CLAUDE OPUS 4.5

EDITED BY ADAM S. JERMYN

LECTURES ON QUANTUM FIELD THEORY AND RENORMALIZATION

Contents

Preface

These notes attempt to explain quantum field theory and renormalization in the way I wish they had been explained to me: with physical intuition at every step, honest acknowledgment of what's strange, and full derivations that don't skip the hard parts.

The central puzzle of quantum field theory is this: when you try to calculate anything beyond tree level, you get infinity. Not a large number—actual infinity. And yet QED, the quantum theory of electrons and light, agrees with experiment to better than one part in a billion. How can a theory that produces infinities make such precise predictions?

The answer is renormalization. But "renormalization" is one of those words that gets used to mean several different things, and the relationship between those things is not obvious. If you've learned renormalization from a condensed matter perspective—starting with a UV theory and integrating out high-energy modes—you may find the particle physics treatment confusing. It seems to go the other direction. These notes will clarify this confusion by showing that both approaches describe the same physics from different viewpoints.

We focus on quantum electrodynamics in 3+1 dimensions. Not because other dimensions aren't interesting, but because this is the real world, and these techniques are how actual calculations get done. We'll compute the vacuum polarization, the electron self-energy, the vertex correction, and culminate in the anomalous magnetic moment of the electron—one of the great triumphs of theoretical physics.

The mathematical prerequisites are quantum mechanics (including second quantization if you've seen it, though we'll develop it from scratch), special relativity, and comfort with contour integration. If you know what a Hamiltonian is and can do a Gaussian integral, you're ready.

I've tried to write the kind of text I wanted when I was learning this material: one where every equation has a "why" attached to it, where the physical picture is never far from the formalism, and where the reader emerges understanding not just how to compute, but what the computation means.

1

From Particles to Fields

Something is deeply wrong with quantum mechanics. Not wrong in a way you can patch over with cleverness—wrong in a way that demands we rethink what particles even are.

The problem shows up the moment you try to make quantum mechanics compatible with special relativity. Relativity allows energy to convert into mass and mass into energy. An electron moving fast enough can spontaneously create electron-positron pairs. A photon hitting a nucleus can materialize into particles. Try to pin down exactly where an electron is, and you discover you've created more electrons. The number of particles refuses to stay fixed.

Single-particle quantum mechanics, which describes *one* particle evolving in time, simply cannot handle a world where particle number changes. This isn't a technical difficulty we can patch over. It's a fundamental clash that demands a completely new conceptual framework. That framework is quantum field theory.

1.1 The Clash Between Quantum Mechanics and Relativity

Let's start with the Schrödinger equation:

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi \tag{1.1}$$

For a free particle, we usually take $\hat{H} = \hat{p}^2/2m$, giving us

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi \tag{1.2}$$

This equation treats time and space very differently. The left side has a first derivative in time; the right side has a second derivative in space. This asymmetry is fine for non-relativistic physics, where time is absolute and space is just... space. But in relativity, space and time are part of a unified spacetime, and any fundamental equation should treat them on equal footing.

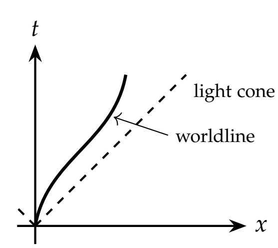


Figure 1.1: In relativity, a particle traces out a worldline in spacetime. Its trajectory must stay inside the light cone.

Here's the first problem. The non-relativistic dispersion relation is $E = p^2/2m$. But the relativistic energy-momentum relation is

$$E^2 = p^2c^2 + m^2c^4 \tag{1.3}$$

or, taking the positive root,

$$E = \sqrt{p^2c^2 + m^2c^4} \tag{1.4}$$

If we try to build a wave equation from this by making the usual substitutions $E \to i\hbar\partial_t$ and $\mathbf{p} \to -i\hbar\nabla$, we get

$$i\hbar\frac{\partial\psi}{\partial t} = \sqrt{-\hbar^2c^2\nabla^2 + m^2c^4}\,\psi \tag{1.5}$$

That square root of a differential operator is a nightmare. What does it even mean to take the square root of ∇^2? You could try to define it through a Fourier transform, but the result is a non-local operator: the time evolution at one point depends on the wave function everywhere. This is ugly and, worse, it leads to problems with causality.

1.2 *The Klein-Gordon Equation: A First Attempt*

There's a simpler approach. Instead of taking the square root, we can square both sides of $E = i\hbar\partial_t$ before equating them:

$$E^2 = p^2c^2 + m^2c^4 \quad \Rightarrow \quad -\hbar^2\frac{\partial^2\psi}{\partial t^2} = -\hbar^2c^2\nabla^2\psi + m^2c^4\psi \tag{1.6}$$

Rearranging (and setting $\hbar = c = 1$ for sanity), we get the *Klein-Gordon equation*:

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\psi = 0 \tag{1.7}$$

or, in covariant notation,

$$(\partial_\mu\partial^\mu + m^2)\psi = (\Box + m^2)\psi = 0 \tag{1.8}$$

where $\Box = \partial_t^2 - \nabla^2$ is the d'Alembertian.[1]

This is better. Space and time now appear symmetrically, as befits a relativistic equation. The equation is Lorentz invariant—it has the same form in every inertial frame.

But there are serious problems.

Problem 1: Negative Energy Solutions

The Klein-Gordon equation is second-order in time, so its general solution requires two initial conditions: $\psi(0, \mathbf{x})$ and $\partial_t\psi(0, \mathbf{x})$. This is unlike the Schrödinger equation, which only needs $\psi(0, \mathbf{x})$.

[1] We're using the "mostly minus" metric convention: $g_{\mu\nu} = \text{diag}(+1, -1, -1, -1)$.

More troubling, plane-wave solutions $\psi = e^{-iEt+i\mathbf{p}\cdot\mathbf{x}}$ exist for *both* signs of E:

$$E = \pm\sqrt{\mathbf{p}^2 + m^2} \tag{1.9}$$

We have positive and negative energy solutions. What do the negative energy solutions mean?

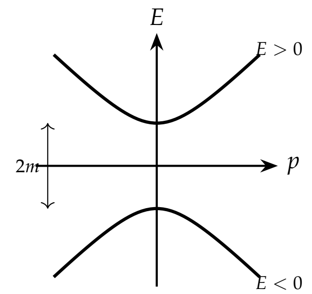


Figure 1.2: The relativistic dispersion relation $E = \pm\sqrt{p^2 + m^2}$ has two branches separated by a gap of $2m$.

In single-particle quantum mechanics, we could try to just ignore them—declare that only positive-energy states are "physical." But this doesn't work. Any localized wave packet will, over time, develop negative-energy components. You can't consistently exclude them.

Problem 2: Negative Probability

In ordinary quantum mechanics, $|\psi|^2$ is the probability density, and it's always positive. We can derive a continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \tag{1.10}$$

with $\rho = |\psi|^2 \geq 0$, ensuring that probability is conserved and non-negative everywhere.

For the Klein-Gordon equation, the analogous conserved density is[2]

$$\rho = \frac{i}{2m}\left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}\right) \tag{1.11}$$

[2] You can verify this by computing $\psi^*(\Box + m^2)\psi - \psi(\Box + m^2)\psi^*$ and showing it equals a total divergence.

This is *not* positive definite. For a negative-energy solution, $\rho < 0$. We have negative probability densities.

This is catastrophic for the interpretation of ψ as a probability amplitude. Whatever ψ is in the Klein-Gordon theory, it's not a wave function in the usual sense.

Problem 3: Causality

There's an even more fundamental problem. In non-relativistic quantum mechanics, if you localize a particle at position $\mathbf{x} = 0$ at time $t = 0$, the wave function will spread out, but it spreads *instantly* to arbitrary distances. The probability of finding the particle at a point arbitrarily far away becomes nonzero immediately.

This is fine non-relativistically—there's no speed limit. But in relativity, nothing can travel faster than light. A particle localized at the origin should not be detectable at a distance $r > ct$ until time t has passed.

Can the Klein-Gordon equation respect this? Actually, it does better than Schrödinger—the propagation speed is limited by c. But there's still a problem: the particle can propagate backward in time.

The negative-energy solutions correspond to particles moving backward in time, or equivalently (as we'll see), to antiparticles moving forward in time.

1.3 *The Dirac Equation: A Heroic Attempt*

Dirac tried a different approach. He wanted a relativistic wave equation that was *first-order* in time, like Schrödinger's equation, hoping this would fix the probability problem.

For the equation to be first-order in both space and time derivatives, Dirac proposed

$$i\frac{\partial\psi}{\partial t} = (-i\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi \tag{1.12}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ and β are some objects to be determined.

For this to be consistent with the relativistic dispersion relation, we need

$$\left(i\frac{\partial}{\partial t}\right)^2 = (-i\boldsymbol{\alpha}\cdot\nabla + \beta m)^2 \tag{1.13}$$

Working this out (and demanding $E^2 = p^2 + m^2$), Dirac found that α_i and β must satisfy

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}, \quad \{\alpha_i, \beta\} = 0, \quad \beta^2 = 1 \tag{1.14}$$

where $\{A, B\} = AB + BA$ is the anticommutator. These relations can't be satisfied by ordinary numbers—α_i and β must be *matrices*.

The smallest matrices that work are 4×4. This means ψ isn't a single component—it's a four-component object called a *spinor*.

The Dirac equation is a major success. It automatically incorporates spin-$\frac{1}{2}$—the electron's spin emerges naturally from requiring Lorentz invariance and first-order time evolution. It predicts the electron's magnetic moment to be $g = 2$ (we'll see later how QED corrects this slightly).

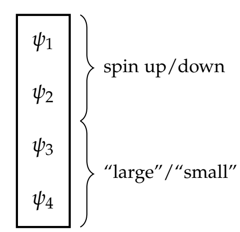


Figure 1.3: The Dirac spinor has four components. For a slowly-moving particle, two are large and two are small.

But it still has negative-energy solutions. The equation

$$(i\gamma^\mu\partial_\mu - m)\psi = 0 \tag{1.15}$$

has solutions with $E = \pm\sqrt{\mathbf{p}^2 + m^2}$.

1.4 *Dirac's Sea and the Birth of Antiparticles*

Dirac's solution to the negative-energy problem was audacious. He proposed that the vacuum isn't empty. Instead, all the negative-energy states are already filled with electrons. The Pauli exclusion principle then prevents positive-energy electrons from falling into these states.

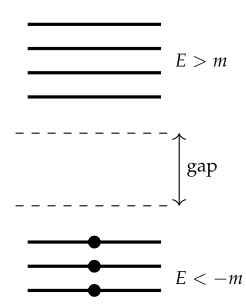


Figure 1.4: The Dirac sea: negative-energy states are filled (dots), positive-energy states are empty.

Now here's the key insight: what if one of those negative-energy electrons is excited—promoted to a positive-energy state? We'd see a positive-energy electron appear, leaving behind a "hole" in the Dirac sea. This hole would behave like a particle with positive energy and *positive charge*—the opposite charge of the electron.

Dirac initially thought this hole might be the proton. But the proton is 1836 times heavier than the electron, and the hole should have exactly the electron's mass. In 1931, Dirac boldly predicted a new particle: the *positron*, with the electron's mass but opposite charge.

In 1932, Carl Anderson discovered the positron in cosmic ray experiments, exactly as Dirac predicted. This was the first antiparticle.

1.5 Why Single-Particle Theory Cannot Work

The Dirac sea picture, while historically important, is conceptually awkward. The vacuum is infinitely full? Every atom is surrounded by an infinite sea of negative-energy electrons?

More seriously, the picture doesn't generalize to bosons. The Pauli exclusion principle only works for fermions. What stops a negative-energy photon from falling to arbitrarily negative energy?

But the real problem is more fundamental. Let me show you why no single-particle relativistic quantum theory can be consistent.

The Localization Problem

Consider trying to localize a particle to a region smaller than its Compton wavelength $\lambda_C = \hbar/mc$. By the uncertainty principle, this requires a momentum uncertainty

$$\Delta p \gtrsim \frac{\hbar}{\Delta x} \gtrsim mc \tag{1.16}$$

The corresponding energy uncertainty is

$$\Delta E \gtrsim c\Delta p \gtrsim mc^2 \tag{1.17}$$

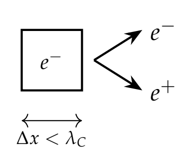


Figure 1.5: Localizing a particle to less than its Compton wavelength provides enough energy to create pairs.

But mc^2 is exactly the rest mass energy. If we have an energy uncertainty this large, we can create particle-antiparticle pairs!

This is the heart of the matter. In a relativistic theory, you cannot have a fixed number of particles. The very act of trying to localize one particle creates more particles. The one-particle Hilbert space is inconsistent—we need a Hilbert space that accommodates arbitrary numbers of particles.

A Numerical Example

Let's put in numbers. For an electron:

$$m_e = 9.11 \times 10^{-31}\,\text{kg} \tag{1.18}$$

$$\lambda_C = \frac{\hbar}{m_e c} = \frac{1.05 \times 10^{-34}\,\text{J}\cdot\text{s}}{(9.11 \times 10^{-31}\,\text{kg})(3 \times 10^{8}\,\text{m/s})} \tag{1.19}$$

$$= 3.9 \times 10^{-13}\,\text{m} = 0.0039\,\text{Å} \tag{1.20}$$

This is about 100 times smaller than the Bohr radius. So for atomic physics, we can usually ignore pair creation—the electron isn't localized tightly enough. But in high-energy physics, where particles are probed at much smaller scales, pair creation is unavoidable.

1.6 Fields as the Fundamental Objects

The resolution is to abandon particles as fundamental. Instead, the fundamental objects are *fields*.

Think about the electromagnetic field. We don't ask "how many photons are there?" when describing an electromagnetic wave—we describe the **E** and **B** fields everywhere in space. Particles (photons) emerge when we quantize the field. The number of photons is not fixed; it changes when the field interacts with matter.

The same is true for electrons. Instead of thinking of an electron as a little ball with a definite trajectory, we introduce an electron *field* $\psi(x)$ defined at every point in spacetime. When we quantize this field, electron and positron particles emerge as excitations—quanta of the field.

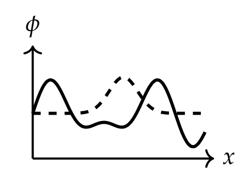


Figure 1.6: A field assigns a value to each point in space. A localized bump in the field corresponds to a particle.

This shift in perspective—from particles to fields—is the core of quantum field theory. The field is always there, defined at every point. Particles are emergent: they're localized excitations of the field. Creating a particle means exciting the field; annihilating a particle means letting the excitation dissipate.

1.7 Why Fields Fix the Problems

Let's see how fields solve the difficulties we encountered.

Negative energy solutions become antiparticles. In field theory, the negative-energy solutions of the Klein-Gordon or Dirac equation don't represent particles with negative energy. Instead, they represent *antiparticles* with positive energy. The field operator naturally creates both particles (positive frequency modes) and antiparticles (negative frequency modes), each with positive energy.

Probability becomes particle number. The quantity that was problematic as a probability density becomes, in field theory, a *charge* *density*. It can be positive or negative because particles have positive charge and antiparticles have negative charge. The total charge (not probability) is conserved.

Particle creation is built in. A field operator can create and destroy particles. The state $|0\rangle$ with no particles—the vacuum—is not boring; quantum fluctuations mean the field is always jittering. Interactions allow particle number to change naturally.

Causality is preserved. This is subtle but crucial. In quantum field theory, we impose that field operators at spacelike separation commute (or anticommute for fermions):

$$[\phi(x), \phi(y)] = 0 \quad \text{for } (x-y)^2 < 0 \tag{1.21}$$

This ensures that no measurement at x can affect a measurement at y if they're spacelike separated. Causality is built into the structure of the theory.

1.8 An Analogy: Sound in a Crystal

There's a useful analogy from condensed matter physics that illuminates the particle-field relationship.

Consider a crystal lattice—atoms arranged in a regular grid, connected by springs. The "fundamental" objects are the atoms. But when you excite the crystal, you don't track individual atoms. Instead, you describe collective oscillations: sound waves.

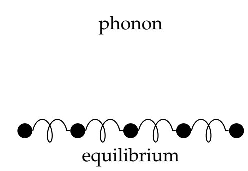


Figure 1.7: Atoms in a crystal (top: displaced, bottom: equilibrium). Collective oscillations are phonons.

When you quantize these oscillations, you get *phonons*—quanta of sound. A phonon is a particle: it has energy, momentum, and can scatter off other phonons and off defects. But there's no "phonon" sitting in the crystal when it's not vibrating. The phonon is an emergent excitation.

Moreover, the number of phonons isn't fixed. Heat the crystal, and you create more phonons. Cool it, and they disappear. The lattice displacement field $u(x)$ is always there; phonons come and go as excitations of that field.

Electrons in quantum field theory are like phonons. The electron field $\psi(x)$ is always there. Electrons and positrons are quanta—excitations of the field. They can be created and destroyed.

The analogy isn't perfect. Phonons eventually emerge from atoms, which are the "real" constituents. But in fundamental physics, there is no deeper level—fields are fundamental, not emergent. The electron field isn't made of anything; it simply is.

1.9 What We're Building Toward

In the coming chapters, we'll construct quantum field theory properly. We'll start with the simplest possible field—a spinless, non-interacting scalar—and see how quantizing it produces particles. Then we'll add interactions and develop the machinery of perturbation theory. The calculations will get messy, so we'll introduce Feynman diagrams to organize them: pictures that turn directly into numbers.

But here's where things get interesting. When we try to compute quantum corrections—the effects of virtual particles flickering in and out of existence—we'll get infinity. Not a big number, not an approximation that breaks down, but actual mathematical infinity. Understanding why this happens, and what to do about it, is really the heart of these lectures. The resolution, called renormalization, is one of the deepest ideas in twentieth-century physics.

