac{\partial^2}{\partial t^2} - \nabla^2 \right) \vec{A} = 0\,. \tag{1.239}$$

The general solution is a superposition of plane waves:

$$\vec{A}(x) = \int \frac{d^3k}{(2\pi)^3}\, \tilde{\vec{A}}(\vec{k})\, e^{i(\vec{k}\cdot\vec{x}-\omega t)} + \text{c.c.}\,, \tag{1.240}$$

where $\omega = |\vec{k}|$ (dispersion relation for massless particles).

The transversality condition $\nabla \cdot \vec{A} = 0$ becomes $\vec{k} \cdot \tilde{\vec{A}}(\vec{k}) = 0$ in Fourier space. This means $\tilde{\vec{A}}$ must be *perpendicular* to the direction of propagation $\vec{k}$.

Polarization vectors

For each wave vector $\vec{k}$, we introduce two orthonormal **polarization vectors** $\vec{\varepsilon}(\vec{k}, \lambda)$ with $\lambda = 1, 2$, satisfying:

$$\vec{k} \cdot \vec{\varepsilon}(\vec{k}, \lambda) = 0 \qquad \text{(transversality)}\,, \tag{1.241}$$

$$\vec{\varepsilon}(\vec{k}, \lambda) \cdot \vec{\varepsilon}(\vec{k}, \lambda') = \delta_{\lambda\lambda'} \qquad \text{(orthonormality)}\,, \tag{1.242}$$

$$\vec{\varepsilon}(\vec{k}, +) \times \vec{\varepsilon}^{\,*}(\vec{k}, +) = -i\,\hat{k}\,, \qquad \vec{\varepsilon}(\vec{k}, -) \times \vec{\varepsilon}^{\,*}(\vec{k}, -) = +i\,\hat{k} \quad \text{for circular polarization}\,. \tag{1.243}$$

Linear polarization basis: For $\vec{k} = k\hat{z}$, a common choice is:

$$\vec{\varepsilon}(\vec{k}, 1) = \hat{x}\,, \qquad \vec{\varepsilon}(\vec{k}, 2) = \hat{y}\,. \tag{1.244}$$

[3] The Ward-Takahashi identities are relations between Green's functions that follow from gauge invariance. They are the quantum analogue of current conservation $\partial_\mu j^\mu = 0$. In momentum space, they state that contracting an amplitude with the photon momentum k^μ gives zero (up to terms involving external charged particles). These identities will be derived in detail when we discuss QED and renormalization.

Circular polarization basis:

$$\vec{\varepsilon}(\vec{k},\pm) = \frac{1}{\sqrt{2}}(\hat{x} \pm i\hat{y})\,. \tag{1.245}$$

Circular polarizations correspond to definite helicity: $\lambda = +$ (right-handed) and $\lambda = -$ (left-handed).

Completeness relation (polarization sum): The two transverse polarization vectors span the plane perpendicular to $\vec{k}$:

$$\sum_{\lambda=1}^{2} \varepsilon_i(\vec{k},\lambda)\varepsilon_j(\vec{k},\lambda) = \delta_{ij} - \frac{k_i k_j}{|\vec{k}|^2} \equiv \delta^{\perp}_{ij}(\vec{k})\,. \tag{1.246}$$

The right-hand side is the **transverse projector**: it projects any vector onto the plane perpendicular to $\vec{k}$.

Canonical quantization in the Coulomb gauge

In the Coulomb gauge, the canonical momentum conjugate to A^i is:

$$\pi^i = \frac{\partial \mathcal{L}}{\partial \dot{A}^i} = \dot{A}^i = -E^i\,. \tag{1.247}$$

Note that $\pi^0 = \partial\mathcal{L}/\partial\dot{A}^0 = 0$ identically — A^0 is not a dynamical variable. This is a **primary constraint** in Dirac's terminology.

The transverse delta function

We cannot impose the naive commutation relation $[A^i(\vec{x}),\pi^j(\vec{x}')] = i\delta^{ij}\delta^{(3)}(\vec{x}-\vec{x}')$ because it would violate the constraint $\nabla\cdot\vec{A} = 0$. Instead, we use the **transverse delta function**:

$$\delta^{\perp}_{ij}(\vec{x}-\vec{x}') = \int \frac{d^3k}{(2\pi)^3}\left(\delta_{ij} - \frac{k_i k_j}{|\vec{k}|^2}\right) e^{i\vec{k}\cdot(\vec{x}-\vec{x}')}\,. \tag{1.248}$$

This is the Fourier transform of the transverse projector (1.246).

Box 1.27 Properties of the transverse delta function

1. Transversality:

$$\partial_i\delta^{\perp}_{ij}(\vec{x}) = \int \frac{d^3k}{(2\pi)^3} ik_i \left(\delta_{ij} - \frac{k_i k_j}{|\vec{k}|^2}\right) e^{i\vec{k}\cdot\vec{x}} = 0\,, \tag{1.249}$$

since $k_i(\delta_{ij} - k_i k_j/|\vec{k}|^2) = k_j - k_j = 0$.

2. Projection property: For any vector field $\vec{V}(\vec{x})$:

$$\int d^3x'\, \delta^{\perp}_{ij}(\vec{x}-\vec{x}')V^j(\vec{x}') = V^{\perp}_i(\vec{x})\,, \tag{1.250}$$

where $\vec{V}^{\perp}$ is the transverse part of $\vec{V}$ (satisfying $\nabla\cdot\vec{V}^{\perp}=0$).
3. Decomposition:

$$\delta^{\perp}_{ij}(\vec{x}) = \delta_{ij}\delta^{(3)}(\vec{x}) - \delta^{||}_{ij}(\vec{x})\,, \tag{1.251}$$

where $\delta^{||}_{ij}(\vec{x}) = -\partial_i\partial_j(4\pi|\vec{x}|)^{-1}$ is the longitudinal delta function.

Equal-time commutation relations

The correct equal-time commutation relations in the Coulomb gauge are:

$$[A^i(t,\vec{x}),\pi^j(t,\vec{x}')] = i\delta^{\perp}_{ij}(\vec{x}-\vec{x}')\,, \tag{1.252}$$
$$[A^i(t,\vec{x}),A^j(t,\vec{x}')] = [\pi^i(t,\vec{x}),\pi^j(t,\vec{x}')] = 0\,. \tag{1.253}$$

These commutators are consistent with the constraint $\nabla\cdot\vec{A}=0$:

$$[\nabla\cdot\vec{A}(t,\vec{x}),\pi^j(t,\vec{x}')] = i\partial_i\delta^{\perp}_{ij}(\vec{x}-\vec{x}') = 0\,. \tag{1.254}$$

Mode expansion and creation/annihilation operators

The quantized field in the Coulomb gauge is:

$$\vec{A}(x) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} \sum_{\lambda=1}^{2} \vec{\varepsilon}(\vec{k},\lambda)\left[a(\vec{k},\lambda)\,e^{-ik\cdot x} + a^{\dagger}(\vec{k},\lambda)\,e^{ik\cdot x}\right], \tag{1.255}$$

where $\omega_k = |\vec{k}|$ and $k\cdot x = \omega_k t - \vec{k}\cdot\vec{x}$.

The creation and annihilation operators satisfy:

$$[a(\vec{k},\lambda),a^{\dagger}(\vec{k}',\lambda')] = \delta_{\lambda\lambda'}\delta^{(3)}(\vec{k}-\vec{k}')\,, \tag{1.256}$$
$$[a(\vec{k},\lambda),a(\vec{k}',\lambda')] = [a^{\dagger}(\vec{k},\lambda),a^{\dagger}(\vec{k}',\lambda')] = 0\,. \tag{1.257}$$

Box 1.28 Derivation of field commutators from operator commutators

Using the mode expansion (1.255):

$$[A^i(t,\vec{x}),\pi^j(t,\vec{x}')] = [A^i, -\dot{A}^j]$$
$$= \sum_{\lambda,\lambda'}\int \frac{d^3k\, d^3k'}{(2\pi)^3\sqrt{4\omega_k\omega_{k'}}}\varepsilon^i(\vec{k},\lambda)\varepsilon^j(\vec{k}',\lambda')$$
$$\times(i\omega_{k'})[a(\vec{k},\lambda),a^\dagger(\vec{k}',\lambda')]e^{i\vec{k}\cdot\vec{x}}e^{-i\vec{k}'\cdot\vec{x}'} + \text{h.c.} \tag{1.258}$$

Using (1.256) and the polarization sum (1.246):

$$[A^i(t,\vec{x}),\pi^j(t,\vec{x}')] = i\int\frac{d^3k}{(2\pi)^3}\left(\delta^{ij}-\frac{k^ik^j}{|\vec{k}|^2}\right)e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} = i\delta^{ij}_\perp(\vec{x}-\vec{x}')\,. \tag{1.259}$$