1.10 A Note on Units

Throughout these lectures, we'll use *natural units*:

$$\hbar = c = 1 \tag{1.22}$$

This simplifies every formula enormously. In natural units:

- Energy, mass, and momentum all have the same units.
- Time and length have the same units (inverse energy).
- The electron mass is $m_e = 0.511\,\text{MeV}$.
- Lengths are measured in inverse MeV, with $1\,\text{MeV}^{-1} \approx 197\,\text{fm}$.

To convert back to SI units, restore factors of $\hbar$ and c by dimensional analysis. But for the calculations themselves, natural units are far cleaner.

1.11 Summary

We've seen that single-particle quantum mechanics clashes with special relativity at every turn. The Klein-Gordon equation gives us negative energies and probabilities that can go negative. The Dirac equation still has negative-energy solutions, leading to the strange picture of a filled sea. And most fundamentally, trying to localize a particle to less than its Compton wavelength provides enough energy to create pairs—so the very notion of "one particle" becomes meaningless.

The resolution is radical: particles aren't fundamental. Fields are. The electron field $\psi(x)$ exists everywhere in spacetime, and what we call "an electron" is just a quantized excitation of that field. Electrons and positrons come and go as the field ripples and vibrates.

This might feel like we've traded something concrete (particles) for something abstract (fields). But look at what we've gained: a framework that naturally handles creation and annihilation, respects relativity, and—as we'll see—makes predictions of astonishing precision. The strangeness is worth it.

The transition from particles to fields represents one of the great conceptual shifts in physics. It took decades to work out. Dirac's equation appeared in 1928; the positron was predicted in 1931 and discovered in 1932. But a consistent quantum field theory of electromagnetism wasn't completed until the late 1940s, with the work of Tomonaga, Schwinger, and Feynman, and Dyson's proof that their apparently different approaches were equivalent. The difficulties weren't just technical—they involved deep questions about what kind of object a quantum field is, and how to make sense of the infinities that appeared in calculations. We'll confront these infinities head-on, because understanding them is the key to understanding modern particle physics.

2

The Free Scalar Field

Why on earth would we spend a whole chapter on a theory where nothing happens? "Free" means no interactions—the field just sits there, rippling through space, never scattering, never creating anything new. Boring, right?

But you can't understand what's complicated until you understand what's simple. The free theory is the foundation on which everything else is built. And something genuinely surprising will happen: we'll start with a classical wave equation and end up with *particles*. Honest-to-goodness particles, with definite energies and momenta, that obey quantum statistics. They'll emerge from the mathematics without our putting them in by hand.

The program is this: take a classical field satisfying the Klein-Gordon equation, expand it in plane waves (like vibrations of a string), and then promote the expansion coefficients to operators. When we impose the right commutation relations, the quanta of the field turn out to be particles. By the end, "particle" will be a derived concept, emerging naturally from the quantum theory of fields.

2.1 *Classical Field Theory: A Lightning Review*

Before quantizing, we need the classical theory. Just as point particles are described by the Lagrangian $L = T - V$, fields are described by a *Lagrangian density* $\mathcal{L}$.

For a scalar field $\phi(x) = \phi(t, \mathbf{x})$, the action is

$$S = \int d^4x\, \mathcal{L}(\phi, \partial_\mu \phi) \tag{2.1}$$

The field equations come from requiring $\delta S = 0$, which gives the Euler-Lagrange equation

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \tag{2.2}$$

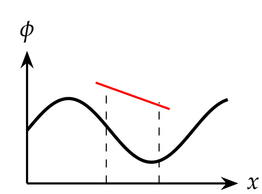


Figure 2.1: A field $\phi(x)$ and its spatial derivative. The Lagrangian depends on both.

For the free scalar field, the Lagrangian density is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2\phi^2 \tag{2.3}$$

Let's unpack this. In components:

$$\mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - \frac{1}{2}m^2\phi^2 \tag{2.4}$$

The first term is kinetic energy (involving time derivatives), the second is "gradient energy" (it costs energy to have the field vary in space), and the third is a "mass term" proportional to ϕ^2.

Why this form? The Lagrangian is the simplest Lorentz-invariant expression we can write with at most two derivatives.[1] The mass term determines how the field oscillates in time when it's uniform in space.

[1] With no derivatives, we'd have $\mathcal{L} = -V(\phi)$, which gives no dynamics. With first derivatives, we'd need a vector like $\partial_\mu\phi$, but contracting it with itself gives $(\partial_\mu\phi)^2$, which has two derivatives.

Let's derive the equation of motion. We have

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \partial^\mu\phi, \quad \frac{\partial \mathcal{L}}{\partial \phi} = -m^2\phi \tag{2.5}$$

So the Euler-Lagrange equation gives

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0 \tag{2.6}$$

This is the Klein-Gordon equation! Written out:

$$\frac{\partial^2\phi}{\partial t^2} - \nabla^2\phi + m^2\phi = 0 \tag{2.7}$$

or more compactly, $(\Box + m^2)\phi = 0$.

2.2 *Plane Wave Solutions*

The Klein-Gordon equation is linear, so we can find solutions by separation of variables. Try a plane wave:

$$\phi(x) = e^{-ipx} = e^{-ip^0t + i\mathbf{p}\cdot\mathbf{x}} \tag{2.8}$$

where $px = p_\mu x^\mu = p^0 t - \mathbf{p}\cdot\mathbf{x}$. Substituting:

$$[-(p^0)^2 + \mathbf{p}^2 + m^2]e^{-ipx} = 0 \tag{2.9}$$

This vanishes when $(p^0)^2 = \mathbf{p}^2 + m^2$, i.e., when

$$p^0 = \pm\omega_{\mathbf{p}}, \quad \text{where } \omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2} \tag{2.10}$$

We have solutions for both signs. Define the *positive frequency* solution with $p^0 = +\omega_{\mathbf{p}}$ and the *negative frequency* solution with $p^0 = -\omega_{\mathbf{p}}$.

The general solution is a superposition:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left[a(\mathbf{p}) e^{-i\omega_{\mathbf{p}} t + i\mathbf{p}\cdot\mathbf{x}} + b^*(\mathbf{p}) e^{+i\omega_{\mathbf{p}} t - i\mathbf{p}\cdot\mathbf{x}} \right] \quad (2.11)$$

The coefficients $a(\mathbf{p})$ and $b(\mathbf{p})$ are complex numbers determined by initial conditions.

Why the factor $1/\sqrt{2\omega_{\mathbf{p}}}$? This is a normalization convention that will make our formulas simpler later. It ensures that the measure $d^3p/(2\pi)^3 \cdot 1/(2\omega_{\mathbf{p}})$ is Lorentz invariant.

For a *real* scalar field (which we'll focus on), $\phi = \phi^*$, which requires $b(\mathbf{p}) = a(\mathbf{p})$. So:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left[a(\mathbf{p}) e^{-ipx} + a^*(\mathbf{p}) e^{ipx} \right] \quad (2.12)$$

where now $px = \omega_{\mathbf{p}} t - \mathbf{p}\cdot\mathbf{x}$ with the understanding that $p^0 = \omega_{\mathbf{p}}$.

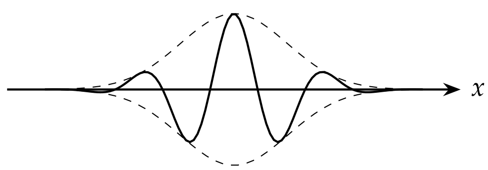


Figure 2.2: A wave packet: a superposition of plane waves creates a localized disturbance. The dashed lines show the envelope.

2.3 *The Hamiltonian*

To quantize, we'll need the Hamiltonian. The conjugate momentum to ϕ is

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi} \quad (2.13)$$

The Hamiltonian density is

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L} = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2 \quad (2.14)$$

All three terms are positive. This makes sense: the Hamiltonian is the energy density, which should be positive for a stable system.

The total Hamiltonian is

$$H = \int d^3x\, \mathcal{H} = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2 \right] \quad (2.15)$$

2.4 *Canonical Quantization*

Now we quantize. The prescription is the same as in quantum mechanics: promote ϕ and π to operators and impose canonical commutation relations.

At equal times:

$$[\phi(t,\mathbf{x}),\phi(t,\mathbf{y})] = 0 \tag{2.16}$$

$$[\pi(t,\mathbf{x}),\pi(t,\mathbf{y})] = 0 \tag{2.17}$$

$$[\phi(t,\mathbf{x}),\pi(t,\mathbf{y})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y}) \tag{2.18}$$

The delta function appears because $\phi(t,\mathbf{x})$ and $\pi(t,\mathbf{y})$ at different spatial points are independent degrees of freedom—like q_i and p_j for different particles in classical mechanics, where $[q_i, p_j] = i\delta_{ij}$.

What happens to the mode expansion? The coefficients $a(\mathbf{p})$ and $a^*(\mathbf{p})$ become operators $a_\mathbf{p}$ and $a^\dagger_\mathbf{p}$:[2]

[2] We'll write $a_\mathbf{p}$ rather than $a(\mathbf{p})$ to emphasize that it's now an operator.

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\left[a_\mathbf{p}e^{-ipx} + a^\dagger_\mathbf{p}e^{ipx}\right] \tag{2.19}$$

The commutation relations (2.18) translate into commutation relations for a and $a^\dagger$. After some algebra (see below), we find:

$$[a_\mathbf{p}, a^\dagger_\mathbf{q}] = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q}) \tag{2.20}$$

and

$$[a_\mathbf{p}, a_\mathbf{q}] = [a^\dagger_\mathbf{p}, a^\dagger_\mathbf{q}] = 0 \tag{2.21}$$

Deriving the Commutation Relations

Let me show how (2.20) follows from (2.18). The key is to invert the mode expansion.

From (2.12), taking $\pi = \dot{\phi}$:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\left[a_\mathbf{p}e^{-ipx} + a^\dagger_\mathbf{p}e^{ipx}\right] \tag{2.22}$$

$$\pi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{(-i\omega_\mathbf{p})}{\sqrt{2\omega_\mathbf{p}}}\left[a_\mathbf{p}e^{-ipx} - a^\dagger_\mathbf{p}e^{ipx}\right] \tag{2.23}$$

We can solve for $a_\mathbf{p}$ by taking appropriate Fourier transforms. Using $\int d^3x\, e^{i(\mathbf{p}-\mathbf{q})\cdot\mathbf{x}} = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})$, we find

$$a_\mathbf{p} = \int d^3x\, e^{ipx}\left[\sqrt{\frac{\omega_\mathbf{p}}{2}}\phi(x) + \frac{i}{\sqrt{2\omega_\mathbf{p}}}\pi(x)\right]\Bigg|_{t=0} \tag{2.24}$$

(evaluated at any fixed time, say $t = 0$).

Now compute $[a_\mathbf{p}, a^\dagger_\mathbf{q}]$ using this expression. The $[\phi,\phi]$ and $[\pi,\pi]$ terms vanish. The cross terms give:

$$[a_\mathbf{p}, a^\dagger_\mathbf{q}] = \int d^3x\, d^3y\, e^{i\mathbf{p}\cdot\mathbf{x}}e^{-i\mathbf{q}\cdot\mathbf{y}}\frac{i}{2}\left(-\frac{\omega_\mathbf{q}}{\omega_\mathbf{p}} + \frac{\omega_\mathbf{p}}{\omega_\mathbf{q}}\right)^{1/2}\cdot[\phi(\mathbf{x}),\pi(\mathbf{y})] \tag{2.25}$$

Wait—that's getting complicated. Let me just give the result: using $[\phi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y})$ and carefully tracking the factors, one obtains

$$[a_{\mathbf{p}}, a^\dagger_{\mathbf{q}}] = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q}) \tag{2.26}$$

The key point is that these are exactly the commutation relations of the quantum harmonic oscillator! For each momentum $\mathbf{p}$, we have a raising operator $a^\dagger_{\mathbf{p}}$ and a lowering operator $a_{\mathbf{p}}$.

2.5 *The Vacuum and Particle States*

With the algebra of harmonic oscillators, we know what to do. Define the *vacuum* $|0\rangle$ as the state annihilated by all lowering operators:

$$a_{\mathbf{p}}|0\rangle = 0 \quad \text{for all } \mathbf{p} \tag{2.27}$$

The vacuum is the state with no particles—the lowest-energy state of the field.

A *one-particle state* is created by acting with $a^\dagger$:

$$|\mathbf{p}\rangle = \sqrt{2\omega_{\mathbf{p}}}\, a^\dagger_{\mathbf{p}}|0\rangle \tag{2.28}$$

The normalization factor ensures $\langle\mathbf{p}|\mathbf{q}\rangle = (2\pi)^3 2\omega_{\mathbf{p}}\,\delta^{(3)}(\mathbf{p}-\mathbf{q})$, which is Lorentz invariant.

Multi-particle states are created by acting multiple times:

$$|\mathbf{p}_1, \mathbf{p}_2\rangle \propto a^\dagger_{\mathbf{p}_1} a^\dagger_{\mathbf{p}_2}|0\rangle \tag{2.29}$$

Since $[a^\dagger_{\mathbf{p}}, a^\dagger_{\mathbf{q}}] = 0$, these states are symmetric under exchange:

$$|\mathbf{p}, \mathbf{q}\rangle = |\mathbf{q}, \mathbf{p}\rangle \tag{2.30}$$

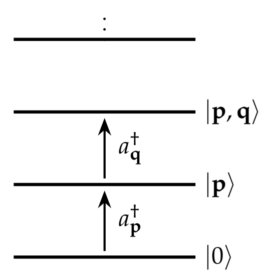


Figure 2.3: The Fock space: the vacuum, one-particle states, two-particle states, etc.

The particles are *bosons*.

Notice what happened. We started with a classical field satisfying a wave equation. We quantized it. And out pop particles—quanta of the field—that automatically obey Bose-Einstein statistics. The connection between spin and statistics (bosons are spin-0, spin-1, spin-2, ...; fermions are spin-1/2, spin-3/2, ...) emerges naturally from the structure of quantum field theory.

2.6 *The Hamiltonian in Terms of Creation/Annihilation Operators*

Let's express the Hamiltonian in terms of a and $a^\dagger$. This will show that particles carry energy, as they should.

Substituting the mode expansion into $H = \int d^3x\, \mathcal{H}$ and grinding through the algebra (which is straightforward but tedious), we get:

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} \left(a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + \frac{1}{2}[a_{\mathbf{p}}, a_{\mathbf{p}}^\dagger] \right) \tag{2.31}$$

The first term is $\omega_{\mathbf{p}} \cdot a_{\mathbf{p}}^\dagger a_{\mathbf{p}}$—the energy $\omega_{\mathbf{p}}$ times the number of particles with momentum $\mathbf{p}$. This is exactly what we want.

The second term is problematic:

$$\frac{1}{2} \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} \cdot (2\pi)^3 \delta^{(3)}(0) = \frac{1}{2} \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} \cdot V \tag{2.32}$$

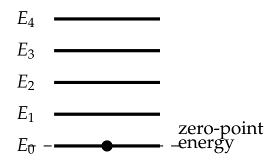


Figure 2.4: Each oscillator contributes $\frac{1}{2}\omega$ to the vacuum energy.

where $V = (2\pi)^3 \delta^{(3)}(0)$ is the (infinite) volume of space. This is the *zero-point energy*—the vacuum has infinite energy!

Wait—the vacuum has infinite energy? What does that even mean? Can you feel this energy? Can you extract it? And if you can't, does it matter?

This infinity is our first encounter with the divergences of quantum field theory. It's relatively benign, and we can deal with it by *normal ordering*—but don't let that reassure you too much. We're declaring by fiat that the vacuum energy is zero, sweeping an infinity under the rug. It works, but it should make you uncomfortable. Deeper infinities are coming.

For now, define the normal-ordered Hamiltonian by putting all $a^\dagger$s to the left of all as:

$$: H := \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}}\, a_{\mathbf{p}}^\dagger a_{\mathbf{p}} \tag{2.33}$$

This gives $\langle 0| : H : |0\rangle = 0$. We're declaring that the vacuum has zero energy, and measuring all other energies relative to it.

Is this cheating? In a sense, yes—we're sweeping an infinity under the rug. But since we only ever measure energy *differences*, defining the vacuum energy to be zero is physically sensible.[3]

With normal ordering, the Hamiltonian is

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}}\, a_{\mathbf{p}}^\dagger a_{\mathbf{p}} \tag{2.34}$$

and a one-particle state $|\mathbf{p}\rangle$ has energy

$$H|\mathbf{p}\rangle = \omega_{\mathbf{p}}|\mathbf{p}\rangle = \sqrt{\mathbf{p}^2 + m^2}|\mathbf{p}\rangle \tag{2.35}$$

This is exactly the relativistic energy-momentum relation! The quanta of the field are relativistic particles of mass m.

[3] There's a subtlety here related to gravity. The vacuum energy acts as a cosmological constant. The fact that the observed cosmological constant is tiny (but nonzero) while QFT predicts an enormous vacuum energy is one of the great unsolved problems of physics.

2.7 Momentum and Charge

Similarly, the momentum operator is

$$\mathbf{P} = -\int d^3x\, \pi \nabla \phi = \int \frac{d^3p}{(2\pi)^3} \mathbf{p}\, a^\dagger_{\mathbf{p}} a_{\mathbf{p}} \tag{2.36}$$

A one-particle state $|\mathbf{p}\rangle$ has momentum $\mathbf{p}$, as expected.

For a *complex* scalar field $\phi \neq \phi^*$, there's also a conserved charge. The Lagrangian

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi \tag{2.37}$$

has a $U(1)$ symmetry: $\phi \to e^{i\alpha}\phi$, $\phi^* \to e^{-i\alpha}\phi^*$. By Noether's theorem, this symmetry implies a conserved current and charge.

The mode expansion now has two sets of operators:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left[a_{\mathbf{p}} e^{-ipx} + b^\dagger_{\mathbf{p}} e^{ipx} \right] \tag{2.38}$$

Here $a^\dagger_{\mathbf{p}}$ creates a particle and $b^\dagger_{\mathbf{p}}$ creates an *antiparticle*. Both have positive energy $\omega_{\mathbf{p}}$, but they carry opposite charge. The conserved charge is

$$Q = \int \frac{d^3p}{(2\pi)^3} \left(a^\dagger_{\mathbf{p}} a_{\mathbf{p}} - b^\dagger_{\mathbf{p}} b_{\mathbf{p}} \right) \tag{2.39}$$

Particles contribute $+1$, antiparticles contribute -1. The vacuum has $Q = 0$.

2.8 The Feynman Propagator

We now introduce one of the most important objects in quantum field theory: the *propagator*. Physically, it describes how a disturbance in the field propagates from one point to another.

Consider the vacuum expectation value of a time-ordered product:

$$\Delta_F(x - y) = \langle 0 | T \phi(x) \phi(y) | 0 \rangle \tag{2.40}$$

where the time-ordering symbol T puts the operator with the later time to the left:

$$T\phi(x)\phi(y) = \begin{cases} \phi(x)\phi(y) & \text{if } x^0 > y^0 \\ \phi(y)\phi(x) & \text{if } y^0 > x^0 \end{cases} \tag{2.41}$$

Why time ordering? Consider what $\langle 0|\phi(x)\phi(y)|0\rangle$ means. Acting on the vacuum, $\phi(y)$ can create a particle (via the $a^\dagger$ part). Then $\phi(x)$ can annihilate it (via the a part). This represents a particle being created at y and propagating to x.

But if $y^0 > x^0$, the $\phi(y)$ comes later. Then we need $\phi(x)$ to create and $\phi(y)$ to annihilate. Time ordering handles this automatically.

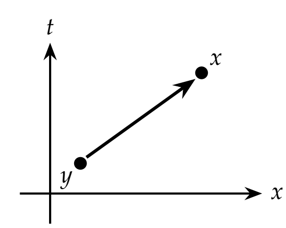


Figure 2.5: The propagator describes a particle created at y and absorbed at x (if $x^0 > y^0$).

Let me compute Δ_F for the case $x^0 > y^0$:

$$\langle 0|\phi(x)\phi(y)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}\cdot 2\omega_{\mathbf{q}}}} \times \langle 0|(a_{\mathbf{p}}e^{-ipx} + a_{\mathbf{p}}^\dagger e^{ipx})(a_{\mathbf{q}}e^{-iqy} + a_{\mathbf{q}}^\dagger e^{iqy})|0\rangle \tag{2.42}$$

The only term that survives is $\langle 0|a_{\mathbf{p}}a_{\mathbf{q}}^\dagger|0\rangle = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})$, giving:

$$\langle 0|\phi(x)\phi(y)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}e^{-ip(x-y)} \tag{2.43}$$

where $p^0 = \omega_{\mathbf{p}} > 0$.

Similarly, for $y^0 > x^0$:

$$\langle 0|\phi(y)\phi(x)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}e^{-ip(y-x)} = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}e^{+ip(x-y)} \tag{2.44}$$

The Feynman propagator combines these. There's a compact formula that captures both cases:

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2+i\epsilon}e^{-ip(x-y)} \tag{2.45}$$

where $\epsilon > 0$ is infinitesimal.

Where Does the $i\epsilon$ Come From?

This is worth understanding in detail. The integral in (2.45) is over all four components of p. The denominator $p^2 - m^2 = (p^0)^2 - \mathbf{p}^2 - m^2$ vanishes when

$$p^0 = \pm\omega_{\mathbf{p}} \tag{2.46}$$

So the p^0 integral has poles on the real axis. We need to specify how to go around them.

The $i\epsilon$ prescription shifts the poles slightly:

$$p^2 - m^2 + i\epsilon = (p^0)^2 - \omega_{\mathbf{p}}^2 + i\epsilon = (p^0 - \omega_{\mathbf{p}} + i\epsilon')(p^0 + \omega_{\mathbf{p}} - i\epsilon') \tag{2.47}$$

where $\epsilon' = \epsilon/(2\omega_{\mathbf{p}})$.

The pole at $p^0 = +\omega_{\mathbf{p}}$ is shifted slightly below the real axis. The pole at $p^0 = -\omega_{\mathbf{p}}$ is shifted slightly above.

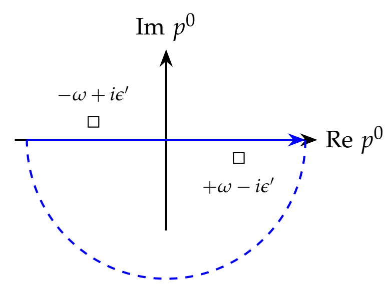


Figure 2.6: The $i\epsilon$ prescription pushes poles off the real axis. For $x^0 > y^0$, we close in the lower half-plane, picking up the $+\omega$ pole.

Now do the p^0 integral by contour integration. For $x^0 > y^0$, the factor $e^{-ip^0(x^0-y^0)}$ decays in the lower half-plane (since $x^0 - y^0 > 0$). So we close the contour below, picking up the pole at $p^0 = +\omega_{\mathbf{p}} - i\epsilon'$:

$$\oint \frac{dp^0}{2\pi}\frac{i\,e^{-ip^0(x^0-y^0)}}{(p^0-\omega+i\epsilon')(p^0+\omega-i\epsilon')} = \frac{e^{-i\omega(x^0-y^0)}}{2\omega} \tag{2.48}$$

This gives exactly $\langle 0|\phi(x)\phi(y)|0\rangle$.

For $y^0 > x^0$, we close above, picking up the other pole, and get $\langle 0|\phi(y)\phi(x)|0\rangle$.

The $i\epsilon$ prescription encodes causality: positive-frequency modes propagate forward in time, negative-frequency modes propagate backward. Or equivalently: particles propagate forward, antiparticles propagate backward.

2.9 Physical Interpretation of the Propagator

What does the propagator mean physically? Consider the amplitude for a particle to be created at y and detected at x:

$$\langle 0|\phi(x)|y\rangle = \langle 0|\phi(x)\phi(y)|0\rangle \cdot \text{(creation factor)} \tag{2.49}$$

The propagator $\Delta_F(x-y)$ is essentially this amplitude. It tells us how the quantum field correlates disturbances at different spacetime points.

In momentum space, the propagator is

$$\tilde{\Delta}_F(p) = \frac{i}{p^2 - m^2 + i\epsilon} \tag{2.50}$$

This has a pole when $p^2 = m^2$—exactly when p is the four-momentum of a particle on its mass shell.

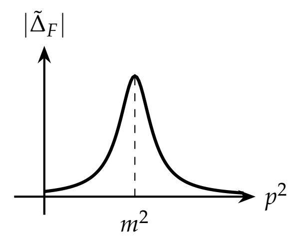


Figure 2.7: The propagator peaks sharply at $p^2 = m^2$.

Think about what this means. If you Fourier transform to see the contribution of a particular momentum p, the propagator is large when $p^2 \approx m^2$. The particle "wants" to be on-shell.

For virtual particles in Feynman diagrams, $p^2 \neq m^2$ in general. The propagator is still well-defined but smaller. The further off-shell the virtual particle, the more suppressed its contribution. This is why high-energy virtual particles contribute less than low-energy ones—a fact that will be crucial when we discuss renormalization.

2.10 The Klein-Gordon Equation Revisited

Let's verify that the propagator satisfies the right equation. Acting with the Klein-Gordon operator:

$$(\Box_x + m^2)\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i(-p^2+m^2)}{p^2 - m^2 + i\epsilon} e^{-ip(x-y)} \tag{2.51}$$

$$= -i \int \frac{d^4p}{(2\pi)^4} e^{-ip(x-y)} \tag{2.52}$$

$$= -i\delta^{(4)}(x-y) \tag{2.53}$$

So the Feynman propagator is a Green's function for the Klein-Gordon equation:

$$(\Box + m^2)\Delta_F(x-y) = -i\delta^{(4)}(x-y) \tag{2.54}$$

This makes physical sense. The propagator describes how a field disturbance—a delta-function source at y—spreads through spacetime.

2.11 Putting in Numbers

Let's build some intuition with numbers. Consider a scalar particle of mass $m = 100$ MeV (roughly a pion). Its Compton wavelength is

$$\lambda_C = \frac{1}{m} = \frac{1}{100\,\text{MeV}} \approx 2 \times 10^{-15}\,\text{m} = 2\,\text{fm} \tag{2.55}$$

This is about the size of a proton.

The propagator falls off exponentially for spacelike separations $|x - y| \gg 1/m$. So a pion's influence is felt only within about 2 fm—which is why the strong nuclear force (mediated by pions) has short range.

Compare to the photon, which is massless. Its propagator falls off as $1/|x - y|^2$ for large spacelike separations—it has infinite range. This is why electromagnetism is a long-range force.

2.12 What We've Built

Look at what's happened. We started with waves satisfying a classical equation, and out popped particles. Not because we put particles in—we didn't. We just quantized the field, imposing commutation relations on the mode amplitudes, and the mathematics gave us creation and annihilation operators. The operator $a^\dagger_{\mathbf{p}}$ adds a quantum of momentum $\mathbf{p}$ and energy $\omega_{\mathbf{p}} = \sqrt{p^2 + m^2}$. That quantum is what we call a particle.

The particles automatically obey Bose statistics—you can put as many as you like in the same state. The Hamiltonian counts them, weighted by their energies. The propagator describes how disturbances travel through the field, carrying correlations from one point to another.

But we haven't introduced any interactions yet. The particles fly freely through space without ever colliding or scattering. To describe real physics—the physics of electrons bouncing off each other, of photons being absorbed and emitted—we need interactions. That's where things get interesting, and that's the subject of the next chapter.

The quantization procedure we've followed—starting with a classical field, expanding in modes, promoting to operators—is called canonical quantization. *It's the oldest and most intuitive approach. There's another approach, the* path integral, *which we'll glimpse later. Both give the same physics, but they make different calculations easy. The path integral is particularly natural for understanding gauge theories and non-perturbative effects. For our purposes, canonical quantization provides the clearest conceptual foundation for understanding where particles come from and what renormalization is doing.*

3
Interactions and Perturbation Theory

A universe of free particles would be boring beyond description. Nothing would happen—ever. Particles would fly past each other, oblivious, never scattering, never combining, never creating anything new. The fact that things *happen*—that electrons repel, that light gets absorbed, that particles annihilate into other particles—is what makes physics worth studying.

In the last chapter we built a quantum theory of a free field. Particles emerged as quanta, propagators described their motion through spacetime, but the physics was trivial. Now we add interactions, and everything changes.

But here's the problem: interacting quantum field theories are essentially impossible to solve exactly. The free theory was tractable because it's just a collection of harmonic oscillators, each with its exact solution. Add interactions, and the oscillators are coupled in complicated, nonlinear ways. Exact solutions don't exist.

Our salvation is perturbation theory. If the interaction is "small"—if there's a small coupling constant λ or e or g—we can expand in powers of that coupling. The leading term is the free theory. Corrections come order by order, each calculable from the previous ones.

This chapter develops the machinery: the interaction picture, time evolution, Wick's theorem, and the structure of the perturbation series. In the next chapter, we'll see how this machinery gets encoded in Feynman diagrams.

3.1 Adding Interactions to the Lagrangian

The simplest interacting theory is ϕ^4 theory: a real scalar field with a quartic self-interaction. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 \tag{3.1}$$

The first two terms are the free Klein-Gordon Lagrangian. The last term is new.

Why ϕ^4 rather than, say, ϕ^3? Several reasons:

- A ϕ^3 term would make the potential unbounded below—there'd be no ground state.
- The $1/4!$ is a convention that simplifies later formulas.
- ϕ^4 is the simplest interaction that's "renormalizable" in four dimensions (more on this later).

The coupling constant λ is dimensionless (in four spacetime dimensions with $\hbar = c = 1$). We'll assume $\lambda \ll 1$ so perturbation theory makes sense.[1]

The equation of motion is now

$$(\Box + m^2)\phi = -\frac{\lambda}{3!}\phi^3 \tag{3.2}$$

The right-hand side couples different modes—the equation is nonlinear, and we can't solve it exactly.

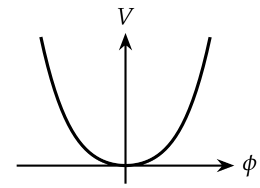


Figure 3.1: The ϕ^4 potential: $V = \frac{m^2}{2}\phi^2 + \frac{\lambda}{4!}\phi^4$. The quartic term keeps it bounded below.

[1] What counts as "small" here? The expansion parameter turns out to be roughly $\lambda/(16\pi^2)$, so even $\lambda \sim 1$ can give a reasonable expansion.

3.2 *The Interaction Picture*

To do perturbation theory systematically, we split the Hamiltonian:

$$H = H_0 + H_{\text{int}} \tag{3.3}$$

where H_0 is the free Hamiltonian (which we can solve exactly) and H_{int} is the interaction (which we treat perturbatively).

For ϕ^4 theory:

$$H_0 = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2\right] \tag{3.4}$$

$$H_{\text{int}} = \int d^3x \, \frac{\lambda}{4!}\phi^4 \tag{3.5}$$

In the *Schrödinger picture*, states evolve and operators are fixed. In the *Heisenberg picture*, operators evolve and states are fixed. The *interaction picture* is a hybrid: states evolve due to H_{int}, operators evolve due to H_0.

Define interaction-picture operators:

$$\phi_I(t, \mathbf{x}) = e^{iH_0 t}\phi_S(\mathbf{x})e^{-iH_0 t} \tag{3.6}$$

where ϕ_S is the Schrödinger-picture operator. The subscript I for "interaction picture" is often dropped when context is clear.

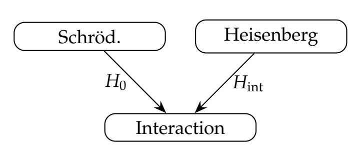


Figure 3.2: The interaction picture splits evolution between operators (H_0) and states (H_{int}).

Since ϕ_I evolves with the *free* Hamiltonian H_0, it satisfies the free field equation and has the same mode expansion as in Chapter 2:

$$\phi_I(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\left[a_{\mathbf{p}}e^{-ipx} + a^\dagger_{\mathbf{p}}e^{ipx}\right] \tag{3.7}$$

This is why the interaction picture is so useful: the operators are simple (free fields), and all the complexity of interactions goes into the states.

3.3 Time Evolution in the Interaction Picture

How do states evolve in the interaction picture? Define the interaction-picture state:

$$|\psi_I(t)\rangle = e^{iH_0 t}|\psi_S(t)\rangle \tag{3.8}$$

where $|\psi_S(t)\rangle = e^{-iHt}|\psi_S(0)\rangle$ is the Schrödinger-picture state.

The equation of motion for $|\psi_I(t)\rangle$ is

$$i\frac{d}{dt}|\psi_I(t)\rangle = H_I(t)|\psi_I(t)\rangle \tag{3.9}$$

where $H_I(t) = e^{iH_0 t}H_{\text{int}}e^{-iH_0 t}$ is the interaction Hamiltonian in the interaction picture.

The solution is

$$|\psi_I(t)\rangle = U(t, t_0)|\psi_I(t_0)\rangle \tag{3.10}$$

where the time-evolution operator U satisfies

$$i\frac{\partial U}{\partial t} = H_I(t)U(t, t_0), \quad U(t_0, t_0) = 1 \tag{3.11}$$

If H_I were time-independent, we'd have $U = e^{-iH_I(t-t_0)}$. But $H_I(t)$ depends on time (because the free-field operators do), so we need to be more careful.

Why does time ordering appear? Think about what we're computing: the state at time t given the state at t_0. If the Hamiltonian were constant, we'd just exponentiate: $U = e^{-iH(t-t_0)}$. But $H_I(t)$ changes with time—the free-field operators inside it are evolving—so we have to be careful about the order in which things happen. An interaction at time t_2 happens *before* an interaction at $t_1 > t_2$, so $H_I(t_2)$ should act on the state first.

The formal solution is the *time-ordered exponential*:

$$U(t, t_0) = T\exp\left(-i\int_{t_0}^{t} dt'\, H_I(t')\right) \tag{3.12}$$

where T is the time-ordering symbol that puts later operators to the left. Expanded:

$$\begin{aligned} U(t, t_0) = 1 - i\int_{t_0}^{t} dt_1\, H_I(t_1) \\ + (-i)^2\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, H_I(t_1)H_I(t_2) + \cdots \end{aligned} \tag{3.13}$$

The nested integrals automatically enforce the time ordering.

3.4 The S-Matrix

We're interested in scattering processes: particles come in from the distant past, interact, and fly off to the distant future. The object that encodes this is the *S-matrix*.

Define the S-matrix as the limit:

$$S = \lim_{\substack{t \to +\infty \\ t_0 \to -\infty}} U(t, t_0) \tag{3.14}$$

In this limit, we can write

$$S = T \exp\left(-i \int_{-\infty}^{+\infty} dt\, H_I(t)\right) = T \exp\left(-i \int d^4x\, \mathcal{H}_I(x)\right) \tag{3.15}$$

where $\mathcal{H}_I = \frac{\lambda}{4!}\phi_I^4$ is the interaction Hamiltonian density.

The S-matrix element between an initial state $|i\rangle$ and a final state $|f\rangle$ is $\langle f|S|i\rangle$. This gives the probability amplitude for the transition $|i\rangle \to |f\rangle$.

In perturbation theory, we expand:

$$S = 1 + (-i)\int d^4x\, \mathcal{H}_I(x) + \frac{(-i)^2}{2!}\int d^4x\, d^4y\, T[\mathcal{H}_I(x)\mathcal{H}_I(y)] + \cdots \tag{3.16}$$

$$= 1 + S^{(1)} + S^{(2)} + \cdots \tag{3.17}$$

where $S^{(n)}$ is the nth-order term, proportional to λ^n.

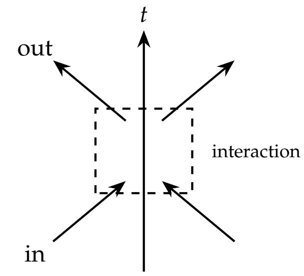


Figure 3.3: Scattering: particles come in from $t \to -\infty$, interact, and go out to $t \to +\infty$.