The Hamiltonian

The Hamiltonian in the Coulomb gauge is:

$$H = \frac{1}{2}\int d^3x\,(\vec{E}^2+\vec{B}^2) = \frac{1}{2}\int d^3x\,(\dot{\vec{A}}^2+(\nabla\times\vec{A})^2)\,. \tag{1.260}$$

In terms of creation and annihilation operators (the $+\frac{1}{2}$ written below is schematic: in the continuum normalisation the zero-point piece is properly $\frac{1}{2}\delta^{(3)}(\vec{0})$ per mode, an infinite vacuum-energy density that becomes $\frac{1}{2}$ per mode only in a finite quantisation volume V, where $\delta^{(3)}(\vec{0})\to V/(2\pi)^3$):

$$H = \sum_{\lambda=1}^{2}\int d^3k\,\omega_k\left(a^\dagger(\vec{k},\lambda)a(\vec{k},\lambda)+\tfrac{1}{2}\delta^{(3)}(\vec{0})\right)\,. \tag{1.261}$$

After normal ordering (subtracting the infinite vacuum energy):

$$:H: = \sum_{\lambda=1}^{2}\int d^3k\,\omega_k\,a^\dagger(\vec{k},\lambda)a(\vec{k},\lambda)\,. \tag{1.262}$$

A photon state $|\vec{k},\lambda\rangle = a^\dagger(\vec{k},\lambda)\,|0\rangle$ has energy $\omega_k = |\vec{k}|$ and momentum $\vec{k}$.

The role of A^0 and the instantaneous Coulomb interaction

In the Coulomb gauge for the *free* field, we set $A^0 = 0$. However, in the *interacting* theory (QED), the situation is more subtle.

When matter (charged particles) is present, Gauss's law becomes:

$$\nabla \cdot \vec{E} = \rho\,, \tag{1.263}$$

where ρ is the charge density. Using $\vec{E} = -\nabla A^0 - \dot{\vec{A}}$ and the Coulomb gauge $\nabla \cdot \vec{A} = 0$:

$$-\nabla^2 A^0 = \rho \qquad \Rightarrow \qquad A^0(\vec{x}) = \int d^3x'\, \frac{\rho(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|}\,. \tag{1.264}$$

Key observations:

- A^0 is *not* an independent dynamical variable; it is completely determined by the instantaneous charge distribution ρ.
- There is no time derivative in (1.264): A^0 responds *instantaneously* to changes in ρ.
- This does not violate causality! The physical fields $\vec{E}$ and $\vec{B}$ propagate causally; the apparent non-locality is an artifact of the Coulomb gauge.

The contribution of A^0 to the Hamiltonian gives the **instantaneous Coulomb interaction**:

$$H_{\text{Coul}} = \frac{1}{2}\int d^3x\, (\nabla A^0)^2 = \frac{1}{2}\int d^3x\, d^3x'\, \frac{\rho(\vec{x})\rho(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|}\,. \tag{1.265}$$

For QED with electrons, the physical electric charge is $q_e = -e$ (with $e > 0$ the elementary charge), so the charge density is $\rho = q_e \psi^\dagger\psi = -e\,\psi^\dagger\psi$. Substituting:

$$H_{\text{Coul}} = \frac{q_e^2}{2}\int d^3x\, d^3x'\, \frac{\psi^\dagger(\vec{x})\psi(\vec{x})\psi^\dagger(\vec{x}')\psi(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|} = \frac{e^2}{2}\int d^3x\, d^3x'\, \frac{\psi^\dagger(\vec{x})\psi(\vec{x})\psi^\dagger(\vec{x}')\psi(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|}\,. \tag{1.266}$$

The Coulomb energy is even in the charge, so it does not depend on the sign of q_e — but having the charge density expressed with the correct sign is essential whenever ρ couples linearly to an external potential or to other fields (e.g. the photon field via $j^\mu A_\mu$).

This is the familiar Coulomb interaction between charges. In covariant gauges, this interaction is instead mediated by the exchange of longitudinal and timelike photons, which cancel against each other to give the same physical result.

Covariant quantization: Gupta–Bleuler formalism

In the covariant approach, we quantize all four components of A^μ and impose the Lorenz condition as a constraint on physical states.

The modified Lagrangian

We add the gauge-fixing term (1.236) to obtain:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial_\mu A^\mu)^2\,. \tag{1.267}$$

The equation of motion becomes:

$$\Box A^\mu - \left(1 - \frac{1}{\xi}\right)\partial^\mu(\partial_\nu A^\nu) = 0\,. \tag{1.268}$$

In the Feynman gauge ($\xi = 1$), this simplifies to $\Box A^\mu = 0$ for each component.

Covariant commutation relations

We impose the covariant commutation relations:

$$[A^\mu(t,\vec{x}), \dot{A}^\nu(t,\vec{x}')] = -ig^{\mu\nu}\delta^{(3)}(\vec{x}-\vec{x}')\,. \tag{1.269}$$

Note the metric $g^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$: the timelike component has the *opposite* sign!

Mode expansion with four polarizations

The covariant mode expansion is:

$$A^\mu(x) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} \sum_{\lambda=0}^{3} \varepsilon^\mu(\vec{k},\lambda)\left[a(\vec{k},\lambda)\,e^{-ik\cdot x} + a^\dagger(\vec{k},\lambda)\,e^{ik\cdot x}\right], \tag{1.270}$$

with *four* polarization vectors:

- $\lambda = 1,2$: transverse (physical) polarizations, $\varepsilon^\mu = (0,\vec{\varepsilon})$
- $\lambda = 0$: timelike (scalar) polarization, $\varepsilon^\mu = (1,0,0,0)$
- $\lambda = 3$: longitudinal polarization, $\varepsilon^\mu = (0,\hat{k})$

The polarization vectors satisfy

$$g_{\mu\nu}\,\varepsilon^\mu(\vec{k},\lambda)\,\varepsilon^\nu(\vec{k},\lambda') \;=\; \varepsilon^\mu(\vec{k},\lambda)\,\varepsilon_\mu(\vec{k},\lambda') \;=\; \eta_{\lambda\lambda'}\,, \tag{1.271}$$

where the index μ is raised/lowered with the *Minkowski* metric $g^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$, and the result $\eta_{\lambda\lambda'}$ is a 4×4 *symbol* that turns out to be numerically equal to $\mathrm{diag}(1,-1,-1,-1)$. This is not a new “polarisation metric” but a consequence of choosing the four basis polarisations aligned with the metric basis (one timelike + three spatial unit vectors): for the timelike $\lambda = 0$ one has $\varepsilon^0\varepsilon_0 = +1$ and $\vec{\varepsilon} = 0$, hence $+1$; for the transverse and longitudinal $\lambda = 1,2,3$ one has $\varepsilon^0 = 0$ and $|\vec{\varepsilon}| = 1$, hence -1.

The problem of negative-norm states

The commutation relation (1.269) implies:

$$[a(\vec{k},\lambda), a^\dagger(\vec{k}',\lambda')] = -\eta_{\lambda\lambda'}\delta^{(3)}(\vec{k}-\vec{k}')\,. \tag{1.272}$$

For the timelike polarization ($\lambda = 0$):

$$[a(\vec{k},0), a^\dagger(\vec{k}',0)] = -\delta^{(3)}(\vec{k}-\vec{k}')\,. \tag{1.273}$$

This leads to states with **negative norm**:

$$\|a^\dagger(\vec{k},0)\,|0\rangle\,\|^2 = \langle 0|\,a(\vec{k},0)a^\dagger(\vec{k},0)\,|0\rangle = -\delta^{(3)}(0) < 0\,. \tag{1.274}$$

Such states would give negative probabilities — clearly unphysical!

The Gupta–Bleuler condition

The resolution is that *not all states in the Fock space are physical.* We impose the **Gupta–Bleuler condition**: physical states $|\text{phys}\rangle$ must satisfy:

$$\partial_\mu A^{\mu(+)}(x)\,|\text{phys}\rangle = 0\,, \tag{1.275}$$

where $A^{\mu(+)}$ is the positive-frequency (annihilation) part of A^μ.

In terms of mode operators, this becomes:

$$\big[a(\vec{k},0) - a(\vec{k},3)\big]\,|\text{phys}\rangle = 0 \qquad \forall\,\vec{k}\,. \tag{1.276}$$

Box 1.29 Consequences of the Gupta–Bleuler condition

1. The vacuum is physical:
Since $a(\vec{k},\lambda)\,|0\rangle = 0$ for all λ, we have:

$$\big[a(\vec{k},0) - a(\vec{k},3)\big]\,|0\rangle = 0 - 0 = 0\,. \tag{1.277}$$

The vacuum satisfies the Gupta–Bleuler condition (1.276).