3.5 A Concrete Example: $2 \to 2$ Scattering

Let's compute a scattering amplitude. Consider two particles with momenta p_1, p_2 scattering into two particles with momenta p_3, p_4:

$$|i\rangle = |p_1, p_2\rangle = \sqrt{2\omega_1}\sqrt{2\omega_2}\, a_{p_1}^\dagger a_{p_2}^\dagger |0\rangle \tag{3.18}$$

$$|f\rangle = |p_3, p_4\rangle = \sqrt{2\omega_3}\sqrt{2\omega_4}\, a_{p_3}^\dagger a_{p_4}^\dagger |0\rangle \tag{3.19}$$

The $S = 1$ term contributes only if $|f\rangle = |i\rangle$ (no scattering). For actual scattering, we need $S^{(1)}$ at least.

The first-order term is

$$S^{(1)} = -i\frac{\lambda}{4!}\int d^4x\, \phi_I^4(x) \tag{3.20}$$

To compute $\langle f|S^{(1)}|i\rangle$, we need to evaluate

$$\langle p_3, p_4|\phi_I^4(x)|p_1, p_2\rangle \tag{3.21}$$

Each ϕ_I can either create or annihilate a particle. We need to annihilate both incoming particles and create both outgoing particles. That requires four field operators, which is exactly what we have!

Working Out the Contractions

Expand $\phi_I(x)$ in creation and annihilation operators:

$$\phi_I(x) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_k}} \left[a_k e^{-ikx} + a_k^\dagger e^{ikx} \right] \tag{3.22}$$

Each of the four ϕs in ϕ^4 can contribute an a or an $a^\dagger$.

To go from $|p_1, p_2\rangle$ to $|p_3, p_4\rangle$, we need:

- Two a's to annihilate the incoming particles
- Two $a^\dagger$'s to create the outgoing particles

The calculation gives (after carefully counting factors):

$$\langle p_3, p_4 | S^{(1)} | p_1, p_2 \rangle = -i\lambda \int d^4x \, e^{i(p_1+p_2-p_3-p_4)\cdot x} \tag{3.23}$$

$$= -i\lambda \cdot (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) \tag{3.24}$$

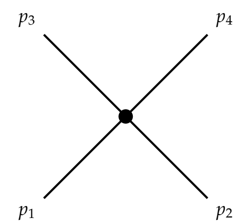


Figure 3.4: The tree-level 2 → 2 scattering diagram in ϕ^4 theory: four particles meet at a single vertex.

The delta function enforces momentum conservation—no surprise, since the interaction is translationally invariant.

We define the *invariant amplitude* $\mathcal{M}$ by stripping off the momentum-conserving delta function:

$$\langle f | S | i \rangle = (2\pi)^4 \delta^{(4)}(p_f - p_i) \cdot i\mathcal{M} \tag{3.25}$$

So at tree level in ϕ^4 theory, $\mathcal{M} = -\lambda$.

The cross-section is proportional to $|\mathcal{M}|^2 = \lambda^2$. This is why the coupling constant governs the strength of scattering.

3.6 Wick's Theorem

The calculation above was manageable because we had just four field operators. At higher orders, we have products of many fields, and keeping track of all the ways to pair creation and annihilation operators becomes tedious.

Wick's theorem provides a systematic way to do this. It relates time-ordered products to *normal-ordered* products (all $a^\dagger$s to the left) plus *contractions* (propagators).

Define the *contraction* of two fields as the difference between the time-ordered and normal-ordered products:

$$\phi(x)\phi\phi(x)\phi(y) = T[\phi(x)\phi(y)] - :\phi(x)\phi(y): \tag{3.26}$$

Since normal-ordered products give zero in vacuum expectation values, the contraction is just the propagator:

$$\phi(x)\phi\phi(x)\phi(y) = \langle 0 | T[\phi(x)\phi(y)] | 0 \rangle = \Delta_F(x-y) \tag{3.27}$$

Wick's theorem: The time-ordered product of fields equals the sum over all ways of contracting pairs:

$$\begin{aligned} T[\phi_1\phi_2\phi_3\phi_4] &= :\phi_1\phi_2\phi_3\phi_4: \\ &+ \phi_1\phi\phi_1\phi_2 :\phi_3\phi_4: + (\text{5 more single contractions}) \\ &+ \phi_1\phi\phi_1\phi_2\phi_3\phi\phi_3\phi_4 + (\text{2 more double contractions}) \end{aligned} \tag{3.28}$$

Each contraction replaces two fields with a propagator.

When we take the vacuum expectation value $\langle 0|\cdots|0\rangle$, only the fully contracted terms survive (normal-ordered terms with leftover fields give zero). So:

$$\langle 0|T[\phi_1\phi_2\phi_3\phi_4]|0\rangle = \Delta_F(x_1 - x_2)\Delta_F(x_3 - x_4) + 2 \text{ other pairings} \tag{3.29}$$

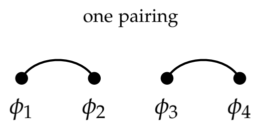


Figure 3.5: One way to contract four fields into two pairs. Each contraction becomes a propagator.

3.7 *Scattering Amplitudes from Wick's Theorem*

For S-matrix elements between particle states, we get a mix of:

- **External contractions**: Fields paired with external particles (creation/annihilation).
- **Internal contractions**: Fields paired with each other (propagators).

Consider the second-order term in ϕ^4:

$$S^{(2)} = \frac{(-i\lambda)^2}{2! \cdot (4!)^2} \int d^4x\, d^4y\, T[\phi^4(x)\phi^4(y)] \tag{3.30}$$

We have 8 fields. For a $2 \to 2$ process, 4 fields connect to external particles and 4 form internal contractions (two propagators).

Different ways of making these pairings give different contributions to the amplitude. Each pairing corresponds to a Feynman diagram, which we'll develop systematically in the next chapter.

3.8 *The Structure of Perturbation Theory*

Let me step back and describe the overall structure.

The S-matrix has the expansion

$$S = 1 + \sum_{n=1}^{\infty} S^{(n)} \tag{3.31}$$

where $S^{(n)} \propto \lambda^n$ (or e^n or g^n for other couplings).

For a given process (say, $2 \to 2$ scattering), $S^{(n)}$ involves an integral over n spacetime points (where interactions occur) and a sum over all ways to connect external particles and propagators.

Each term can be computed (in principle) using Wick's theorem. The result is a sum of integrals, one for each way to pair up fields.

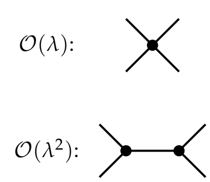


Figure 3.6: The perturbation series: higher orders have more vertices.

At low orders, there are only a few diagrams and the calculation is tractable. At high orders, the number of diagrams explodes combinatorially. For precision calculations in QED, people have computed to order α^5—thousands of diagrams.

3.9 Why Time-Ordering?

You might wonder: why is the S-matrix defined with time-ordering? Why not just $\exp(-i \int H_I)$?

The answer involves causality. Physical processes respect causality: the future cannot affect the past. Time-ordering builds this in at the level of quantum amplitudes.

Consider the propagator. For $x^0 > y^0$, it describes a particle created at y propagating to x. For $y^0 > x^0$, it describes a particle created at x propagating to y. The propagator automatically handles both cases.

But actually, there's something deeper going on. The Feynman propagator isn't just time-ordered; it's the specific combination that treats positive-frequency modes (particles) propagating forward and negative-frequency modes (antiparticles) propagating backward. This combination is what ensures causality in the quantum theory.[2]

[2] For spacelike separations, the propagator doesn't vanish, but commutators of fields do. This is enough to ensure that measurements at spacelike separation can't influence each other.

3.10 The LSZ Reduction Formula

There's an important result that connects S-matrix elements to correlation functions: the *LSZ reduction formula* (Lehmann-Symanzik-Zimmermann).

The idea is: instead of working with in/out states explicitly, compute the vacuum expectation value of a time-ordered product of fields, then "amputate" the external propagators.

For a 2 → 2 process:

$$\langle p_3, p_4|S|p_1, p_2\rangle = \prod_{i=1}^{4} \left[\int d^4x_i \, e^{\pm i p_i x_i} (\Box_{x_i} + m^2) \right] \times \langle 0|T[\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)]|0\rangle \quad (3.32)$$

The $(\Box + m^2)$ operators kill the external propagators (since $(\Box + m^2)\Delta_F(x) = -i\delta^{(4)}(x)$), leaving only the "amputated" part.

This formula is powerful because correlation functions $\langle 0|T[\phi(x_1) \cdots \phi(x_n)]|0\rangle$ are often easier to compute than S-matrix elements directly, especially using path integral methods.

3.11 A Worked Example: The Four-Point Function

Let's compute the four-point correlation function at order λ:

$$G^{(4)}(x_1, x_2, x_3, x_4) = \langle 0|T[\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)]|0\rangle \tag{3.33}$$

At zeroth order (free theory), Wick's theorem gives:

$$\begin{aligned} G_0^{(4)} &= \Delta_F(x_1 - x_2)\Delta_F(x_3 - x_4) + \Delta_F(x_1 - x_3)\Delta_F(x_2 - x_4) \\ &\quad + \Delta_F(x_1 - x_4)\Delta_F(x_2 - x_3) \end{aligned} \tag{3.34}$$

These are the three ways to pair four points into two pairs.

At first order in λ:

$$G_1^{(4)} = -\frac{i\lambda}{4!}\int d^4y\, \langle 0|T[\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\phi^4(y)]|0\rangle \tag{3.35}$$

We have 8 fields, and we need to contract them all. The interesting term has each external point x_i contracted with one of the $\phi(y)$s:

$$G_1^{(4)} = -i\lambda \int d^4y\, \Delta_F(x_1 - y)\Delta_F(x_2 - y)\Delta_F(x_3 - y)\Delta_F(x_4 - y) + \cdots \tag{3.36}$$

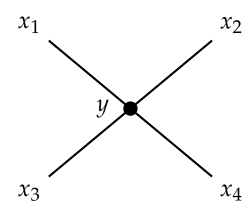


Figure 3.7: The first-order contribution to the four-point function. Each external point connects to the vertex y via a propagator.

The "$\cdots$" includes terms where some x_is are contracted with each other rather than with y—these turn out to be "disconnected" contributions that don't contribute to scattering.

Fourier transforming to momentum space:

$$\tilde{G}_1^{(4)}(p_1, p_2, p_3, p_4) = -i\lambda \cdot (2\pi)^4\delta^{(4)}(p_1 + p_2 + p_3 + p_4) \cdot \prod_{i=1}^{4} \frac{i}{p_i^2 - m^2 + i\epsilon} \tag{3.37}$$

The product of propagators gives poles when each $p_i^2 = m^2$—when the external particles are on-shell. The LSZ procedure amputates these propagators, leaving just $-i\lambda$.

3.12 What We've Learned

The perturbation theory machinery we've developed has a clear structure:

1. **Interaction picture**: Fields evolve as free fields; states evolve due to interactions.

2. **S-matrix**: Encodes scattering amplitudes as a time-ordered exponential.

3. **Wick's theorem**: Converts time-ordered products into sums over contractions.

4. **Contractions**: Each contraction is a propagator connecting two spacetime points.

5. **Perturbative expansion**: Higher orders in λ involve more vertices and more propagators.

The procedure is straightforward—you could teach a computer to do it—but there's something almost magical about watching the combinatorics organize themselves into physical amplitudes. Each contraction corresponds to a particle propagating between two points. The sum over contractions becomes a sum over all the ways particles can be exchanged.

But enumerating contractions by hand gets tedious fast. This machinery screams for a better notation—something that captures the combinatorics visually, that lets you see at a glance what particles are doing. That notation is Feynman diagrams, and it's the subject of our next chapter.

3.13 A Philosophical Note

Before moving on, let me make a philosophical point.

We're expanding in powers of λ, assuming λ is small. But what does "small" mean? There's no dimensionful scale in the coupling constant to compare to anything.

The expansion parameter turns out to be $\lambda/(16\pi^2)$, roughly. This factor of $16\pi^2 \approx 158$ comes from phase space integrals in loops. So even for $\lambda \sim 1$, the expansion is reasonably well-behaved.

But the series doesn't converge. At high orders, the number of Feynman diagrams grows factorially (roughly $n!$ at order n), while the coupling suppression grows only like λ^n. The perturbation series is *asymptotic*: it gives a good approximation up to some optimal truncation, beyond which adding more terms makes things worse.

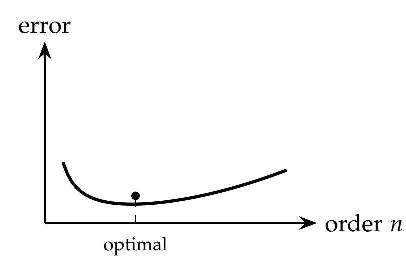


Figure 3.8: An asymptotic series: the error decreases to an optimal point, then increases.

This doesn't mean perturbation theory is useless—far from it. For QED, the first few terms give predictions accurate to parts per billion. But it does mean we're not computing the "exact" answer by summing infinitely many terms. We're computing an approximation, and the approximation is spectacularly good for weakly coupled theories.

Whether this limitation is fundamental or whether there's a way to "resum" the perturbation series and extract exact answers is an open question. For our purposes, we'll trust perturbation theory where it's warranted and note its limitations where they matter.

The interaction picture and perturbative S-matrix were developed in the late 1940s and early 1950s, during the heroic era of quantum electrodynamics. Tomonaga, Schwinger, and Feynman developed equivalent formulations (Tomonaga and Schwinger using the approach sketched here; Feynman using path integrals). Dyson showed in 1949 that the approaches were equivalent and provided the systematic framework we still use today. The combinatorial explosion of diagrams at high orders, and the asymptotic nature of the series, were understood later. Modern techniques—renormalization group methods, resummation, lattice QFT—go beyond naive perturbation theory in various ways, but for weakly coupled theories like QED, the approach developed in the 1940s remains the gold standard.

4
The Language of Feynman Diagrams

In the last chapter, we developed the machinery of perturbation theory: time-ordered products, Wick contractions, the S-matrix expansion. The calculations work, but they're bookkeeping nightmares. You enumerate contractions, track signs, match up creation and annihilation operators, and eventually get an integral. The process obscures the physics.

Feynman diagrams are a better way. They turn the combinatorics into pictures. Each diagram represents a term in the perturbation series. You draw a picture, read off the corresponding integral, and compute. The physics becomes visual: particles propagate along lines, interact at vertices, create and annihilate.

More than just computational tools, Feynman diagrams shape how we think about quantum processes. When physicists discuss scattering, they speak of "tree diagrams" and "loop diagrams," of particles being "exchanged" and processes being "mediated." This language comes from the diagrams.

Let me show you how it works.

4.1 From Contractions to Pictures

Recall the structure of perturbation theory. The S-matrix is

$$S = T \exp\left(-i \int d^4x \, \mathcal{H}_I(x)\right) = 1 + S^{(1)} + S^{(2)} + \cdots \qquad (4.1)$$

Each term involves products of field operators. Wick's theorem converts these into sums over contractions. Each contraction is a propagator.

The Feynman diagram is a graphical representation:

- Each **vertex** represents an interaction point (a factor of $\mathcal{H}_I$).
- Each **internal line** represents a propagator (a contraction between two fields).

- Each **external line** represents an external particle (in the initial or final state).

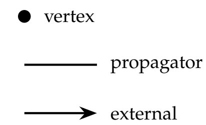
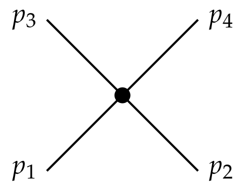


Figure 4.1: The basic elements of Feynman diagrams.

For ϕ^4 theory, each vertex has four lines meeting (because the interaction is ϕ^4). The tree-level $2 \to 2$ scattering amplitude, which we calculated laboriously in Chapter 3, is just:



Four external particles meet at one vertex. No internal lines (this is tree-level—no loops). The amplitude is just $-i\lambda$ times a momentum-conserving delta function.

4.2 *Feynman Rules in Position Space*

Let me state the Feynman rules systematically. For ϕ^4 theory in position space:

1. **Draw all distinct diagrams** with the required external lines and the appropriate number of vertices.
2. **For each vertex at position** x: write a factor of $-i\lambda$.
3. **For each internal line from** x **to** y: write a propagator $\Delta_F(x-y)$.
4. **For each external line from** x **to external momentum** p: write $e^{\pm ipx}$, where the sign depends on whether the particle is incoming $(-)$ or outgoing $(+)$.
5. **Integrate over all vertex positions**: $\int d^4x_1 \cdots d^4x_n$.
6. **Divide by the symmetry factor**: accounts for overcounting equivalent diagrams.

The symmetry factor deserves explanation. If interchanging vertices or lines gives the same diagram, we've overcounted, and must divide by the number of such interchanges.

4.3 *Feynman Rules in Momentum Space*

Position-space rules are conceptually clear but computationally awkward. The integrals over vertex positions are easier after Fourier transforming.

Using

$$\Delta_F(x-y) = \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} e^{-ik(x-y)} \tag{4.2}$$

each propagator introduces a momentum variable k flowing along the line.

The position integrals then give delta functions enforcing momentum conservation at each vertex:

$$\int d^4x \, e^{i(p_1+p_2+p_3+p_4)\cdot x} = (2\pi)^4 \delta^{(4)}(p_1 + p_2 + p_3 + p_4) \tag{4.3}$$

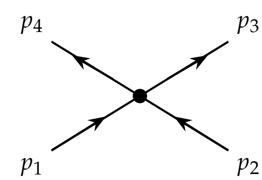


Figure 4.2: Momentum conservation at a vertex: what goes in must come out.

Momentum-space Feynman rules for ϕ^4 theory:

1. **Draw all distinct diagrams**.
2. **Assign momenta** to all lines (external momenta are fixed; internal momenta are integration variables).
3. **For each vertex**: write $-i\lambda$.
4. **For each internal line with momentum** k: write $\dfrac{i}{k^2 - m^2 + i\epsilon}$.
5. **For each loop**: integrate $\displaystyle\int \frac{d^4k}{(2\pi)^4}$ over the undetermined momentum.
6. **Impose momentum conservation** at each vertex (this determines some internal momenta in terms of others).
7. **Divide by the symmetry factor**.