2. Transverse photon states are physical:
Consider a single transverse photon state $|\vec{k},\lambda_T\rangle = a^\dagger(\vec{k},\lambda_T)\,|0\rangle$ with $\lambda_T = 1$ or 2. We need to check:

$$\big[a(\vec{k}',0) - a(\vec{k}',3)\big]a^\dagger(\vec{k},\lambda_T)\,|0\rangle = 0\,. \tag{1.278}$$

Since $[a(\vec{k}',\lambda), a^\dagger(\vec{k},\lambda')] = -\eta_{\lambda\lambda'}\delta^{(3)}(\vec{k}'-\vec{k})$ and the transverse polarizations ($\lambda_T = 1,2$) are orthogonal to the scalar ($\lambda = 0$) and longitudinal ($\lambda = 3$) polarizations:

$$[a(\vec{k}',0), a^\dagger(\vec{k},\lambda_T)] = 0\,, \qquad [a(\vec{k}',3), a^\dagger(\vec{k},\lambda_T)] = 0\,. \tag{1.279}$$

Therefore $\left[a(\vec{k}',0)-a(\vec{k}',3)\right]$ commutes with $a^\dagger(\vec{k},\lambda_T)$, and acting on the vacuum gives zero. Multi-photon states with only transverse photons are also physical.

3. Scalar and longitudinal photons appear only in zero-norm combinations:

Consider creating a scalar photon: $|\vec{k},0\rangle = a^\dagger(\vec{k},0)\,|0\rangle$. Acting with the Gupta–Bleuler operator:

$$\begin{aligned}\left[a(\vec{k}',0)-a(\vec{k}',3)\right]a^\dagger(\vec{k},0)\,|0\rangle &= [a(\vec{k}',0),a^\dagger(\vec{k},0)]\,|0\rangle\\ &= -\delta^{(3)}(\vec{k}'-\vec{k})\,|0\rangle \neq 0\,.\end{aligned} \tag{1.280}$$

So $|\vec{k},0\rangle$ alone is *not* physical. Similarly, $|\vec{k},3\rangle = a^\dagger(\vec{k},3)\,|0\rangle$ alone is not physical.

However, the *combination*:

$$|\vec{k},S\rangle \equiv \left(a^\dagger(\vec{k},0)+a^\dagger(\vec{k},3)\right)|0\rangle \tag{1.281}$$

is physical:

$$\begin{aligned}\left[a(\vec{k}',0)-a(\vec{k}',3)\right]\left(a^\dagger(\vec{k},0)+a^\dagger(\vec{k},3)\right)|0\rangle &= \left(-\delta^{(3)}(\vec{k}'-\vec{k})+\delta^{(3)}(\vec{k}'-\vec{k})\right)|0\rangle\\ &= 0\,,\end{aligned} \tag{1.282}$$

where we used $[a(\vec{k}',0),a^\dagger(\vec{k},0)] = -\delta^{(3)}$ and $[a(\vec{k}',3),a^\dagger(\vec{k},3)] = +\delta^{(3)}$.

But this state has **zero norm**. Expanding the four-term product and noting that the cross terms $\langle 0|\,a(\vec{k},\lambda)a^\dagger(\vec{k},\lambda')\,|0\rangle$ vanish for $\lambda \neq \lambda'$ (because $[a(\vec{k},\lambda),a^\dagger(\vec{k},\lambda')] \propto \eta_{\lambda\lambda'}$ is diagonal):

$$\begin{aligned}\langle\vec{k},S|\vec{k},S\rangle &= \langle 0|\left(a(\vec{k},0)+a(\vec{k},3)\right)\left(a^\dagger(\vec{k},0)+a^\dagger(\vec{k},3)\right)|0\rangle\\ &= \langle 0|\;a(\vec{k},0)a^\dagger(\vec{k},0)\;|0\rangle \;+\; \langle 0|\;a(\vec{k},3)a^\dagger(\vec{k},3)\;|0\rangle\\ &= \langle 0|\;[a(\vec{k},0),a^\dagger(\vec{k},0)]\;|0\rangle \;+\; \langle 0|\;[a(\vec{k},3),a^\dagger(\vec{k},3)]\;|0\rangle \quad (\text{using } a\,|0\rangle = 0)\\ &= -\delta^{(3)}(\vec{0})+\delta^{(3)}(\vec{0}) = 0\,.\end{aligned} \tag{1.283}$$

The negative norm of the scalar photon *exactly cancels* the positive norm of the longitudinal photon!

4. Structure of the physical Hilbert space:

The space of states satisfying the Gupta–Bleuler condition has the structure:

$$\mathcal{H}_{\text{phys}} = \mathcal{H}_T \oplus \mathcal{H}_0\,, \tag{1.284}$$

where:

- $\mathcal{H}_T$ = states with only transverse photons (positive norm)
- $\mathcal{H}_0$ = states with scalar+longitudinal combinations (zero norm)

States in $\mathcal{H}_0$ are orthogonal to all physical states (including themselves). They form the *null space* or *spurious states*.
The **true physical Hilbert space** is the quotient:

$$\mathcal{H}_{\text{true}} = \mathcal{H}_{\text{phys}}/\mathcal{H}_0 , \tag{1.285}$$

where we identify states differing by null vectors: $|\psi\rangle \sim |\psi\rangle + |\chi\rangle$ for any $|\chi\rangle \in \mathcal{H}_0$.
On $\mathcal{H}_{\text{true}}$, the inner product is **positive-definite**, and the space is equivalent to the Fock space of transverse photons only (as in Coulomb gauge quantization).

Equivalence with Coulomb gauge

The Gupta–Bleuler formalism is physically equivalent to quantization in the Coulomb gauge:

- Both give the same S-matrix elements for physical processes.
- In Coulomb gauge, only physical degrees of freedom appear from the start.
- In covariant gauges, unphysical degrees of freedom are present but decouple from physical observables.

The advantage of covariant quantization is that Lorentz invariance is manifest at every step, making it easier to verify that physical results are Lorentz invariant.

Polarization sums in calculations

Coulomb gauge

In the Coulomb gauge, the sum over physical polarizations gives:

$$\sum_{\lambda=1}^{2} \varepsilon_i(\vec{k},\lambda)\varepsilon_j(\vec{k},\lambda) = \delta_{ij} - \frac{k_i k_j}{|\vec{k}|^2} = \delta_{ij}^{\perp}(\hat{k}) . \tag{1.286}$$

Covariant gauges

In covariant calculations, the polarization sum for external photons is:

$$\sum_{\lambda=1}^{2} \varepsilon_\mu(k,\lambda)\varepsilon_\nu^*(k,\lambda) = -g_{\mu\nu} + \frac{k_\mu \bar{k}_\nu + \bar{k}_\mu k_\nu}{k\cdot\bar{k}} , \tag{1.287}$$

where $\bar{k}^\mu$ is an auxiliary vector.

For practical calculations, we often use the simpler replacement:

$$\sum_{\lambda=1}^{2} \varepsilon_\mu(k,\lambda)\varepsilon_\nu^*(k,\lambda) \to -g_{\mu\nu}\,. \tag{1.288}$$

The **Ward identity** guarantees that the extra terms (proportional to k_μ or k_ν) do not contribute to physical amplitudes when contracted with conserved currents.

1.5 Feynman propagators

The **Feynman propagator** represents the amplitude for a particle to propagate from spacetime point y to point x. More precisely, it is the vacuum expectation value of the time-ordered product of field operators. The propagator encodes both particle and antiparticle propagation in a single object.

Scalar field propagator

For a real scalar field (where $\phi = \phi^\dagger$), the Feynman propagator is defined as the vacuum expectation value of the time-ordered product of two field operators:

$$i\Delta_F(x-y) \equiv \langle 0|T\big(\phi(x)\phi(y)\big)|0\rangle\,. \tag{1.289}$$

(For a complex scalar one would instead pair ϕ with $\phi^\dagger$; the present derivation is the real case but the structure carries over.)

The time-ordered product is:

$$T\big(\phi(x)\phi(y)\big) = \theta(x_0-y_0)\,\phi(x)\phi(y) + \theta(y_0-x_0)\,\phi(y)\phi(x)\,. \tag{1.290}$$

Derivation of the momentum-space form

Substituting the mode expansion and using the commutation relations:

$$i\Delta_F(x-y) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\left[\theta(x_0-y_0)\,e^{-ik\cdot(x-y)} + \theta(y_0-x_0)\,e^{ik\cdot(x-y)}\right]. \tag{1.291}$$

To convert this to a four-dimensional integral, we use contour integration. Consider:

$$\begin{aligned}&\frac{1}{2\omega_k}\left[\theta(x_0-y_0)\,e^{-i\omega_k(x_0-y_0)} + \theta(y_0-x_0)\,e^{i\omega_k(x_0-y_0)}\right]\\ &\quad = -\int_C \frac{dk_0}{2\pi i}\frac{e^{-ik_0(x_0-y_0)}}{(k_0-\omega_k)(k_0+\omega_k)}\end{aligned} \tag{1.292}$$

where C is one of the two contours shown in Figure 1.1.

The contour is chosen to ensure convergence of the exponential $e^{-ik_0(x_0-y_0)}$ at infinity:

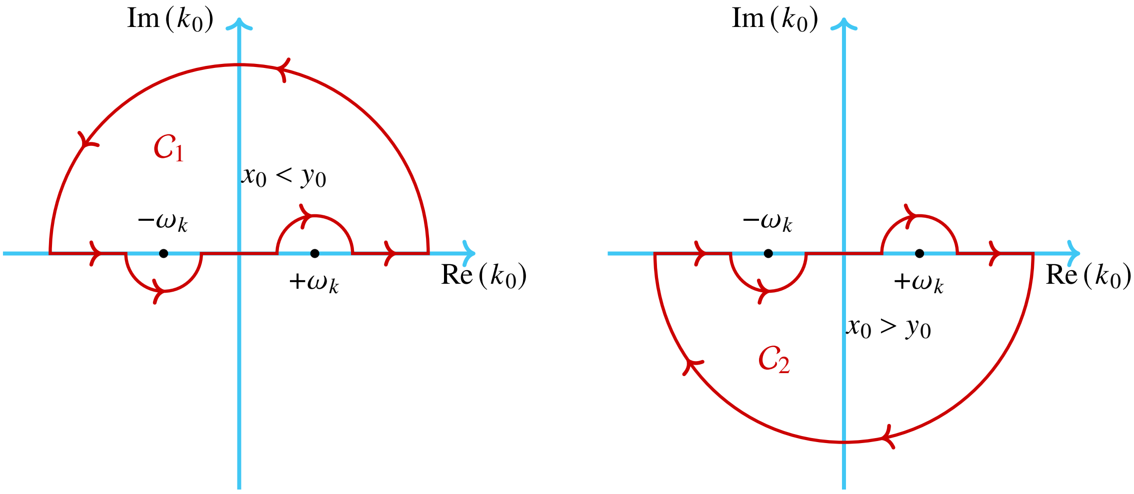

Figure 1.1. Integration contours for the scalar field propagator. *Left:* contour C_1, closed in the upper half-plane (used when $x_0 < y_0$). *Right:* contour C_2, closed in the lower half-plane (used when $x_0 > y_0$). The poles at $k_0 = \pm\omega_k$ lie on the real axis.

- If $x_0 < y_0$: we close in the *upper* half-plane (C_1), since $\lim_{k_0 \to +i\infty} e^{-ik_0(x_0-y_0)} \to 0$ when $(x_0 - y_0) < 0$. The enclosed pole is at $k_0 = -\omega_k$.
- If $x_0 > y_0$: we close in the *lower* half-plane (C_2), since $\lim_{k_0 \to -i\infty} e^{-ik_0(x_0-y_0)} \to 0$ when $(x_0 - y_0) > 0$. The enclosed pole is at $k_0 = +\omega_k$.

Evaluating the residues explicitly (note that C_2 is traversed clockwise, contributing an extra minus sign):

$$\begin{aligned}
-\int_C \frac{dk_0}{2\pi i} & \frac{e^{-ik_0(x_0-y_0)}}{(k_0-\omega_k)(k_0+\omega_k)} \\
&= \theta(x_0-y_0)\,(-1)\cdot 2\pi i \cdot \frac{-1}{2\pi i} \lim_{k_0\to\omega_k} \frac{(k_0-\omega_k)\, e^{-ik_0(x_0-y_0)}}{(k_0-\omega_k)(k_0+\omega_k)} \\
&\quad + \theta(y_0-x_0)\cdot 2\pi i \cdot \frac{-1}{2\pi i} \lim_{k_0\to-\omega_k} \frac{(k_0+\omega_k)\, e^{-ik_0(x_0-y_0)}}{(k_0-\omega_k)(k_0+\omega_k)} \\
&= \theta(x_0-y_0)\,\frac{e^{-i\omega_k(x_0-y_0)}}{2\omega_k} + \theta(y_0-x_0)\,\frac{e^{i\omega_k(x_0-y_0)}}{2\omega_k}\,,
\end{aligned} \tag{1.293}$$

which reproduces the time-ordered expression (1.291).

The $i\epsilon$ **prescription** encodes this contour choice by displacing the poles slightly off the real axis:

$$k_0^2 - \omega_k^2 + i\epsilon = 0 \quad \Longrightarrow \quad k_0 = +\omega_k - i\epsilon\,, \qquad k_0 = -\omega_k + i\epsilon\,, \tag{1.294}$$

as illustrated in Figure 1.2.
This is equivalent to adding $+i\epsilon$ to the denominator:

$$\frac{1}{k_0^2 - \omega_k^2 + i\epsilon} = \frac{1}{k^2 - m^2 + i\epsilon}\,. \tag{1.295}$$

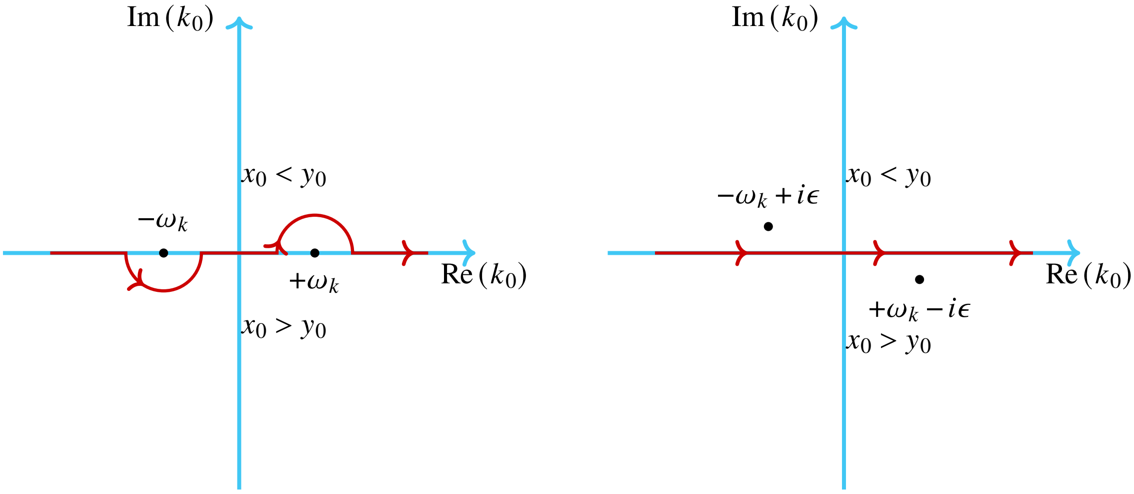

Figure 1.2. The $i\epsilon$ prescription. *Left:* original configuration with poles on the real axis and contour detouring around them via small semicircles. *Right:* equivalent configuration with poles displaced to $-\omega_k + i\epsilon$ and $+\omega_k - i\epsilon$, and integration along the entire real axis.

The final result

The Feynman propagator in momentum space is (we display it directly in the convention $i\Delta_F$, consistent with (1.289) and with Chapter 8):

$$\boxed{i\Delta_F(x-y) = \int \frac{d^4k}{(2\pi)^4} \frac{i\, e^{-ik\cdot(x-y)}}{k^2 - m^2 + i\epsilon}} \tag{1.296}$$

Important: In the Fourier representation, k_0 is no longer constrained to equal $\omega_k = \sqrt{|\vec{k}|^2 + m^2}$; it is an independent integration variable. The on-shell condition is encoded in the pole structure of the denominator.