The number of loop integrals equals the number of independent momenta not fixed by conservation. For a diagram with V vertices and I internal lines, the number of loops is $L = I - V + 1$.

4.4 *Counting Loops*

This formula $L = I - V + 1$ is worth understanding. Each internal line carries a momentum we integrate over. Each vertex gives a delta function constraining momenta. But one delta function just enforces overall conservation (total incoming = total outgoing), which doesn't reduce the number of integrals.

So:

$$\text{loop integrals} = I - (V - 1) = I - V + 1 \tag{4.4}$$

Diagram	V	I	L	Example
Tree (2→2)	1	0	0	ϕ^4 vertex
Tree (2→2)	2	1	0	s-channel
1-loop (2→2)	2	2	1	fish/bubble
2-loop (vac)	1	2	2	"figure-eight"

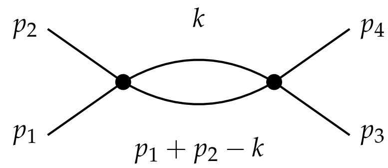
Table 4.1: Examples of the loop-counting formula $L = I - V + 1$.

At tree level ($L = 0$), all momenta are determined by external kinematics. No integrals remain—just algebra. This is why tree-level calculations are easy.

At one loop ($L = 1$), there's one undetermined momentum to integrate. This integral often diverges—the subject of our next chapter.

4.5 A One-Loop Example

Let's see a one-loop diagram. Consider the "fish" diagram contributing to $2 \to 2$ scattering at order λ^2:



This is called the s-channel diagram (after the Mandelstam variable $s = (p_1 + p_2)^2$).

Let's apply the Feynman rules:

- Two vertices: $(-i\lambda)^2 = -\lambda^2$.
- Two internal propagators with momenta k and $p_1 + p_2 - k$:

$$\frac{i}{k^2 - m^2 + i\epsilon} \cdot \frac{i}{(p_1 + p_2 - k)^2 - m^2 + i\epsilon} \tag{4.5}$$

- One loop integral: $\int \frac{d^4k}{(2\pi)^4}$.
- Symmetry factor: $1/2$ (the two internal lines are equivalent).

The amplitude is:

$$i\mathcal{M}_s = \frac{(-i\lambda)^2}{2} \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} \cdot \frac{i}{(p_1 + p_2 - k)^2 - m^2 + i\epsilon} \tag{4.6}$$

Now here's the problem. Let's estimate this integral. At large $|k|$:

$$\int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 \cdot k^2} \sim \int \frac{k^3\, dk}{k^4} = \int \frac{dk}{k} \tag{4.7}$$

This integral diverges logarithmically at large k. We have our first infinity!

We'll deal with this in Chapters 5-7. For now, just note: loop diagrams involve integrals over all momenta, and these integrals can diverge.

4.6 Symmetry Factors: The Trickiest Part

The symmetry factor is the trickiest part of Feynman rules. It compensates for overcounting when we enumerate diagrams. Getting this wrong is one of the most common errors in QFT calculations, so let's be thorough.

Where the Counting Comes From

The nth-order term in the S-matrix is

$$S^{(n)} = \frac{(-i)^n}{n!} \int \prod_{j=1}^{n} d^4 x_j \, T[\mathcal{H}_I(x_1) \cdots \mathcal{H}_I(x_n)] \tag{4.8}$$

The $1/n!$ is there because the time-ordering symbol T treats the vertices as distinguishable, but when we integrate over all positions x_j, we're summing over all ways to assign vertices to positions. If the vertices are actually identical, we've overcounted by $n!$.

Similarly, within each $\mathcal{H}_I = \frac{\lambda}{4!}\phi^4$, the $1/4!$ is there because expanding $\phi^4 = \phi \cdot \phi \cdot \phi \cdot \phi$ gives $4!$ ways to pair the four ϕ fields with other fields via Wick contractions, and all these pairings give the same result.

When we draw Feynman diagrams and apply the naive rules, we're essentially treating all internal lines and vertices as distinguishable. The *symmetry factor S* corrects for this overcounting:

$$\text{contribution} = \frac{(\text{diagram value})}{S} \tag{4.9}$$

where S counts the number of ways to permute internal lines and vertices that leave the diagram unchanged (with external legs fixed in place).

A Systematic Method

Here's a reliable algorithm for computing symmetry factors:

1. Start with the diagram. Label all vertices $1, 2, \ldots$ and all internal lines $a, b, \ldots$

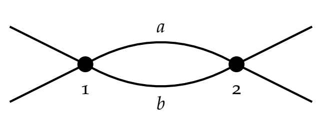
2. Count the number of distinct ways to relabel vertices and lines that produce the same diagram (with external legs held fixed).

3. That count is S.

Equivalently, S is the order of the automorphism group of the diagram—the group of permutations that preserve the structure.

Worked Example: The Fish Diagram

Consider the "fish" or "bubble" diagram in ϕ^4 theory:



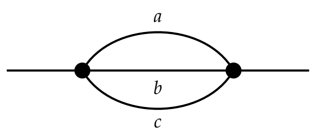
What symmetries does this diagram have?

- Can we swap vertices 1 and 2? No—that would interchange the external legs on left and right.
- Can we swap internal lines a and b? Yes! The diagram looks exactly the same.

So there's one nontrivial symmetry: $(a \leftrightarrow b)$. Together with the identity, this gives a symmetry group of order 2. Therefore $S = 2$.

Worked Example: The Sunset Diagram

Now consider the "sunset" or "sunrise" diagram:



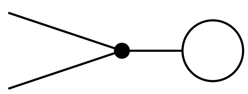
This has three internal lines connecting the same two vertices. What's the symmetry factor?

We can permute the three lines a, b, c in any of $3! = 6$ ways without changing the diagram. So $S = 6$.

But wait—in a theory like ϕ^3 where each vertex has three lines, this diagram would appear differently. The symmetry factor depends on the theory! For ϕ^4 theory, this diagram would need four lines at each vertex, so we'd need external lines making up the difference.

Worked Example: The Tadpole

The simplest loop diagram is the tadpole:

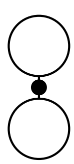


This has a loop attached to the vertex (a "bubble" on a stick). The single loop line can be "flipped" (traversed in the opposite direction), but for a scalar field, both directions give the same propagator. In terms of automorphisms: there are two ways to orient the loop, both giving the same diagram.

So $S = 2$.

Worked Example: Figure-Eight (Double Tadpole)

Consider the figure-eight vacuum diagram:



This has:

- Each loop can be flipped: factor of $2 \times 2 = 4$
- The two loops can be interchanged: factor of 2

Total: $S = 4 \times 2 = 8$.

Why Getting This Wrong Matters

If you use the wrong symmetry factor, you get the wrong coefficient for a diagram's contribution. In simple calculations at low orders, this might just give a wrong numerical factor. But in renormalization, getting symmetry factors wrong can make counterterms inconsistent, leading to apparent violations of gauge invariance or other symmetries.

The safest approach is: when in doubt, go back to the Wick contraction level and count explicitly. Every valid Wick contraction corresponds to a contribution; count them, and you'll find the right symmetry factor naturally.

A Table of Common Symmetry Factors

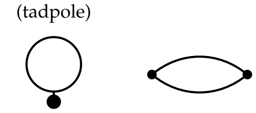


Figure 4.3: Common diagrams with their symmetry factors.

Diagram Type	S
Tree (no symmetry)	1
Fish / bubble	2
Tadpole (one loop)	2
Sunset (3 parallel lines)	6
Figure-eight	8

Table 4.2: Symmetry factors for common diagram topologies.

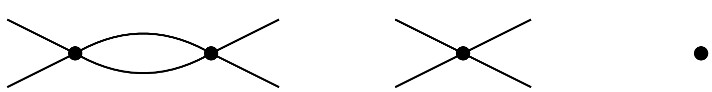
4.7 *Connected Diagrams and Vacuum Bubbles*

Not all diagrams contribute to scattering in the same way. This section explains which diagrams matter and why—a crucial point for understanding what we actually calculate.

Connected vs. Disconnected Diagrams

A diagram is **connected** if you can reach any part of it from any other part by traveling along lines. A diagram is **disconnected** if it breaks into separate pieces.

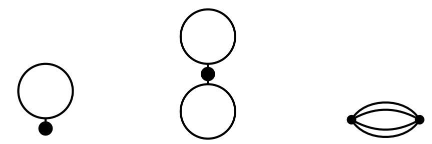
For example, consider these two diagrams for $2 \to 2$ scattering at second order:



The left diagram is connected: both vertices are linked by internal lines. The right diagram is disconnected: there's a tree-level scattering plus a floating "vacuum bubble" that isn't connected to any external particle.

What Are Vacuum Bubbles?

A **vacuum bubble** (or vacuum diagram) is a connected subdiagram with no external legs—a closed diagram floating in the vacuum. Examples:



These diagrams represent virtual processes happening in the vacuum with no connection to any physical particles.

Why Vacuum Bubbles Cancel

Here's the key result: vacuum bubbles cancel out of all physical observables.

The intuition is this. The vacuum-to-vacuum amplitude $\langle 0|S|0\rangle$ includes all possible vacuum bubbles:

$$\langle 0|S|0\rangle = 1 + (\text{one bubble}) + (\text{two bubbles}) + \cdots = e^{i\theta} \tag{4.10}$$

where θ is some (possibly infinite) phase.

Any scattering amplitude factorizes:

$$\langle f|S|i\rangle = \langle f|S|i\rangle_{\text{connected}} \times \langle 0|S|0\rangle \tag{4.11}$$

The vacuum bubbles multiply the connected amplitude by the same factor $e^{i\theta}$.

But we never measure $\langle f|S|i\rangle$ directly—we measure probabilities, which are:

$$|\langle f|S|i\rangle|^2 = |\langle f|S|i\rangle_{\text{connected}}|^2 \times |\langle 0|S|0\rangle|^2 \tag{4.12}$$

If we properly normalize our states (or, equivalently, divide by the vacuum normalization), the factor $|\langle 0|S|0\rangle|^2$ cancels:

$$\frac{|\langle f|S|i\rangle|^2}{|\langle 0|S|0\rangle|^2} = |\langle f|S|i\rangle_{\text{connected}}|^2 \tag{4.13}$$

Upshot: Vacuum bubbles contribute an overall phase that cancels in physical observables. We can ignore them entirely when computing scattering amplitudes.

The Linked Cluster Theorem

This cancellation is formalized in the **linked cluster theorem** (or **connected diagram theorem**). It states:

The logarithm of the S-matrix equals the sum of all connected diagrams.

In equations:

$$\log\langle 0|S|0\rangle = \sum_{\text{connected vacuum diagrams}} \tag{4.14}$$

and more generally:

$$\langle f|S|i\rangle_{\text{connected}} = \text{sum of connected diagrams with external legs} \tag{4.15}$$

The theorem tells us that the exponential structure of vacuum bubbles (they can appear any number of times, independently) is exactly what's needed for them to factor out.

This is not obvious! It's a nontrivial result about the structure of perturbation theory. But once you know it, you can simplify calculations enormously by ignoring all disconnected diagrams.

Different Kinds of "Disconnected"

Be careful about terminology. There are different ways a diagram can fail to be connected:

1. **Disconnected with vacuum bubble**: A scattering part plus a floating bubble (no external legs on the bubble). These cancel, as discussed.

2. **Disconnected scattering**: Two separate scattering processes that happen independently. For example, if you scatter $A + B \to C + D$ and simultaneously scatter $E + F \to G + H$ with no interaction between the two processes. These contribute to the S-matrix but represent independent events.

The second type doesn't "cancel"—it represents legitimate physics where two things happen independently. But for computing a specific scattering amplitude with definite external particles, we still focus on connected diagrams for that process.

Practical Rule

For computing scattering amplitudes $\mathcal{M}$:

Only draw connected diagrams.

Any diagram with a floating vacuum bubble can be ignored—it will cancel when we compute physical quantities. This is an enormous simplification in practice.

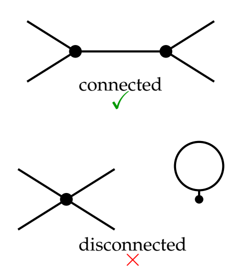
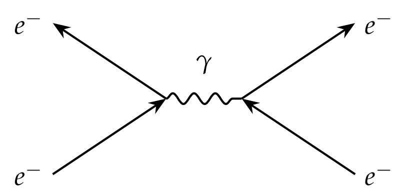


Figure 4.4: Connected diagrams contribute; disconnected diagrams with vacuum bubbles cancel.

4.8 *The Physical Picture*

Feynman diagrams aren't just computational tools—they provide a physical picture of quantum processes.

Consider electron-electron scattering in QED. The tree-level diagram shows the electrons exchanging a photon:



We say the electromagnetic force is "mediated" by photon exchange. The photon is a *virtual particle*—it exists only during the interaction, not before or after. Its momentum is determined by kinematics, and it's generally off-shell ($k^2 \neq 0$ for a virtual photon).

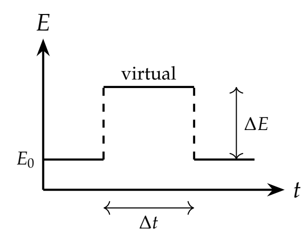


Figure 4.5: A virtual particle borrows energy ΔE for time $\Delta t \lesssim \hbar/\Delta E$.

Is the virtual photon "real"? This is a genuinely subtle question, and physicists disagree. You can't observe it directly—try to measure the intermediate photon, and you've changed the experiment into something else entirely. The virtual photon exists only as an intermediate state in a quantum amplitude, with no independent existence before or after.

Some physicists will tell you virtual particles are "just mathematics"—artifacts of perturbation theory with no physical reality. Others insist they're as real as anything else in quantum mechanics: after all, their effects are measurable. The electrons really do repel each other; something must be happening.

I'm not sure the distinction matters. What matters is that we have a calculation that gives the right answer. But if the ontological question bothers you, hold onto that discomfort. It points at something deep about the interpretation of quantum field theory that we don't fully understand.

4.9 *Putting in Numbers: Tree-Level Scattering*

Let's do a complete calculation for ϕ^4 theory. We'll compute the cross-section for $2 \to 2$ scattering at tree level.

The tree-level amplitude is

$$i\mathcal{M} = -i\lambda \tag{4.16}$$

So $|\mathcal{M}|^2 = \lambda^2$.

The differential cross-section in the center-of-mass frame is[1]

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s}|\mathcal{M}|^2 \tag{4.17}$$

[1] This formula comes from phase-space integrals and flux factors; we won't derive it here.

where $s = (p_1 + p_2)^2$ is the center-of-mass energy squared.

For $\lambda = 0.1$ and $\sqrt{s} = 1$ GeV:

$$\frac{d\sigma}{d\Omega} = \frac{\lambda^2}{64\pi^2 s} = \frac{0.01}{64\pi^2 \cdot (1\,\mathrm{GeV})^2} \approx 1.6 \times 10^{-5}\,\mathrm{GeV}^{-2} \tag{4.18}$$

Converting to more useful units (using $1\,\mathrm{GeV}^{-2} \approx 0.389\,\mathrm{mb}$):

$$\frac{d\sigma}{d\Omega} \approx 6 \times 10^{-6}\,\mathrm{mb} = 6 \times 10^{-33}\,\mathrm{cm}^2 \tag{4.19}$$

This is small—much smaller than typical hadronic cross-sections (tens of millibarns) but comparable to weak-interaction cross-sections. The coupling $\lambda = 0.1$ is moderately weak.

Integrating over angles gives the total cross-section. For identical particles, we must divide by 2 to avoid double-counting final states where the two particles are exchanged:[2]

[2] Swapping the two identical outgoing particles gives the same physical final state.

$$\sigma_{\mathrm{tot}} = \frac{1}{2} \cdot 4\pi \cdot \frac{d\sigma}{d\Omega} \approx 3.8 \times 10^{-5}\,\mathrm{mb} \tag{4.20}$$

4.10 *Beyond ϕ^4: Different Theories, Different Rules*

The Feynman rules depend on the theory. Let me sketch how other theories work.

ϕ^3 **theory**: Each vertex has three lines (factor: $-ig$). Unstable vacuum, but useful as a toy model.

Yukawa theory: A scalar ϕ couples to a fermion ψ via $g\bar{\psi}\phi\psi$. The vertex has two fermion lines and one scalar line.

QED: The photon A_μ couples to the electron ψ via $e\bar{\psi}\gamma^\mu\psi A_\mu$. The vertex has two electron lines and one photon line.

Each theory has its own propagators:

- Scalar: $\dfrac{i}{k^2 - m^2 + i\epsilon}$
- Fermion: $\dfrac{i(\not{k} + m)}{k^2 - m^2 + i\epsilon}$
- Photon: $\dfrac{-ig_{\mu\nu}}{k^2 + i\epsilon}$ (in Feynman gauge)

The numerator of the fermion propagator contains $\not{k} = \gamma^\mu k_\mu$—the gamma matrices make fermion calculations more involved. The photon propagator involves the metric tensor because photons carry polarization.

We'll develop the QED rules fully in Chapter 8.

4.11 Reading Diagrams: What Diverges and Why

Before we tackle infinities systematically, let's build intuition for which diagrams diverge.

Consider a generic loop integral:

$$\int \frac{d^4k}{(2\pi)^4} \frac{1}{(k^2 - m^2)^n} \tag{4.21}$$

For large k, this behaves like $\int k^3 dk/k^{2n} = \int k^{3-2n} dk$.

- If $3 - 2n < -1$, the integral converges.
- If $3 - 2n = -1$ (i.e., $n = 2$), the integral diverges logarithmically.
- If $3 - 2n > -1$ (i.e., $n < 2$), the integral diverges like a power of the cutoff.

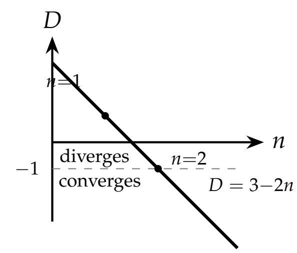


Figure 4.6: Degree of divergence $D = 3 - 2n$. Convergence requires $D < -1$, i.e., $n > 2$.

In ϕ^4 theory, a one-loop diagram has two propagators ($n = 2$), so it diverges logarithmically. A two-loop diagram might have more propagators, but the additional loop adds another k integration, potentially making things worse.

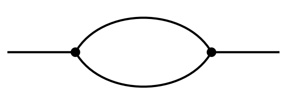
The systematic way to analyze this is *power counting*, which we'll develop in Chapter 6. For now, the intuition is: loops are dangerous because they integrate over all momenta, and high momenta can cause trouble.

4.12 A Bestiary of Diagrams

Let me catalog some common diagram types. These names will recur constantly, and each type asks a different physical question.

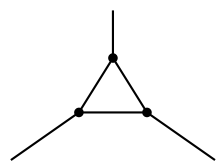
Tree diagrams have no loops. All internal momenta are fixed by external momenta, so there's nothing to integrate over. They're the leading-order approximation to any process, and they're always finite.

Self-energy diagrams correct the propagator. They're asking: what is the mass of a particle that's constantly surrounded by its own quantum fluctuations? A particle propagates, emits and reabsorbs virtual particles, then continues. The answer, as we'll see, is divergent.



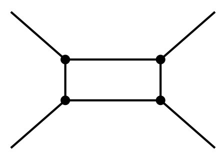
self-energy

Vertex corrections modify the interaction itself. The basic coupling gets dressed by virtual particles, changing its effective strength. In QED, the vertex correction will tell us how the electron's magnetic moment differs from Dirac's prediction—a measurable effect that agrees with experiment to extraordinary precision.



vertex correction

Box diagrams appear when two particles exchange two virtual particles. They're more complicated than single-exchange diagrams, but they're essential for precision calculations and for processes that can't happen at tree level.



box diagram

Tadpoles are loops connected to a single vertex—little bubbles that grow off a line and pop. They often vanish by symmetry, but when they don't, they shift vacuum expectation values.

In QED, three diagrams control the renormalization: the photon self-energy (vacuum polarization), the electron self-energy, and the vertex correction. We'll compute all three in full detail, and doing so will teach us what renormalization really means.

4.13 *Summary: The Feynman Rules*

For ϕ^4 theory in momentum space:

Element	**Factor**
Vertex	$-i\lambda$
Propagator (momentum k)	$\dfrac{i}{k^2 - m^2 + i\epsilon}$
Loop	$\int \dfrac{d^4k}{(2\pi)^4}$
Momentum conservation	at each vertex
Symmetry factor	divide by S

The procedure:

1. Draw all topologically distinct connected diagrams.
2. Assign momenta (external fixed, internal integrated).
3. Write down the factors.
4. Integrate and compute.

We now have the tools to do calculations. In the next chapter, we'll face the main challenge: those loop integrals diverge, and we need to understand why and what to do about it.

Feynman introduced his diagrams in his 1949 papers on QED. They were initially met with skepticism—Bohr reportedly thought they gave the wrong picture of quantum mechanics, suggesting particles followed definite trajectories. But their computational power won out. Dyson showed how Feynman's diagrams, Schwinger's operator methods, and Tomonaga's covariant perturbation theory were all equivalent. The diagrams became the standard language of particle physics. Today, every particle physics paper includes Feynman diagrams, and the visual vocabulary—propagators, vertices, loops—shapes how we think about fundamental interactions. The transition from Wick contractions to diagrams is more than notational convenience; it's a shift in how we conceptualize quantum processes.

5
Why Infinity?

We've built the machinery of quantum field theory: fields, particles, propagators, Feynman diagrams, perturbation theory. Now we run into a wall.

When we try to compute loop corrections—quantum corrections from virtual particles—we get infinity. Not a large number. Actual infinity. The integrals diverge.

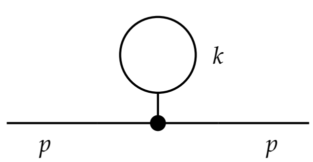
This chapter is about understanding *why*. What is the physical origin of these infinities? Are they telling us something deep about nature, or are they just an artifact of our computational methods? What went wrong?

I want to be clear: these infinities are not optional. They're not a mistake in the calculation. Any honest treatment of quantum field theory in the continuum produces them. They're built into the structure of the theory.

Understanding why is the first step toward understanding how to deal with them.

5.1 *The One-Loop Correction in ϕ^4 Theory*

Let's examine a specific divergent integral. Consider the "tadpole" diagram in ϕ^4 theory—a loop attached to a single vertex, contributing to the mass of the scalar particle:



By the Feynman rules, this contributes:

$$-i\Sigma = (-i\lambda) \cdot \frac{1}{2} \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} \tag{5.1}$$

The $1/2$ is the symmetry factor (we can flip the loop). This is a correction to the propagator, contributing a "self-energy" Σ.

Let's evaluate this integral. In Euclidean space (after Wick rotation, $k^0 \to ik_E^0$):

$$\Sigma = \frac{\lambda}{2} \int \frac{d^4k_E}{(2\pi)^4} \frac{1}{k_E^2 + m^2} \tag{5.2}$$

The integral is over all of 4-dimensional momentum space. In spherical coordinates:

$$\int d^4k_E = \int_0^\infty k_E^3\, dk_E \int d\Omega_4 = 2\pi^2 \int_0^\infty k_E^3\, dk_E \tag{5.3}$$

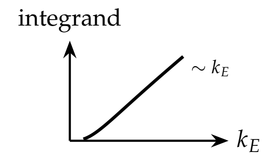


Figure 5.1: The integrand grows like k_E at large k_E.

where $2\pi^2$ is the surface area of the unit 3-sphere.

So:

$$\Sigma = \frac{\lambda}{2} \cdot \frac{2\pi^2}{(2\pi)^4} \int_0^\infty dk_E \frac{k_E^3}{k_E^2 + m^2} \tag{5.4}$$

At large k_E, the integrand behaves like $k_E^3/k_E^2 = k_E$. The integral diverges:

$$\int_m^\Lambda dk_E\, k_E = \frac{\Lambda^2 - m^2}{2} \xrightarrow{\Lambda\to\infty} \infty \tag{5.5}$$

The integral is *quadratically divergent*—it grows like Λ^2 when we cut it off at some large momentum Λ.

5.2 *What Does This Mean Physically?*

Let's think about what the integral represents. We're summing over all possible momenta k of the virtual particle circulating in the loop. The momentum can be anything—there's no constraint on it.

The divergence comes from the high-momentum (short-distance) region. When k is much larger than m or any external momentum, the propagator becomes $1/k^2$, and integrating $k^3dk/k^2 = k\,dk$ diverges.

Interpretation 1: Short-distance singularity. In position space, the propagator $\Delta_F(x-y)$ is singular when $x \to y$. The tadpole diagram involves $\Delta_F(0)$—the propagator at zero separation. This is the field "interacting with itself" at the same point, and such pointlike interactions produce infinities.

Classical electromagnetism has the same problem. The self-energy of a point charge is infinite—the energy stored in the electric field $E \sim 1/r^2$ diverges as you integrate toward the charge. Classical physicists knew this and it troubled them. They tried giving the electron a finite radius, but nobody knew how big it should be or what held it together. Quantum field theory inherits this short-distance pathology. It's not a bug in our quantum treatment; it's a feature we took over from classical physics and made worse.

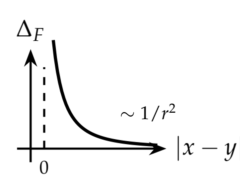


Figure 5.2: The propagator diverges as points approach—the position-space origin of UV divergences.

Interpretation 2: Virtual particles of arbitrarily high energy. In the momentum picture, we're summing over virtual particles with momenta from 0 to ∞. There's no upper limit on how energetic a virtual particle can be. The sum over high-energy virtual particles is what diverges.

But wait—doesn't energy conservation forbid arbitrarily high energies? For *real* particles, yes. But virtual particles are off-shell and can have any momentum. They exist only briefly (by the energy-time uncertainty relation) and don't violate energy conservation in the long run. The integral sums over all these fleeting high-energy fluctuations.

Interpretation 3: We're taking a continuum limit. Quantum field theory assumes spacetime is a continuum—there are field values at every point, no matter how close. But we're integrating over momenta up to infinity, which corresponds to arbitrarily short wavelengths.

If spacetime were discrete at some scale a, momenta would be bounded by roughly $\Lambda \sim 1/a$. The integral would be finite. The divergence signals that we're taking a continuum limit without being careful about what happens at short distances.

5.3 *The Degree of Divergence*

Different loop integrals diverge in different ways. Let's classify them.

For a general one-loop integral in four dimensions:

$$\int \frac{d^4k}{(2\pi)^4} \frac{k^a}{(k^2 - m^2)^b} \tag{5.6}$$

At large k, this behaves like k^{4+a-2b}. The integral:

- **Converges** if $4 + a - 2b < 0$
- **Diverges logarithmically** if $4 + a - 2b = 0$
- **Diverges like** Λ^{4+a-2b} if $4 + a - 2b > 0$

The *superficial degree of divergence D* is defined as $D = 4 + a - 2b$. For the tadpole ($a = 0$, $b = 1$), we have $D = 2$: quadratic divergence.

For the "fish" diagram we saw in Chapter 4 (two propagators), we have $D = 4 - 4 = 0$: logarithmic divergence. Logarithmic divergences are the mildest—they grow slowly with the cutoff.

Diagram	Props	D	Type
Tadpole	1	2	quad.
Fish	2	0	log.
Triangle	3	−2	finite
Box	4	−4	finite

Table 5.1: Degree of divergence for one-loop diagrams in ϕ^4.

5.4 *Power Counting: Which Diagrams Diverge?*

For a general diagram, we can determine the degree of divergence by power counting, without doing the integral.

Consider a diagram with:

- L loops
- I internal lines (propagators)
- V vertices
- E external lines

Each loop contributes d^4k, adding 4 powers of momentum. Each propagator contributes $1/k^2$, subtracting 2 powers. So:

$$D = 4L - 2I \tag{5.7}$$

We can relate these using topology. In ϕ^4 theory, each vertex has 4 lines, so:

$$4V = 2I + E \tag{5.8}$$

(Each internal line connects two vertices; each external line connects one vertex to the outside.)

Also, $L = I - V + 1$ (the loop-counting formula from Chapter 4).

Combining these:

$$D = 4L - 2I = 4(I - V + 1) - 2I = 2I - 4V + 4 = 4 - E \tag{5.9}$$

This is significant: the degree of divergence depends only on the number of external lines, not on the internal structure of the diagram.

For ϕ^4 theory:

- $E = 0$ (vacuum diagrams): $D = 4$, quartically divergent
- $E = 2$ (propagator corrections): $D = 2$, quadratically divergent
- $E = 4$ (4-point functions): $D = 0$, logarithmically divergent
- $E \geq 6$: $D < 0$, convergent

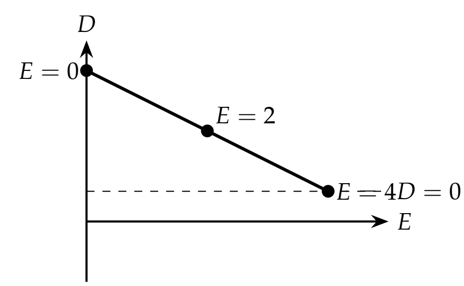


Figure 5.3: In ϕ^4 theory, $D = 4 - E$. Diagrams with $E \leq 4$ external lines are divergent.

Only diagrams with $E \leq 4$ diverge. This is a finite number of classes. This will be crucial for renormalization: we only need to "fix" a finite number of divergent structures.

5.5 *The Same Argument for QED*

Let's repeat the analysis for QED. The vertex is $e\bar{\psi}\gamma^\mu\psi A_\mu$, connecting two fermion lines and one photon line.

The power counting is more involved because fermion propagators behave differently (they go like $1/k$ at large k rather than $1/k^2$), and there are different types of lines.

After careful accounting, the degree of divergence for a QED diagram is:[1]

[1] We'll derive this properly in Chapter 8.

$$D = 4 - \frac{3}{2}E_f - E_\gamma \tag{5.10}$$

where E_f is the number of external fermion lines and E_γ is the number of external photon lines.

The divergent diagrams are:

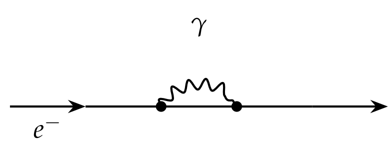
- $E_f = 0$, $E_\gamma = 2$: $D = 2$ (photon self-energy)
- $E_f = 2$, $E_\gamma = 0$: $D = 1$ (fermion self-energy)
- $E_f = 2$, $E_\gamma = 1$: $D = 0$ (vertex correction)

These are the only three divergent structures in QED. All other diagrams are finite. This is why QED is *renormalizable*: there's only a finite number of infinities to deal with.

5.6 What the Divergence Tells Us

Let's think more carefully about what the infinity means.

Consider the electron self-energy in QED. The electron propagates, emits a virtual photon, reabsorbs it, and continues. This process happens at all distance scales—the photon can have any wavelength.



Very short-wavelength photons (high-momentum) contribute to this process. There's no cutoff—we integrate all the way to infinite momentum. This is the divergence.

But physically, we shouldn't expect our theory to be valid to arbitrarily short distances. At some point—perhaps the Planck scale $\ell_P \sim 10^{-35}$ m—spacetime itself might become discrete, or gravity becomes important, or something else happens. Our theory is an approximation that breaks down at some scale.

The divergence is telling us: *this quantity depends on physics we don't know*.

Think about what "the electron mass" means in this context. When we write m_e in the Lagrangian, what number should we put? The measured electron mass is a physical observable—we can measure it by deflecting electrons in a magnetic field. But the measured mass includes the effect of all those virtual photon emissions and absorptions.

Here's the strange thing: suppose you could somehow strip away all the quantum fluctuations and measure the "bare" electron mass.

How would you do it? You'd need to probe the electron at arbitrarily short distances, but that means arbitrarily high energies, which means you can't avoid the fluctuations. The "bare" mass isn't just hard to measure—it might not be a meaningful concept at all. What we call the electron mass is always the dressed mass, fluctuations included.

The divergence arises because we're trying to compute the relationship between the bare mass (a parameter in the Lagrangian) and the physical mass (what we measure). This relationship involves physics at all scales, including scales where our theory isn't valid.

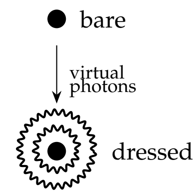


Figure 5.4: The "bare" electron vs. the "dressed" electron surrounded by virtual photons.

5.7 Why Not Just Cut Off the Integral?

Here's a natural thought: if the divergence comes from high momenta, why not just stop the integral at some cutoff Λ?

We can do this. The integral becomes:

$$\Sigma(\Lambda) = \frac{\lambda}{32\pi^2}\Lambda^2 + \text{finite terms} \tag{5.11}$$

Now Σ is finite—it depends on Λ. If we take $\Lambda \to \infty$, we recover the divergence. If we keep Λ finite, everything is well-defined.

But here's the problem: Λ is not a physical quantity. Where does it come from? What determines its value?

If Λ represents "the scale where new physics kicks in," we should be able to relate it to something observable. And indeed, the theory with a finite cutoff makes predictions that depend on Λ. Different values of Λ give different predictions.

The key insight of renormalization is that we can *absorb* the Λ-dependence into the parameters of the theory. We redefine the mass and coupling constant to depend on Λ in exactly the way needed to make physical predictions Λ-independent.

This sounds like sleight of hand—and in a way, it is. But it works because only a finite number of structures diverge. We have a finite number of parameters to adjust, and a finite number of divergences to cancel. The result is a theory that makes unambiguous predictions for physical observables.

5.8 A Simple Analogy

Here's an analogy that captures the spirit of the problem and its resolution.

Suppose you're trying to measure the height of a table. You put a ruler on top and measure from the floor. But the ruler is sitting on a thick book, and the book is sitting on the table. You measure ruler + book, not just the table.

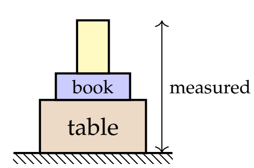


Figure 5.5: Measuring "table height" while a book sits on top.

If you don't know how thick the book is, your measurement of "table height" is ambiguous. It includes an unknown contribution from the book.

But suppose you also measure something else—say, how high the table raises a chair when you put the chair on it. This gives you another combination of table and book heights. From two measurements, you can determine the table height even if you never figure out the book's thickness.

Similarly in QFT: the "bare" parameters include contributions from physics we don't understand (the high-momentum region). But physical observables are combinations of bare parameters and loop corrections. By measuring enough observables, we can determine the physical quantities without ever knowing the bare parameters.

The "book thickness" never cancels exactly—what cancels is its effect on the differences between physical observables. We express everything in terms of measured quantities, and the unknown short-distance physics drops out.

5.9 *Different Types of Infinities*

Not all infinities are created equal. Let me distinguish two types:

Ultraviolet (UV) divergences come from the high-momentum (short-distance) region. These are the main subject of renormalization. They arise because we integrate to $k \to \infty$.

Infrared (IR) divergences come from the low-momentum (long-distance) region. They arise in theories with massless particles (like photons) and are handled differently—they cancel between different contributions to physical observables, or are absorbed into definitions of "soft" radiation.

We'll focus on UV divergences. These are the conceptually challenging ones, the ones that led to the crisis in 1930s QED and the eventual triumph of renormalization.

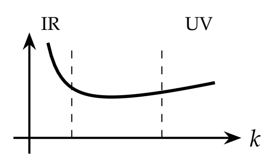


Figure 5.6: IR divergences at small k, UV divergences at large k.

5.10 *The Historical Crisis*

In the 1930s, physicists tried to compute quantum corrections in QED and found infinities everywhere. Dirac, Heisenberg, and Pauli all struggled with this. Some thought quantum electrodynamics was fundamentally inconsistent. Dirac famously said that the theory had "insoluble difficulties."