The propagator as a Green's function

The Feynman propagator satisfies:

$$(\Box_x + m^2)\, i\Delta_F(x-y) = -i\delta^{(4)}(x-y)\,. \tag{1.297}$$

Quick proof from the Fourier representation. Acting with $\Box_x + m^2$ on the momentum-space form (1.296):

$$\begin{aligned}(\Box_x + m^2)\, i\Delta_F(x-y) &= \int \frac{d^4k}{(2\pi)^4} \frac{i\,(-k^2+m^2)\, e^{-ik\cdot(x-y)}}{k^2 - m^2 + i\epsilon} \\ &= -\int \frac{d^4k}{(2\pi)^4}\, i\, e^{-ik\cdot(x-y)} = -i\,\delta^{(4)}(x-y)\,.\end{aligned} \tag{1.298}$$

Proof via time derivatives of the T-product. A more instructive derivation works directly with the operator definition. Consider ∂_0 acting on $T(\phi(x)\phi(y))$:

$$\frac{\partial}{\partial x_0} T(\phi(x)\phi(y)) = \delta(x_0 - y_0)\,[\phi(x), \phi(y)] + T(\partial_0\phi(x)\,\phi(y))\,. \tag{1.299}$$

The commutator vanishes for equal-time scalar fields: $[\phi(t,\vec{x}), \phi(t,\vec{y})] = 0$. Taking a second time derivative:

$$\frac{\partial^2}{\partial x_0^2} T(\phi(x)\phi(y)) = \delta(x_0 - y_0)\,[\dot{\phi}(x), \phi(y)] + T(\ddot{\phi}(x)\,\phi(y))\,. \tag{1.300}$$

Since $\dot{\phi} = \pi$ for the real scalar field, the equal-time canonical commutation relation gives $[\pi(t,\vec{x}), \phi(t,\vec{y})] = -i\delta^{(3)}(\vec{x} - \vec{y})$. Moreover, spatial derivatives pass through the θ functions unaffected: $\nabla_x^2 T(\phi(x)\phi(y)) = T(\nabla_x^2\phi(x)\,\phi(y))$. Combining:

$$\Box_x T(\phi(x)\phi(y)) = -i\delta^{(4)}(x - y) + T(\Box_x\phi(x)\,\phi(y))\,. \tag{1.301}$$

Taking the vacuum expectation value and using the Klein–Gordon equation $(\Box + m^2)\phi = 0$:

$$(\Box_x + m^2)\,\langle 0|\,T(\phi(x)\phi(y))\,|0\rangle = -i\delta^{(4)}(x - y)\,, \tag{1.302}$$

which is precisely $(\Box_x + m^2)\, i\Delta_F(x - y) = -i\delta^{(4)}(x - y)$, confirming (1.297).

Dirac field propagator

The Feynman propagator for the Dirac field is:

$$iS_{F,\alpha\beta}(x - y) = \langle 0|\,T(\psi_\alpha(x)\bar{\psi}_\beta(y))\,|0\rangle\,, \tag{1.303}$$

where $\alpha, \beta = 1, \ldots, 4$ are spinor indices. In matrix notation (suppressing spinor indices):

$$iS_F(x - y) = \langle 0|\,T(\psi(x)\bar{\psi}(y))\,|0\rangle\,, \tag{1.304}$$

where $S_F(x - y)$ is a 4×4 matrix in spinor space.

Note the minus sign in the time-ordered product for fermions:

$$T(\psi(x)\bar{\psi}(y)) = \theta(x_0 - y_0)\psi(x)\bar{\psi}(y) - \theta(y_0 - x_0)\bar{\psi}(y)\psi(x)\,. \tag{1.305}$$

Physical interpretation

To understand what the Dirac propagator describes, it is useful to recall the role of ψ and $\bar{\psi}$ in creating and annihilating particles. From the mode expansion (1.142), $\psi \sim b\,u + d^\dagger v$ and $\bar{\psi} \sim b^\dagger \bar{u} + d\,\bar{v}$, so:

	Particles (e^-)	**Antiparticles** (e^+)
ψ: annihilates from $\lvert i\rangle$ creates in $\langle f\rvert$	b	$d^\dagger$
$\bar{\psi}$: creates in $\lvert i\rangle$ annihilates from $\langle f\rvert$	$b^\dagger$	d

The two terms in the time-ordered product therefore describe:

- $\theta(x_0 - y_0)\,\psi(x)\bar{\psi}(y)$: the operator $\bar{\psi}(y)$ creates an electron at y, and $\psi(x)$ destroys it at x (particle propagation forward in time)
- $\theta(y_0 - x_0)\,\bar{\psi}(y)\psi(x)$: the operator $\psi(x)$ creates a positron at x, and $\bar{\psi}(y)$ destroys it at y (antiparticle propagation forward in time, or equivalently, particle propagation backward in time)

The Feynman propagator treats both processes simultaneously in a single object.

Derivation

Step 1: Mode expansion. Using the field expansions (1.142) and keeping only the terms that give nonzero vacuum expectation values (only $b\,b^\dagger$ survives in $\langle 0|\,\psi\bar{\psi}\,|0\rangle$ and only $d\,d^\dagger$ in $\langle 0|\,\bar{\psi}\psi\,|0\rangle$):

$$\langle 0|\,\psi_\alpha(x)\bar{\psi}_\beta(y)\,|0\rangle = \sum_s \int \frac{d^3p}{(2\pi)^3}\frac{m}{E_p}\,u_\alpha(\vec{p},s)\bar{u}_\beta(\vec{p},s)\,e^{-ip\cdot(x-y)}\,, \tag{1.306}$$

$$\langle 0|\,\bar{\psi}_\beta(y)\psi_\alpha(x)\,|0\rangle = \sum_s \int \frac{d^3p}{(2\pi)^3}\frac{m}{E_p}\,v_\alpha(\vec{p},s)\bar{v}_\beta(\vec{p},s)\,e^{+ip\cdot(x-y)}\,. \tag{1.307}$$

Step 2: Spin sums. Using the completeness relations (1.150):

$$iS_F(x-y) = \int \frac{d^3p}{(2\pi)^3}\frac{m}{E_p}\left[\theta(x_0-y_0)\,e^{-ip\cdot(x-y)}\frac{\not{p}+m}{2m} - \theta(y_0-x_0)\,e^{ip\cdot(x-y)}\frac{\not{p}-m}{2m}\right]. \tag{1.308}$$

Step 3: Relation to the scalar propagator. To simplify (1.308), introduce the auxiliary functions

$$i\Delta_+(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_p}\,e^{-ip\cdot x}\,, \qquad i\Delta_-(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_p}\,e^{+ip\cdot x}\,. \tag{1.309}$$

Acting with $i\not{\partial}_x + m$ on $i\Delta_+(x-y)$ produces the factor $(\not{p}+m)/(2E_p)$ from the derivative of $e^{-ip\cdot x}$, and similarly $i\not{\partial}_x + m$ acting on $i\Delta_-(x-y)$ produces $(-\not{p}+$

$m)/(2E_p)$ from $e^{+ip\cdot x}$. One can verify that, after accounting for the factor m/E_p coming from the BD spinor normalization, equation (1.308) is equivalent to:

$$S_F(x-y) = (i\partial\!\!\!/_x + m)\left[\theta(x_0-y_0)\Delta_+(x-y) + \theta(y_0-x_0)\Delta_-(x-y)\right]. \quad (1.310)$$

The expression in brackets is precisely the scalar Feynman propagator (1.291): $i\Delta_F(x-y) = i[\theta(x_0-y_0)\Delta_+(x-y)+\theta(y_0-x_0)\Delta_-(x-y)]$. The derivative $\partial\!\!\!/_x$ can be pulled through the θ functions because the additional terms proportional to $\delta(x_0-y_0)$ cancel: the Pauli–Jordan function $\Delta(x) \equiv \Delta_+(x) - \Delta_-(x)$ is proportional to the equal-time commutator $[\phi(x), \phi^\dagger(y)]|_{x_0=y_0}$ and vanishes at equal times.

Therefore, we arrive at the fundamental relation:

$$S_F(x-y) = (i\partial\!\!\!/_x + m)\Delta_F(x-y)\,. \quad (1.311)$$

Step 4: Momentum-space form. Acting with $i\partial\!\!\!/_x + m$ on the Fourier representation of the scalar propagator (1.296), the derivative $i\partial_\mu$ brings down p_μ. Expressed in the same i-factor convention as $i\Delta_F$ for the scalar, i.e. $iS_F(x-y) \equiv \langle 0|T(\psi(x)\bar{\psi}(y))|0\rangle$:

$$\boxed{iS_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\,\frac{i(p\!\!\!/ + m)}{p^2 - m^2 + i\epsilon}\,e^{-ip\cdot(x-y)}} \quad (1.312)$$

A useful shorthand notation in momentum space is

$$iS_F(p) = \frac{i}{p\!\!\!/ - m + i\epsilon}, \quad (1.313)$$

where the denominator is understood as $p\!\!\!/ - m = \gamma^\mu p_\mu - m$, and the equivalence with (1.312) follows from rationalising: multiplying numerator and denominator by $p\!\!\!/ + m$ and using $p\!\!\!/ p\!\!\!/ = p^2$ gives $i(p\!\!\!/ + m)/(p^2 - m^2 + i\epsilon)$. Either form ($iS_F$ with the i on the right, or $S_F = 1/(p\!\!\!/ - m + i\epsilon)$ understood as "the Dirac propagator without the conventional Feynman i") is in use in the literature; in these notes we follow the chapter-8 convention and keep the i explicitly on both sides.

Photon propagator

The photon propagator describes the propagation of a virtual photon between two spacetime points. Its structure is more subtle than the scalar or Dirac propagators due to the gauge invariance of electromagnetism.