The resolution came after World War II. In 1947, Willis Lamb measured a tiny splitting in the hydrogen spectrum (the Lamb shift) that couldn't be explained by the Dirac equation alone—it required quantum corrections. This spurred a burst of theoretical activity.

Tomonaga, Schwinger, and Feynman independently developed methods to handle the infinities. Their key insight: the infinities can be absorbed into a redefinition of the mass and charge. What you measure isn't the "bare" mass or charge but the "renormalized" values, which include all the quantum corrections.

Dyson showed that these three approaches were equivalent and that the procedure worked to all orders in perturbation theory. QED was saved. The calculated Lamb shift agreed well with experiment.

5.11 A Preview of Renormalization

Here's the plan—the logic we'll develop over the coming chapters.

First, we introduce a temporary fix: a cutoff Λ that makes all integrals finite. The theory now depends on an artificial parameter, which is uncomfortable, but at least we can compute.

Next, we discover something remarkable: all the infinities have the same structure. They can be absorbed into redefinitions of a finite number of parameters—the mass, the coupling constant, and the normalization of the field. We add counterterms to the Lagrangian that exactly cancel the divergences, with coefficients that depend on Λ.

Then we fix the counterterms by insisting that physical quantities—the measured mass, the measured coupling—take their experimental values. This determines everything.

Finally, we take $\Lambda \to \infty$. With counterterms in place, this limit exists. The Λ-dependence of the counterterms exactly cancels the Λ-dependence of the divergent integrals. Physical predictions are Λ-independent.

The miracle is that only a finite number of counterterms are needed. Once we fix them using a finite number of measurements, everything else is predicted. We don't get to adjust anything further.

5.12 Why This Isn't Cheating

You might feel uncomfortable. We had infinities. We added terms to cancel them. Isn't this just sweeping problems under the rug?

In a sense, yes. We're parameterizing our ignorance of short-distance physics into a finite number of constants. But here's why it's not cheating:

1. **Predictions work.** QED predictions agree with experiment to astonishing precision. The anomalous magnetic moment of the electron is predicted to 10 significant figures, matching measurements. Let me say that again: *ten significant figures*. That's like predicting

the distance from New York to Los Angeles to within the width of a human hair. No other theory in the history of science has achieved this precision.

2. **The arbitrariness is expected.** We *know* our theory isn't complete at arbitrarily short distances. We shouldn't expect it to predict everything from first principles. The bare parameters encode physics we don't know.

3. **The structure is constrained.** Only certain types of infinities can appear. We can't add arbitrary counterterms—only those with the structure allowed by the symmetries and power counting. The theory isn't infinitely adjustable.

4. **It's falsifiable.** Once we fix the free parameters, everything else is predicted. If the predictions were wrong, the theory would be falsified. It isn't.

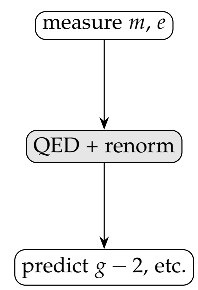


Figure 5.7: Renormalization: a finite number of inputs (measured masses and couplings) determines infinitely many predictions.

5.13 What Are We Really Learning?

Here's a deeper way to think about it. Quantum field theory at low energies doesn't care about the details of high-energy physics. Whatever happens at the Planck scale, or whatever replaces our theory at 10^{15} GeV, its effects on low-energy physics are encoded in just a few parameters: masses and couplings.

This is the principle of *decoupling*: physics at different scales separates. You don't need to know quantum gravity to compute atomic spectra. The high-energy physics is "integrated out" and its effects are absorbed into the parameters of the low-energy theory.

The divergences are telling us this decoupling isn't perfect—the integral tries to sum over all scales, including scales where our theory is wrong. But the structure of the divergences is such that we can compensate for our ignorance with a finite number of parameters.

This isn't a bug; it's a feature. It explains why we can do physics at all without knowing everything. Each energy scale can be understood in isolation, with the effects of other scales parameterized by a handful of constants.

5.14 What Comes Next

In the next two chapters, we'll make this precise:

Chapter 6: Regularization. We'll learn dimensional regularization, the standard tool for making divergent integrals finite in a controlled way. We'll see how to evaluate the key integrals systematically.

Chapter 7: What Renormalization Really Means. We'll connect the particle physics approach (counterterms, renormalization conditions) with the Wilsonian approach (integrating out high energies). If you've learned renormalization from a condensed matter perspective, this is where the two pictures come together.

Then we'll apply these tools to QED, computing the three fundamental divergent diagrams and extracting physical predictions.

The infinities of quantum field theory troubled the greatest physicists of the 20th century. Dirac never fully accepted renormalization, calling it "just a stop-gap procedure." Feynman, who won the Nobel Prize for his contributions to QED, described renormalization as "a shell game" and "dippy." Yet the theory works. It makes predictions of staggering precision. Perhaps the discomfort reflects not a flaw in renormalization but a limitation in our intuitions, honed on classical physics where point particles don't require careful handling. Or perhaps future theories will resolve the infinities more satisfactorily. String theory, for instance, has no ultraviolet divergences—the extended nature of strings smooths out short-distance singularities. But for now, renormalization is our best tool for making sense of quantum field theory in the continuum, and its success speaks for itself.

6
Regularization: Taming Infinity

In the last chapter, we saw that loop integrals in quantum field theory diverge. Before we can do anything sensible with these infinities—before we can cancel them, subtract them, or absorb them into parameters—we need to make them finite. We need to *regulate* them.

A regulator is a prescription that makes divergent integrals finite while preserving as much of the theory's structure as possible. Once the integrals are finite, we can manipulate them algebraically, identify the divergent parts, and eventually remove the regulator.

There are many ways to regulate. We could cut off the momentum integral at some Λ (hard cutoff). We could add heavy fictitious particles that cancel the divergence at high momenta (Pauli-Villars). We could put the theory on a lattice, making momenta automatically bounded.

But the method that has become standard in modern particle physics is *dimensional regularization*. Instead of cutting off momenta, we analytically continue the spacetime dimension from $d = 4$ to $d = 4 - \epsilon$, where ϵ is small. The divergences appear as poles in $1/\epsilon$ rather than as infinite numbers.

This might sound bizarre—what does it mean to do physics in $4 - \epsilon$ dimensions? But here's the strange thing: it works. And not just approximately. Dimensional regularization preserves gauge invariance, Lorentz invariance, and other symmetries that are essential for QED and the Standard Model. It's become the tool of choice for practical calculations, even though nobody can draw you a picture of 3.99-dimensional spacetime.

6.1 Why Dimensional Regularization?

Before diving in, let me explain why we use this strange technique.

Hard cutoff problems. The simplest regulator is a momentum cutoff: just stop the integral at $|k| = \Lambda$. But this breaks Lorentz invariance (it treats time and space differently in general) and gauge invariance (it doesn't respect the Ward identities of QED). Breaking these symmetries introduces spurious terms that complicate or invalidate calculations.

Pauli-Villars. You can preserve Lorentz invariance by adding fictitious heavy particles that contribute with opposite signs. But this requires care and becomes awkward for gauge theories.

Dimensional regularization preserves symmetries. Formally continuing to d dimensions respects Lorentz invariance (there's no special direction) and gauge invariance (the symmetry algebra works in any dimension). The only "violation" is that the theory doesn't physically exist in $d \neq 4$, but this is fine—it's a computational trick, and we take $d \to 4$ at the end.

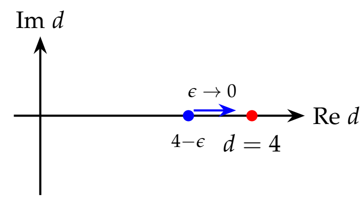


Figure 6.1: We work at $d = 4 - \epsilon$, then take $\epsilon \to 0$ at the end.

6.2 *The Basic Idea*

Consider a one-loop integral in d dimensions:

$$I = \int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 - m^2)^n} \tag{6.1}$$

In $d = 4$, this might diverge. But for d sufficiently small, it converges. We'll compute it for general d, express the answer as a function of d, then analytically continue to $d = 4 - \epsilon$.

The divergence will appear as a pole at $\epsilon = 0$. Something like:

$$I = \frac{A}{\epsilon} + B + O(\epsilon) \tag{6.2}$$

where A and B are finite coefficients. The $1/\epsilon$ is the "regularized infinity."

What Does Non-Integer Dimension Mean?

You might wonder: what does it mean to integrate in 3.99 dimensions? After all, we can't draw 3.99 coordinate axes!

The answer is that we never actually need to picture d-dimensional space. What matters is that the *formulas* for integrals can be analytically continued to any complex d.

Consider the volume of a d-dimensional sphere of radius R:

$$V_d(R) = \frac{\pi^{d/2}}{\Gamma(d/2+1)} R^d \tag{6.3}$$

For $d = 1$, this gives $2R$ (the length of an interval). For $d = 2$, it gives πR^2 (area of a circle). For $d = 3$, it gives $\frac{4}{3}\pi R^3$ (volume of a ball).

But nothing stops us from evaluating the formula at $d = 3.99$. The result isn't the "volume of a 3.99-dimensional sphere"—that doesn't exist. It's just a number defined by analytic continuation of a formula that makes sense for positive integers.

This is the spirit of dimensional regularization: we use formulas that are valid for integer d, analytically continue them to $d = 4 - \epsilon$, and only at the end take $\epsilon \to 0$ to return to physical spacetime.

Should this bother you? In a sense, yes—we're claiming that the answer to a question that doesn't make physical sense ("what happens in 3.99 dimensions?") tells us something true about our 4-dimensional world. The justification is purely pragmatic: it works. The results agree with experiment. If that feels unsatisfying, you're in good company. But as long as no one tells the integrals that they're not allowed to live in fractional dimensions, they seem perfectly happy to give us sensible answers.

Why Does Lower Dimension Help?

Intuitively, reducing the dimension reduces the "phase space" for high-momentum modes. An integral $\int d^d k$ has fewer high-momentum modes in lower dimension.

More precisely, consider

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{k^{2n}} \sim \int_0^\Lambda k^{d-1} dk \cdot k^{-2n} = \int_0^\Lambda k^{d-2n-1} dk \tag{6.4}$$

This converges at large k if $d - 2n - 1 < -1$, i.e., if $d < 2n$. By lowering d below this threshold, we make the integral converge.

For a quadratically divergent integral ($n = 1$), we need $d < 2$. For a logarithmically divergent integral ($n = 2$), any $d < 4$ works. So by taking $d = 4 - \epsilon$ with $\epsilon > 0$, we ensure convergence, and then we track how the answer behaves as $\epsilon \to 0$.

6.3 *Angular Integrals in d Dimensions*

First, let's establish some geometry. In d dimensions, a vector k has d components. The "radial" part is $|k| = \sqrt{k_1^2 + \cdots + k_d^2}$.

The volume element is:

$$d^d k = |k|^{d-1} d|k| \, d\Omega_d \tag{6.5}$$

where $d\Omega_d$ is the angular element on the $(d-1)$-sphere.

The total solid angle in d dimensions is:

$$\Omega_d = \int d\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)} \tag{6.6}$$

This formula involves the Gamma function, which generalizes factorials to non-integer arguments.

d	Ω_d
1	2
2	2π
3	4π
4	$2\pi^2$
d	$\frac{2\pi^{d/2}}{\Gamma(d/2)}$

Table 6.1: Solid angle in d dimensions. For $d = 3$, we recover the familiar 4π steradians.

For integrals that only depend on $|k|$, we can do the angular integral immediately:

$$\int \frac{d^d k}{(2\pi)^d} f(k^2) = \frac{\Omega_d}{(2\pi)^d} \int_0^\infty d|k|\,|k|^{d-1} f(k^2)$$
$$= \frac{1}{(4\pi)^{d/2}} \frac{1}{\Gamma(d/2)} \int_0^\infty dk^2\,(k^2)^{d/2-1} f(k^2) \tag{6.7}$$

This reduces everything to a one-dimensional integral.

6.4 The Master Formula

The key result for dimensional regularization is the following integral:

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2-\Delta)^n} = \frac{i(-1)^n}{(4\pi)^{d/2}} \frac{\Gamma(n-d/2)}{\Gamma(n)} \frac{1}{\Delta^{n-d/2}} \tag{6.8}$$

where we've Wick-rotated to Euclidean space (hence the factor of i) and Δ is some combination of masses and external momenta.

Let me derive this formula, as understanding the derivation will help you use it.

Wick Rotation: From Minkowski to Euclidean Space

Before deriving the master formula, we need to understand Wick rotation. This is a crucial step that turns oscillating Minkowski integrals into convergent Euclidean integrals.

In Minkowski space, $k^2 = k_0^2 - \mathbf{k}^2$. The integral over k_0 runs along the real axis, and the propagator $1/(k^2 - m^2 + i\epsilon)$ has poles near the real axis (at $k_0 = \pm\sqrt{\mathbf{k}^2 + m^2} \mp i\epsilon$).

The $i\epsilon$ prescription tells us the poles are slightly displaced: the positive-energy pole is just below the real axis, and the negative-energy pole is just above. This means we can rotate the integration contour from the real k_0 axis to the imaginary axis without crossing any poles.

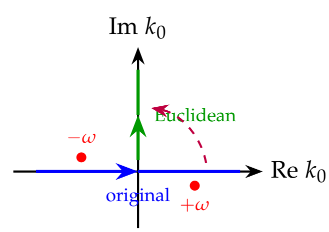


Figure 6.2: Wick rotation: the contour along the real axis (blue) is deformed to the imaginary axis (green). The $i\epsilon$ prescription places poles (red) so no poles are crossed.

Setting $k_0 = ik_0^E$ (Euclidean), we have:

$$k^2 = k_0^2 - \mathbf{k}^2 = -(k_0^E)^2 - \mathbf{k}^2 = -k_E^2 \tag{6.9}$$
$$dk_0 = i\,dk_0^E \tag{6.10}$$

So the Minkowski integral becomes:

$$\int \frac{dk_0\, d^{d-1}\mathbf{k}}{(2\pi)^d} \frac{1}{(k^2-\Delta)^n} = i\int \frac{d^d k_E}{(2\pi)^d} \frac{1}{(-k_E^2-\Delta)^n} = \frac{i(-1)^n}{1} \int \frac{d^d k_E}{(2\pi)^d} \frac{1}{(k_E^2+\Delta)^n} \tag{6.11}$$

The Euclidean integral is over a positive-definite k_E^2, so it's much better behaved—no oscillations, no poles on the integration contour.

Derivation of the Master Formula

Start with the Euclidean integral (after Wick rotation):

$$I_E = \int \frac{d^d k_E}{(2\pi)^d} \frac{1}{(k_E^2 + \Delta)^n} \tag{6.12}$$

Using the angular integral formula:

$$I_E = \frac{1}{(4\pi)^{d/2}} \frac{1}{\Gamma(d/2)} \int_0^\infty dk_E^2 \, (k_E^2)^{d/2-1} \frac{1}{(k_E^2 + \Delta)^n} \tag{6.13}$$

Substitute $u = k_E^2/\Delta$, so $dk_E^2 = \Delta \, du$:

$$I_E = \frac{\Delta^{d/2-n}}{(4\pi)^{d/2}\Gamma(d/2)} \int_0^\infty du \, u^{d/2-1}(1+u)^{-n} \tag{6.14}$$

The remaining integral is a Beta function:

$$\int_0^\infty du \, u^{a-1}(1+u)^{-a-b} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = B(a,b) \tag{6.15}$$

with $a = d/2$ and $b = n - d/2$. So:

$$I_E = \frac{\Delta^{d/2-n}}{(4\pi)^{d/2}} \frac{\Gamma(n-d/2)}{\Gamma(n)} \tag{6.16}$$

Converting back to Minkowski space (the Wick rotation gives a factor of $i(-1)^n$):

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 - \Delta + i\epsilon)^n} = \frac{i(-1)^n}{(4\pi)^{d/2}} \frac{\Gamma(n-d/2)}{\Gamma(n)} \frac{1}{\Delta^{n-d/2}} \tag{6.17}$$

This is the master formula (6.8).

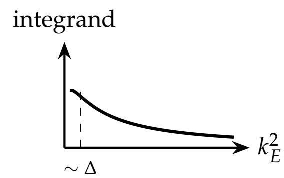


Figure 6.3: The integrand peaks near $k_E^2 \sim \Delta$ and falls off at large momenta.

6.5 *Where Do the Divergences Go?*

The Gamma function $\Gamma(z)$ has poles at $z = 0, -1, -2, \ldots$. In our formula, the argument is $n - d/2$.

For the tadpole integral ($n = 1$) in $d = 4 - \epsilon$:

$$\Gamma(1 - 2 + \epsilon/2) = \Gamma(-1 + \epsilon/2) \tag{6.18}$$

This has a pole at $\epsilon = 0$ because $\Gamma(-1)$ is infinite.

Near $z = -1$:

$$\Gamma(-1+x) = \frac{1}{-1+x} \cdot \Gamma(x) \approx -\frac{1}{1-x} \cdot \frac{1}{x}\Gamma(1+x) \approx -\frac{1}{x} \tag{6.19}$$

Wait, let me be more careful. Using $\Gamma(z+1) = z\Gamma(z)$:

$$\Gamma(-1+\epsilon/2) = \frac{\Gamma(\epsilon/2)}{(-1+\epsilon/2)} = \frac{\Gamma(\epsilon/2)}{-1+\epsilon/2} \tag{6.20}$$

And $\Gamma(\epsilon/2) \approx 2/\epsilon - \gamma_E + O(\epsilon)$, where $\gamma_E \approx 0.5772$ is the Euler-Mascheroni constant.

So the divergence (the pole at $\epsilon = 0$) is encoded in the Gamma function. The integral is no longer literally infinite; it's a meromorphic function of ϵ with a pole at $\epsilon = 0$.