Definition

The Feynman propagator for the photon is defined as:

$$iD_{F,\mu\nu}(x-y) = \langle 0|T(A_\mu(x)A_\nu(y))\,|0\rangle\,, \quad (1.314)$$

where the time-ordered product for bosons is (no minus sign):

$$T(A_\mu(x)A_\nu(y)) = \theta(x_0-y_0)A_\mu(x)A_\nu(y)+\theta(y_0-x_0)A_\nu(y)A_\mu(x)\,. \quad (1.315)$$

The propagator $D_{F,\mu\nu}$ is a 4×4 matrix in Lorentz space (indices $\mu,\nu=0,1,2,3$), in contrast to the Dirac propagator which is a matrix in spinor space.

Gauge dependence

Unlike the scalar and Dirac propagators, the photon propagator *depends on the choice of gauge*. This is because the gauge-fixing procedure modifies the Lagrangian and hence the two-point function.

General covariant gauge (R_ξ gauge):

With the gauge-fixing term $\mathcal{L}_{\text{gf}} = -\frac{1}{2\xi}(\partial_\mu A^\mu)^2$, the propagator in momentum space is:

$$D_{F,\mu\nu}(k) = \frac{1}{k^2+i\epsilon}\left(-g_{\mu\nu}+(1-\xi)\frac{k_\mu k_\nu}{k^2}\right). \quad (1.316)$$

Special cases:

- **Feynman gauge** ($\xi=1$): $D_{F,\mu\nu}(k) = \dfrac{-g_{\mu\nu}}{k^2+i\epsilon}$ (simplest form)
- **Landau gauge** ($\xi=0$): $D_{F,\mu\nu}(k) = \dfrac{1}{k^2+i\epsilon}\left(-g_{\mu\nu}+\dfrac{k_\mu k_\nu}{k^2}\right)$ (transverse)

Derivation in Feynman gauge

Box 1.30 Derivation of the photon propagator

In the Feynman gauge ($\xi=1$), the equation of motion is $\Box A^\mu = 0$, and each component of A^μ behaves like an independent massless scalar field (with metric complications for A^0).

Step 1: Mode expansion. In covariant quantization:

$$A^\mu(x) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}}\sum_{\lambda=0}^{3}\varepsilon^\mu(\vec{k},\lambda)\left[a(\vec{k},\lambda)\,e^{-ik\cdot x}+a^\dagger(\vec{k},\lambda)\,e^{ik\cdot x}\right], \quad (1.317)$$

with $\omega_k=|\vec{k}|$ and commutation relations $[a(\vec{k},\lambda),a^\dagger(\vec{k}',\lambda')] = -\eta_{\lambda\lambda'}\delta^{(3)}(\vec{k}-\vec{k}')$.

Step 2: Polarization completeness. For the covariant polarization vectors $\varepsilon^\mu(\vec{k},\lambda)$ chosen as in Section 1.4 (one timelike, two transverse, one longitudinal), one verifies directly that the polarisation completeness relation reads

$$\sum_{\lambda=0}^{3} \eta_{\lambda\lambda}\, \varepsilon^\mu(\vec{k},\lambda)\, \varepsilon^\nu(\vec{k},\lambda) = g^{\mu\nu}, \tag{1.318}$$

i.e. the polarisation 4-vectors form an orthonormal basis (with metric $\eta_{\lambda\lambda}$) under the Minkowski inner product. The $-g^{\mu\nu}$ that will appear in the propagator comes *not* from this sum but from the commutator $[a(\vec{k},\lambda), a^\dagger(\vec{k}',\lambda')] = -\eta_{\lambda\lambda'}\delta^{(3)}(\vec{k}-\vec{k}')$ (note the minus sign), and the net minus is the source of the negative-norm states discussed in the Gupta–Bleuler analysis.

Step 3: Two-point function. Following the same steps as for the scalar field, with the photon mode expansion and commutation relations, the vacuum expectation value picks up exactly one factor of $-\eta_{\lambda\lambda}$ from the commutator, which combines with the polarisation completeness (1.318) to give

$$\begin{aligned}\langle 0|\, T(A^\mu(x) A^\nu(y))\, |0\rangle = \int \frac{d^3k}{(2\pi)^3 2\omega_k} (-g^{\mu\nu}) \\ \times \left[\theta(x_0 - y_0) e^{-ik\cdot(x-y)} + \theta(y_0 - x_0) e^{ik\cdot(x-y)}\right].\end{aligned} \tag{1.319}$$

Step 4: Fourier representation. Using the same contour arguments as for the scalar field, the Feynman propagator is:

$$iD_F^{\mu\nu}(x-y) = \int \frac{d^4k}{(2\pi)^4} \frac{-ig^{\mu\nu}}{k^2 + i\epsilon} e^{-ik\cdot(x-y)}. \tag{1.320}$$

The result

The photon Feynman propagator in position space, in the convention $iD_F^{\mu\nu}(x-y) = \langle 0|T(A^\mu(x)A^\nu(y))|0\rangle$ (consistent with $i\Delta_F$ for the scalar and iS_F for the Dirac propagator), is

$$\boxed{iD_F^{\mu\nu}(x-y) = \int \frac{d^4k}{(2\pi)^4} \frac{-i\,g^{\mu\nu}}{k^2 + i\epsilon} e^{-ik\cdot(x-y)} \qquad \text{(Feynman gauge)}} \tag{1.321}$$

In momentum space:

$$\boxed{iD_F^{\mu\nu}(k) = \frac{-i\,g^{\mu\nu}}{k^2 + i\epsilon}} \tag{1.322}$$

Properties

- **Green's function:** In Feynman gauge, $iD_F^{\mu\nu}$ satisfies

$$\Box_x\, iD_F^{\mu\nu}(x-y) = +i\, g^{\mu\nu}\, \delta^{(4)}(x-y)\,, \tag{1.323}$$

which follows directly from the Fourier representation: acting with $\Box_x$ brings down $(-k^2)$ in the integrand, $(-k^2)\cdot(-ig^{\mu\nu})/(k^2+i\epsilon) \longrightarrow ig^{\mu\nu}$ as $\epsilon \longrightarrow 0$, and the remaining Fourier integral gives $\delta^{(4)}(x-y)$; equivalently, dropping the i on both sides gives $\Box_x D_F^{\mu\nu} = +g^{\mu\nu}\delta^{(4)}(x-y)$, i.e. $D_F^{\mu\nu}$ is the Green's function of the d'Alembertian in Feynman gauge with the $+g^{\mu\nu}$ source convention.
- **Masslessness:** The pole is at $k^2 = 0$, reflecting the zero mass of the photon. There is no mass parameter in the propagator.
- **Tensor structure:** The factor $-g_{\mu\nu}$ ensures that the propagator connects all four components of A^μ. In Feynman gauge, all components propagate equally.
- **Physical observables:** Although the propagator depends on ξ, physical S-matrix elements are gauge-independent due to the Ward-Takahashi identities. The terms proportional to $k_\mu k_\nu$ vanish when contracted with conserved currents.
- **Coulomb gauge:** In the Coulomb gauge, only the transverse components propagate. The propagator is:

$$D_{ij}(\vec{k},\omega) = \frac{\delta_{ij} - \hat{k}_i\hat{k}_j}{\omega^2 - |\vec{k}|^2 + i\epsilon}\,, \tag{1.324}$$

while $D_{0\mu} = D_{\mu 0} = 0$ *after* the instantaneous Coulomb piece (the Φ field, sourced by the charge density ρ) has been separated off and folded back into the Hamiltonian as H_{Coul}, eq. (1.265): only the transverse photon then propagates dynamically. Without this separation the Coulomb gauge propagator also carries a D_{00} piece encoding instantaneous Coulomb exchange; the statement "$D_{0\mu} = 0$" is correct for the *transverse*-only propagator of the radiation modes, not for the full $A^\mu A^\nu$ correlator.

Summary: comparison of propagators

Field	Propagator (momentum space)	Spin
Real scalar	$\dfrac{i}{k^2 - m^2 + i\epsilon}$	0
Dirac fermion	$\dfrac{i(\not{k} + m)}{k^2 - m^2 + i\epsilon}$	1/2
Photon (Feynman)	$\dfrac{-ig_{\mu\nu}}{k^2 + i\epsilon}$	1

1.6 Noether's theorem and conserved currents

One of the most profound results in theoretical physics is **Noether's theorem**, which establishes a deep connection between continuous symmetries of the action and conservation laws. We follow the treatment of Coleman [2].

General statement and proof

Consider a field theory with Lagrangian density $\mathcal{L}(\phi_r, \partial_\mu \phi_r)$, where ϕ_r denotes a collection of fields (the index r labels different field components). The action is:

$$S = \int d^4x\, \mathcal{L}(\phi_r, \partial_\mu \phi_r)\,. \tag{1.325}$$

Convention. Throughout this section we use the following convention to avoid ambiguity: $\Delta\phi_r$ denotes the *characteristic* (parameter-independent) variation of the field under the symmetry, and $\delta\phi_r \equiv \epsilon\,\Delta\phi_r$ denotes the *full* infinitesimal variation including the small parameter ϵ. Thus $\delta\phi_r$ is linear in ϵ and $\Delta\phi_r$ is not.

Noether's Theorem: *If the action is invariant under a continuous transformation of the fields*

$$\phi_r(x) \longrightarrow \phi_r(x) + \epsilon\,\Delta\phi_r(x)\,, \qquad \delta\phi_r \equiv \epsilon\,\Delta\phi_r\,, \tag{1.326}$$

where ϵ is an infinitesimal parameter, then there exists a conserved current $j^\mu(x)$ satisfying

$$\partial_\mu j^\mu = 0\,, \tag{1.327}$$

and a corresponding conserved charge

$$Q = \int d^3x\, j^0(x)\,, \qquad \frac{dQ}{dt} = 0\,. \tag{1.328}$$

Proof

We present a detailed proof following the elegant approach of Coleman.

Step 1: Variation of the Lagrangian. Under the transformation (1.326), the Lagrangian varies as

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi_r}\,\delta\phi_r + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\,\partial_\mu(\delta\phi_r)\,, \tag{1.329}$$

where here and below $\delta\phi_r = \epsilon\,\Delta\phi_r$ is the full ϵ-linear variation.

Using the Euler–Lagrange equations:

$$\frac{\partial\mathcal{L}}{\partial\phi_r} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\right), \tag{1.330}$$

we can rewrite the first term:

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\right)\delta\phi_r + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\partial_\mu(\delta\phi_r)$$
$$= \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\delta\phi_r\right), \tag{1.331}$$

where we used the product rule in reverse.

Step 2: Symmetry condition. For the action to be invariant, the Lagrangian must change at most by a total divergence:

$$\delta\mathcal{L} = \epsilon\,\partial_\mu K^\mu, \tag{1.332}$$

for some four-vector K^μ. (If $\delta\mathcal{L} = 0$, we simply take $K^\mu = 0$.)

Step 3: Construction of the conserved current. Combining equations (1.331) and (1.332):

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\delta\phi_r\right) = \epsilon\,\partial_\mu K^\mu. \tag{1.333}$$

Substituting $\delta\phi_r = \epsilon\,\Delta\phi_r$ on the LHS and dividing by ϵ (which is arbitrary):

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\Delta\phi_r - K^\mu\right) = 0. \tag{1.334}$$

We identify the **Noether current**:

$$\boxed{j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\Delta\phi_r - K^\mu} \tag{1.335}$$

which satisfies $\partial_\mu j^\mu = 0$.

Step 4: The conserved charge. Define:

$$Q = \int d^3x\, j^0(x). \tag{1.336}$$

To show that Q is conserved:

$$\frac{dQ}{dt} = \int d^3x\,\partial_0 j^0 = -\int d^3x\,\nabla\cdot\vec{j} = -\oint_{\text{surface}}\vec{j}\cdot d\vec{S} = 0, \tag{1.337}$$

where we used $\partial_\mu j^\mu = 0$ and assumed that the fields fall off sufficiently fast at spatial infinity, so the surface integral vanishes.

Important remarks

1. The current j^μ is not unique: one can always add terms of the form $\partial_\nu A^{\mu\nu}$ with $A^{\mu\nu} = -A^{\nu\mu}$, since $\partial_\mu \partial_\nu A^{\mu\nu} = 0$ automatically. The charge Q, however, is unique (up to normalization).
2. **On-shell vs. off-shell:** The proof above uses the Euler–Lagrange equations (1.330), meaning it holds *on-shell* (when the fields satisfy the classical equations of motion). Off-shell, the current is not conserved in general.
   In the **quantum theory**, the situation is more subtle:
   - The Euler–Lagrange equations become *operator equations* (Heisenberg equations of motion), valid inside correlation functions up to *contact terms*.
     What are contact terms? Contact terms are delta-function contributions that arise when two spacetime points coincide. They originate from the non-commutativity of field operators at equal times. For example, when differentiating a time-ordered product:
     $$\partial_0^x T(\phi(x)\phi(y)) = T(\partial_0 \phi(x)\,\phi(y)) + \delta(x_0 - y_0)[\phi(x), \phi(y)]\,. \quad (1.338)$$
     The second term is a contact term: it is nonzero only when $x_0 = y_0$, i.e., when the points are in "contact" in time. More generally, contact terms are proportional to $\delta^{(4)}(x - x_i)$ or its derivatives.
   - Noether's theorem then implies $\partial_\mu j^\mu = 0$ as an *operator identity*, meaning:
     $$\langle 0| T\big(\partial_\mu j^\mu(x)\,\phi(x_1)\cdots\phi(x_n)\big)\,|0\rangle = \text{contact terms}\,, \quad (1.339)$$
     where the contact terms arise when x coincides with one of the x_i. These contact terms encode how the symmetry transformation acts on the fields and lead to the *Ward-Takahashi identities*, which constrain the structure of Green's functions and are crucial for proving renormalizability and gauge invariance of physical observables.
   - The conserved charge Q is a well-defined operator satisfying $[Q, H] = 0$, and generates the symmetry transformation via the commutator
     $$[Q, \phi] = -i\,\Delta\phi\,, \quad (1.340)$$
     where $\Delta\phi$ is the parameter-independent characteristic variation of ϕ under the symmetry (cf. the convention $\delta\phi = \epsilon\,\Delta\phi$ from (1.326)). Note that the infinitesimal parameter ϵ does *not* appear inside this operator identity: it is absorbed into the relation between $\delta\phi$ and $\Delta\phi$. Multiplying by ϵ recovers the finite-parameter form $[\epsilon Q, \phi] = -i\,\delta\phi$.
3. For spacetime symmetries (translations, rotations, boosts), both the coordinates and the fields transform, which requires a slight generalization of the above derivation.

Example 1: Internal U(1) symmetry and charge conservation

Consider the complex scalar field with Lagrangian:

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi \,. \tag{1.341}$$

This Lagrangian is invariant under the global $U(1)$ transformation (same convention as in the earlier complex-scalar section, where ϕ carries charge -1):

$$\phi \to e^{-i\alpha} \phi \,, \qquad \phi^\dagger \to e^{+i\alpha} \phi^\dagger \,. \tag{1.342}$$

For infinitesimal α, the characteristic variations (cf. the $\delta\phi = \epsilon \Delta\phi$ convention introduced in (1.326)) are

$$\Delta\phi = -i\phi \,, \qquad \Delta\phi^\dagger = +i\phi^\dagger \,, \tag{1.343}$$

so that with $\epsilon \equiv \alpha$ we have $\delta\phi = -i\alpha\phi$, $\delta\phi^\dagger = +i\alpha\phi^\dagger$.

Verification of symmetry:

$$\begin{aligned}
\delta\mathcal{L} &= \partial_\mu(\delta\phi^\dagger)\partial^\mu \phi + \partial_\mu \phi^\dagger \partial^\mu(\delta\phi) - m^2(\delta\phi^\dagger \phi + \phi^\dagger \delta\phi) \\
&= \partial_\mu(+i\alpha\phi^\dagger)\partial^\mu \phi + \partial_\mu \phi^\dagger \partial^\mu(-i\alpha\phi) - m^2(+i\alpha\phi^\dagger \phi - i\alpha\phi^\dagger \phi) \\
&= +i\alpha\partial_\mu \phi^\dagger \partial^\mu \phi - i\alpha\partial_\mu \phi^\dagger \partial^\mu \phi = 0 \,.
\end{aligned} \tag{1.344}$$

The Lagrangian is strictly invariant, so $K^\mu = 0$.

The Noether current: Using $\Delta\phi = -i\phi$ and $\Delta\phi^\dagger = +i\phi^\dagger$ in the Noether formula (1.335):

$$\begin{aligned}
j^\mu &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}(-i\phi) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\dagger)}(+i\phi^\dagger) \\
&= (\partial^\mu\phi^\dagger)(-i\phi) + (\partial^\mu\phi)(+i\phi^\dagger) \\
&= i\big(\phi^\dagger \partial^\mu \phi - \phi\, \partial^\mu \phi^\dagger\big) \,.
\end{aligned} \tag{1.345}$$

We can write this more symmetrically as:

$$\boxed{j^\mu = i\phi^\dagger \overleftrightarrow{\partial^\mu} \phi \equiv i(\phi^\dagger \partial^\mu \phi - (\partial^\mu \phi^\dagger)\phi)} \tag{1.346}$$

The conserved charge:

$$Q = \int d^3x\, j^0 = i \int d^3x\, (\phi^\dagger \dot\phi - \dot\phi^\dagger \phi) \,. \tag{1.347}$$

Using the mode expansion from Section 1.2:

$$Q = \int d^3k \left(a^\dagger(\vec{k})a(\vec{k}) - b^\dagger(\vec{k})b(\vec{k})\right) = N_a - N_b \,. \tag{1.348}$$

This is the **charge operator**: it counts the number of particles minus the number of antiparticles. This is the origin of electric charge conservation in QED.

Example 2: Spacetime translations and the energy-momentum tensor

Consider an infinitesimal spacetime translation:

$$x^\mu \longrightarrow x'^\mu = x^\mu + \epsilon^\mu \,. \tag{1.349}$$

The field at the new point takes the old value: $\phi'(x') = \phi(x)$. Therefore:

$$\phi'(x) = \phi(x-\epsilon) = \phi(x) - \epsilon^\mu \partial_\mu \phi(x)\,. \tag{1.350}$$

The field variation is:

$$\delta\phi = \phi'(x) - \phi(x) = -\epsilon^\mu \partial_\mu \phi\,. \tag{1.351}$$

Variation of the Lagrangian: Since $\mathcal{L}$ depends on x only through the fields,

$$\delta\mathcal{L} = -\epsilon^\mu \partial_\mu \mathcal{L} = \partial_\mu(-\epsilon^\mu \mathcal{L}) = \partial_\mu(-\epsilon^\nu g^\mu{}_\nu \mathcal{L})\,. \tag{1.352}$$

The four translation directions are independent symmetries, parametrised by the four components ϵ^ν, so the role of ϵ in our general convention (1.326) is played here by each ϵ^ν separately. The characteristic variation (with $\delta\phi = \epsilon^\nu \Delta_\nu \phi$) and the corresponding K-vector (with $K^\mu = \epsilon^\nu K^\mu{}_\nu$) are therefore

$$\Delta_\nu \phi = -\partial_\nu \phi\,, \qquad K^\mu{}_\nu = -g^\mu{}_\nu \mathcal{L} = -\delta^\mu_\nu \mathcal{L}\,. \tag{1.353}$$

The Noether current. Plugging these characteristic variations into the Noether formula (1.335),

$$j^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \Delta_\nu \phi_r - K^\mu{}_\nu = -\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \partial_\nu \phi_r + \delta^\mu_\nu \mathcal{L}\,,$$

one gets one conserved current per translation direction ν. By long-standing convention one defines the **canonical energy-momentum tensor** as $T^\mu{}_\nu \equiv -j^\mu{}_\nu$ — the overall sign is chosen so that T^{00} comes out as the (positive) energy density of the field, see the explicit example below. Raising the second index with $g^{\nu\rho}$,

$$\boxed{T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \partial^\nu \phi_r - g^{\mu\nu} \mathcal{L}} \tag{1.354}$$

The minus sign relative to the raw Noether current $j^\mu{}_\nu$ is purely a convention; the conservation statement $\partial_\mu T^{\mu\nu} = 0$ is unaffected. Equivalently, one could absorb this sign into the definition of the variation as $\delta\phi = +\epsilon^\mu \partial_\mu \phi$ (the "active" translation), in which case the Noether current of (1.335) would directly give (1.354).

The conservation law is:

$$\partial_\mu T^{\mu\nu} = 0\,. \tag{1.355}$$

The conserved charges: The four conserved quantities are the components of the four-momentum:

$$P^\nu = \int d^3x T^{0\nu} . \tag{1.356}$$

Explicitly:

- $P^0 = \int d^3x T^{00} = \int d^3x (\pi_r \dot{\phi}_r - \mathcal{L}) = \int d^3x \mathcal{H} = H$ is the **total energy**
- $P^i = \int d^3x T^{0i} = \int d^3x \pi_r \partial^i \phi_r$ is the **total momentum**

Example: real scalar field. For $\mathcal{L} = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2\phi^2)$:

$$T^{00} = \dot{\phi}\dot{\phi} - \mathcal{L} = \frac{1}{2}(\dot{\phi}^2 + (\nabla\phi)^2 + m^2\phi^2) = \mathcal{H} , \tag{1.357}$$

$$T^{0i} = \dot{\phi}\partial^i \phi . \tag{1.358}$$

Example 3: Lorentz transformations and angular momentum

Under an infinitesimal Lorentz transformation:

$$x^\mu \to x'^\mu = x^\mu + \omega^\mu{}_\nu x^\nu , \tag{1.359}$$

where $\omega^{\mu\nu} = -\omega^{\nu\mu}$ is antisymmetric (6 independent parameters: 3 rotations + 3 boosts).

The field transforms as:

$$\phi_r(x) \to \phi'_r(x) = \phi_r(\Lambda^{-1}x) + \frac{1}{2}\omega_{\mu\nu}(S^{\mu\nu})_r{}^s \phi_s(\Lambda^{-1}x) , \tag{1.360}$$

where $(S^{\mu\nu})_r{}^s$ are the generators of Lorentz transformations in the field representation (spin matrices). For a scalar field, $S^{\mu\nu} = 0$. For a Dirac spinor, $S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu]$.

The conserved current: The calculation is more involved. The result is the **angular momentum tensor**:

$$M^{\mu\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}(S^{\rho\sigma})_r{}^s \phi_s . \tag{1.361}$$

The first two terms give the **orbital angular momentum**, while the last term gives the **spin angular momentum**.

The conservation law is:

$$\partial_\mu M^{\mu\rho\sigma} = 0 . \tag{1.362}$$

The conserved charges:

$$M^{\rho\sigma} = \int d^3x\, M^{0\rho\sigma} . \tag{1.363}$$

For spatial indices $\rho, \sigma = i, j$:

$$J^k = \frac{1}{2}\epsilon^{ijk} M^{ij}\,, \tag{1.364}$$

is the angular momentum vector.
For $\rho = 0, \sigma = i$:

$$K^i = M^{0i}\,, \tag{1.365}$$

are the boost generators.

Example 4: Dirac field and the vector current

The Dirac Lagrangian:

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi\,, \tag{1.366}$$

is invariant under the global $U(1)$ transformation:

$$\psi \to e^{-i\alpha}\psi\,, \qquad \bar{\psi} \to e^{+i\alpha}\bar{\psi}\,. \tag{1.367}$$

Following the same procedure as for the complex scalar:

$$\delta\psi = -i\alpha\psi\,, \qquad \delta\bar{\psi} = +i\alpha\bar{\psi}\,. \tag{1.368}$$

The Noether current:

$$\begin{aligned} j^\mu &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}(-i\psi) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{\psi})}(+i\bar{\psi}) \\ &= \bar{\psi}(i\gamma^\mu)(-i\psi) + 0\cdot(i\bar{\psi}) \\ &= \bar{\psi}\gamma^\mu\psi\,. \end{aligned} \tag{1.369}$$

$$\boxed{j^\mu = \bar{\psi}\gamma^\mu\psi} \tag{1.370}$$

This current does *not* contain the charge factor. The QED interaction Lagrangian is (a detailed discussion of the sign conventions for the minimal coupling $\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu$ is given in Chapter 6):

$$\mathscr{L}_{\text{int}} = -q\, j^\mu A_\mu = -q\,\bar{\psi}\gamma^\mu\psi\, A_\mu\,, \tag{1.371}$$

where q is the charge of the fermion. For the electron, $q = q_e = -e$ with $e > 0$, giving:

$$\mathscr{L}_{\text{int}} = +e\,\bar{\psi}\gamma^\mu\psi\, A_\mu\,. \qquad (e > 0) \tag{1.372}$$

The conserved charge:

$$Q = \int d^3x\, j^0 = \int d^3x\, \psi^\dagger \psi\,. \tag{1.373}$$

Using the mode expansion:

$$Q = \sum_s \int d^3p \left(b^\dagger(\vec{p},s) b(\vec{p},s) - d^\dagger(\vec{p},s) d(\vec{p},s) \right) = N_{e^-} - N_{e^+}\,. \tag{1.374}$$

This is the **fermion number**: it counts particles minus antiparticles. The **electric charge** is $Q_{\text{em}} = q_e Q = -e(N_{e^-} - N_{e^+})$, which is negative for a state with more electrons than positrons, as expected.

Summary: symmetries and conservation laws

The following table summarizes the key results:

Symmetry	**Conserved current**	**Conserved charge**
Time translation	$T^{0\mu}$ (energy-momentum)	H (energy)
Space translation	T^{0i} (momentum density)	P^i (momentum)
Rotation	M^{0ij}	J^k (angular momentum)
Boost	M^{00i}	K^i (boost generator)
$U(1)$ phase	$j^\mu = \bar{\psi}\gamma^\mu\psi$	Q (fermion number)