6.6 Expanding Around $d = 4$

Let's work out the tadpole integral explicitly. In $d = 4 - \epsilon$:

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2 - m^2} = \frac{i}{(4\pi)^{d/2}} \Gamma(1 - d/2)(m^2)^{d/2-1} \tag{6.21}$$

With $d = 4 - \epsilon$:

- $(4\pi)^{d/2} = (4\pi)^{2-\epsilon/2} = (4\pi)^2 \cdot (4\pi)^{-\epsilon/2} = 16\pi^2 \cdot e^{-\epsilon \ln(4\pi)/2}$
- $\Gamma(1 - d/2) = \Gamma(-1 + \epsilon/2)$
- $(m^2)^{d/2-1} = (m^2)^{1-\epsilon/2} = m^2 \cdot (m^2)^{-\epsilon/2}$

Using $\Gamma(-1 + \epsilon/2) = -2/\epsilon + \gamma_E - 1 + O(\epsilon)$:

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2 - m^2} = \frac{im^2}{16\pi^2} \left[\frac{2}{\epsilon} + 1 - \gamma_E - \ln \frac{m^2}{4\pi} + O(\epsilon) \right] \tag{6.22}$$

The $2/\epsilon$ is the divergent part. The rest is finite.

6.7 Feynman Parameters

Most loop integrals aren't as simple as the tadpole—they have multiple propagators with different momenta. The standard technique to handle this is *Feynman parameterization*.

The Feynman Parameter Trick

The basic identity is:

$$\frac{1}{A_1 A_2 \cdots A_n} = (n-1)! \int_0^1 dx_1 \cdots dx_n \, \delta\left(\sum x_i - 1\right) \frac{1}{(x_1 A_1 + \cdots + x_n A_n)^n} \tag{6.23}$$

For two propagators:

$$\frac{1}{AB} = \int_0^1 dx \, \frac{1}{[xA + (1-x)B]^2} \tag{6.24}$$

This combines multiple propagators into a single denominator, at the cost of introducing an integral over Feynman parameters x_i.

Why Does This Help?

The magic of Feynman parameters is that they let us complete the square in the loop momentum. With multiple propagators, the denominator is a mess like

$$\frac{1}{(k^2 - m_1^2)((k-p)^2 - m_2^2)} \tag{6.25}$$

where k appears in different ways in each factor.

After applying Feynman parameters, we get a single denominator like

$$\frac{1}{[k^2 - 2k \cdot (\text{something}) + (\text{constant})]^2} \tag{6.26}$$

This is a quadratic form in k. We can complete the square by shifting $k \to k + (\text{something})$, after which the denominator becomes

$$\frac{1}{(k^2 - \Delta)^2} \tag{6.27}$$

where Δ depends on the Feynman parameters and external momenta but not on k.

Now the k integral is in standard form! We can apply the master formula directly.

The price we pay is an additional integral over Feynman parameters at the end. But this is usually much easier than the original multi-propagator integral.

Physical Intuition

Is there any physics in the Feynman parameters? Sort of. The parameter x roughly corresponds to how the loop momentum is "distributed" between the two propagators. When $x \approx 0$, the combined denominator is dominated by B (the second propagator). When $x \approx 1$, it's dominated by A (the first propagator). The integral over x sums over all ways of distributing the "blame" for the loop.

In position space, Feynman parameters have an even more direct interpretation: they parameterize how the proper time is divided between different propagators along a particle's world line. But for our momentum-space calculations, it's enough to view them as a mathematical tool.

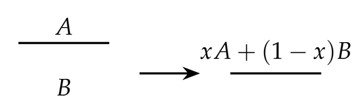
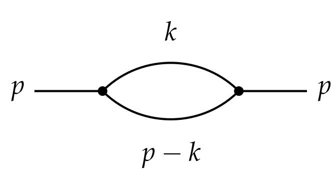


Figure 6.4: Feynman parameters combine multiple propagators into one.

Example: The Fish Diagram

Let's compute the "fish" diagram—the one-loop correction to $2 \to 2$ scattering in ϕ^4 theory:



The integral is:

$$I(p^2) = \int \frac{d^dk}{(2\pi)^d} \frac{1}{(k^2 - m^2)((p-k)^2 - m^2)} \tag{6.28}$$

Apply Feynman parameterization (6.24):

$$I(p^2) = \int_0^1 dx \int \frac{d^d k}{(2\pi)^d} \frac{1}{[x(k^2 - m^2) + (1-x)((p-k)^2 - m^2)]^2} \quad (6.29)$$

The denominator is:

$$D = x(k^2 - m^2) + (1-x)((p-k)^2 - m^2) \quad (6.30)$$

$$= k^2 - 2(1-x)p \cdot k + (1-x)p^2 - m^2 \quad (6.31)$$

Complete the square by shifting $k \to k + (1-x)p$:

$$D = k^2 + (1-x)p^2 - (1-x)^2 p^2 - m^2 = k^2 - \Delta \quad (6.32)$$

where $\Delta = m^2 - x(1-x)p^2$.

The shifted integral is:

$$I(p^2) = \int_0^1 dx \int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 - \Delta)^2} \quad (6.33)$$

Now we can use the master formula (6.8) with $n = 2$:

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 - \Delta)^2} = \frac{i}{(4\pi)^{d/2}} \Gamma(2 - d/2) \frac{1}{\Delta^{2-d/2}} \quad (6.34)$$

In $d = 4 - \epsilon$, $\Gamma(2 - d/2) = \Gamma(\epsilon/2) = 2/\epsilon - \gamma_E + O(\epsilon)$. So:

$$I(p^2) = \frac{i}{16\pi^2} \left[\frac{2}{\epsilon} - \gamma_E + \ln(4\pi) \right] \int_0^1 dx\, \Delta^{-\epsilon/2} + O(\epsilon) \quad (6.35)$$

To leading order in ϵ:

$$\Delta^{-\epsilon/2} = 1 - \frac{\epsilon}{2} \ln \Delta + O(\epsilon^2) \quad (6.36)$$

So the integral is:

$$I(p^2) = \frac{i}{16\pi^2} \left[\frac{2}{\epsilon} - \gamma_E + \ln(4\pi) - \int_0^1 dx\, \ln(m^2 - x(1-x)p^2) \right] + O(\epsilon) \quad (6.37)$$

The $2/\epsilon$ pole is the divergence. The rest is finite and depends on the kinematics through p^2.

6.8 The $\overline{MS}$ Scheme

The constants γ_E and $\ln(4\pi)$ that accompany $1/\epsilon$ are annoying but universal. They appear in every loop integral. It's convenient to absorb them into the definition of the divergent part.

The $\overline{MS}$ *scheme* ("MS-bar") defines the subtraction to remove not just $1/\epsilon$ but the combination:

$$\frac{2}{\bar{\epsilon}} \equiv \frac{2}{\epsilon} - \gamma_E + \ln(4\pi) \quad (6.38)$$

In $\overline{\text{MS}}$, you replace $2/\epsilon \to 2/\bar{\epsilon}$ and the leftover finite parts are simpler.

This is purely a convention for how we split "divergent" from "finite" parts. Physical results don't depend on the scheme—only intermediate steps do.

MS
subtract $\frac{1}{\epsilon}$

$\overline{\text{MS}}$
subtract $\frac{2}{\epsilon} - \gamma_E + \ln 4\pi$

Figure 6.5: The $\overline{\text{MS}}$ scheme subtracts extra constants along with the pole.

6.9 The Mass Scale μ

There's a subtlety we've glossed over. In d dimensions, the coupling constant λ has dimension $(4-d) = \epsilon$. To keep the action dimensionless, we must introduce a mass scale μ and write:

$$\lambda \to \lambda \mu^{\epsilon} \tag{6.39}$$

where λ is now dimensionless.

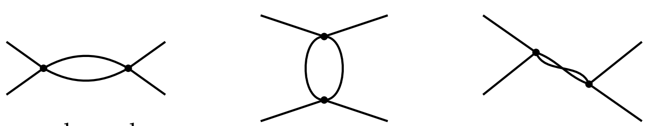
This μ is the *renormalization scale*. It appears in loop integrals through logarithms like $\ln(\mu^2/m^2)$ or $\ln(\mu^2/p^2)$.

The physical results can't depend on our arbitrary choice of μ. This independence leads to the *renormalization group*, which we'll explore in Chapter 12.

6.10 A Complete Worked Example

Let me work through the one-loop correction to the 4-point function in ϕ^4 theory completely.

The amplitude at one loop includes the "fish" diagram in three channels:



s-channel t-channel u-channel

The tree-level amplitude is $\mathcal{M}_0 = -\lambda$.

The one-loop correction from the s-channel is:

$$i\mathcal{M}_s = \frac{(-i\lambda)^2}{2} \cdot I(s) \tag{6.40}$$

where $s = (p_1 + p_2)^2$ and $I(s)$ is the bubble integral we computed.

Adding all three channels:

$$\mathcal{M}_{\text{1-loop}} = -\frac{\lambda^2}{2}\left[I(s) + I(t) + I(u)\right] \tag{6.41}$$

In $\overline{\text{MS}}$:

$$I(s) = \frac{i}{16\pi^2}\left[\frac{2}{\bar{\epsilon}} - \int_0^1 dx \, \ln\frac{m^2 - x(1-x)s}{\mu^2}\right] \tag{6.42}$$

The divergent part is:

$$\mathcal{M}^{\text{div}}_{\text{1-loop}} = -\frac{3\lambda^2}{32\pi^2} \cdot \frac{2}{\bar{\epsilon}} = -\frac{3\lambda^2}{16\pi^2\bar{\epsilon}} \tag{6.43}$$

The factor of 3 comes from the three channels.

6.11 What Have We Accomplished?

We've converted divergent integrals into expressions with poles at $\epsilon = 0$. The "infinity" is now $1/\epsilon$, which is a well-defined mathematical object (it's large when ϵ is small, but it's finite for any nonzero ϵ).

We can now:

1. Identify the divergent part (coefficients of $1/\epsilon$).
2. Identify the finite part (terms independent of $1/\epsilon$).
3. Manipulate these algebraically.
4. Eventually, add counterterms to cancel the $1/\epsilon$ poles.

The regularization is temporary scaffolding. After renormalization, we'll take $\epsilon \to 0$ and the result will be finite.

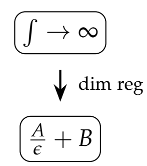


Figure 6.6: Dimensional regularization turns infinities into poles.

6.12 Important Integrals for QED

When we compute QED loop diagrams, we'll need integrals with tensor structure—numerators involving k^μ or $k^\mu k^\nu$. Here are the key results.

Symmetric integration: Odd powers of k integrate to zero by symmetry:

$$\int \frac{d^dk}{(2\pi)^d} \frac{k^\mu}{(k^2 - \Delta)^n} = 0 \tag{6.44}$$

For even powers:

$$\int \frac{d^dk}{(2\pi)^d} \frac{k^\mu k^\nu}{(k^2 - \Delta)^n} = \frac{g^{\mu\nu}}{d} \int \frac{d^dk}{(2\pi)^d} \frac{k^2}{(k^2 - \Delta)^n} \tag{6.45}$$

The factor $g^{\mu\nu}/d$ comes from the only available tensor structure. Similarly:

$$\int \frac{d^dk}{(2\pi)^d} \frac{k^\mu k^\nu k^\rho k^\sigma}{(k^2 - \Delta)^n} = \frac{g^{\{\mu\nu} g^{\rho\sigma\}}}{d(d+2)} \int \frac{d^dk}{(2\pi)^d} \frac{(k^2)^2}{(k^2 - \Delta)^n} \tag{6.46}$$

where $g^{\{\mu\nu} g^{\rho\sigma\}}$ means the sum over all distinct pairings.

These formulas let us reduce tensor integrals to scalar integrals, which we can evaluate with the master formula.

6.13 *Summary: The Toolkit*

Let me collect what we've learned. The dimensional regularization procedure is this: work in $d = 4 - \epsilon$ dimensions, remembering that couplings acquire factors of μ^ϵ to keep their dimensions straight. When you encounter multiple propagators, combine them with Feynman parameters. Complete the square in the loop momentum to put the integral in standard form. Then apply the master integral—the Wick-rotated Gaussian that gives you Gamma functions. Expand in ϵ: the divergences show up as $1/\epsilon$ poles, with finite parts that depend on your regularization scheme. Finally, do the Feynman parameter integrals, which are usually elementary.

The master integral is:

$$\int \frac{d^dk}{(2\pi)^d} \frac{1}{(k^2 - \Delta)^n} = \frac{i(-1)^n}{(4\pi)^{d/2}} \frac{\Gamma(n - d/2)}{\Gamma(n)} \Delta^{d/2-n} \tag{6.47}$$

With these tools, you can compute any one-loop diagram in ϕ^4 theory or QED. Multi-loop calculations are harder—more Feynman parameters, more integrals—but they follow the same principles. The machinery we've built scales.

In the next chapter, we'll see how to use counterterms to cancel the $1/\epsilon$ poles, and we'll finally connect the particle physics approach to renormalization with the Wilsonian perspective you may know from condensed matter.

Dimensional regularization was developed in the early 1970s by 't Hooft and Veltman, specifically for gauge theories. Previous regulators (cutoffs, Pauli-Villars) broke gauge invariance in awkward ways. 't Hooft and Veltman realized that continuing to d dimensions preserved gauge symmetry because the gauge algebra is the same in any dimension. The method was crucial for proving the renormalizability of the electroweak theory, for which 't Hooft and Veltman shared the 1999 Nobel Prize. Today, dimensional regularization is the standard tool for perturbative calculations in the Standard Model. Its elegance lies in replacing brute-force cutoffs with analytic continuation—trading a hard boundary in momentum space for a soft deformation of spacetime dimension.

7
What Renormalization Really Means

Renormalization is perhaps the most misunderstood idea in modern physics. Some people think it's a trick to hide infinities under the rug. Others think it's deep mathematics. The truth is both stranger and simpler: renormalization is the recognition that physics at different scales decouples, and that what we measure depends on the scale at which we measure it.

We've developed the machinery: Feynman diagrams give us integrals, dimensional regularization makes them finite, and we know which structures diverge. Now we need to understand what to *do* with all this. What are "bare" and "renormalized" parameters? How do counterterms work? And how does all this connect to the Wilsonian picture you may know from condensed matter?

That last question is crucial. If you've learned renormalization from a condensed matter perspective—starting with a UV theory and flowing to the IR—the particle physics approach can seem backwards. It looks like we're going from IR to UV. But we're not. The two pictures describe the same physics from different viewpoints. Understanding their relationship is the key to everything.

7.1 Bare Parameters vs. Physical Parameters

Let's start with a concrete question: what is the mass of the electron?

You might say: $m_e = 0.511$ MeV. That's the measured value. But what mass should we put in the Lagrangian?

The Lagrangian for a free electron contains a mass term:

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m_0)\psi \tag{7.1}$$

The parameter m_0 is called the *bare mass*. It's the number we write down in the fundamental theory.

But when we turn on interactions (couple the electron to photons), quantum corrections modify the relationship between m_0 and the physical mass we measure. The electron is always surrounded by a cloud of virtual photons, virtual electron-positron pairs, etc. What we observe includes all these effects.

The relationship is:

$$m_{\text{phys}} = m_0 + \delta m \tag{7.2}$$

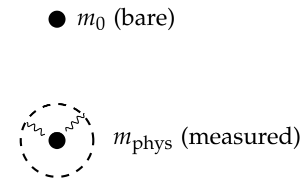


Figure 7.1: The bare mass m_0 and the physical mass m_{phys} differ by quantum corrections.

where δm is the "mass correction" from loop diagrams. As we've seen, δm is divergent—it contains a $1/\epsilon$ pole in dimensional regularization.

So what is m_0? It's not directly measurable. If δm is infinite (in the $\epsilon \to 0$ limit), then m_0 must also be infinite—in just the right way to cancel the divergence and leave a finite m_{phys}.

This sounds crazy. We're adding two infinite quantities to get a finite result? But remember: in dimensional regularization, "infinite" means "has a pole at $\epsilon = 0$." We never actually set $\epsilon = 0$ until the end. At finite ϵ, both m_0 and δm are finite, and the cancellation is perfectly well-defined.

7.2 *The Counterterm Approach*

Here's how we organize the calculation systematically.

Write the Lagrangian as:

$$\mathcal{L} = \mathcal{L}_{\text{physical}} + \mathcal{L}_{\text{counterterm}} \tag{7.3}$$

The "physical" part has the same form as the original Lagrangian but with *renormalized* (physical) parameters. The "counterterm" part contains new terms that cancel the divergences.

For ϕ^4 theory:

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 + \frac{1}{2}\delta_Z(\partial\phi)^2 - \frac{1}{2}\delta_m\phi^2 - \frac{\delta_\lambda}{4!}\phi^4 \tag{7.4}$$

The first three terms use the *renormalized* mass m and coupling λ—the values we measure. The counterterms δ_Z, δ_m, δ_λ are chosen to cancel the divergences from loop diagrams.

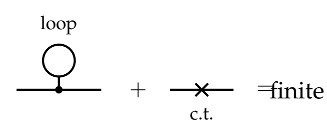


Figure 7.2: Divergent loop plus counterterm (c.t.) gives a finite result.

The counterterms are treated as additional interaction vertices in Feynman diagrams. At one-loop, they contribute at tree level, canceling the one-loop divergences. At two-loop, they appear in one-loop diagrams, canceling two-loop divergences. And so on.

7.3 *Renormalization Conditions*

The counterterms aren't arbitrary—they're fixed by *renormalization conditions*. These are physical requirements that define what we mean by the "mass" and "coupling."

For the mass, a natural condition is:

The pole of the propagator occurs at $p^2 = m^2$.

The full propagator, including loop corrections, is:

$$G(p^2) = \frac{i}{p^2 - m_0^2 - \Sigma(p^2)} \tag{7.5}$$

where $\Sigma(p^2)$ is the self-energy (sum of all 1PI corrections to the propagator). We require:

$$m_0^2 + \Sigma(m^2) = m^2 \tag{7.6}$$

This determines δ_m in terms of Σ.

For the coupling, we might require:

The 4-point function at some specific kinematic point equals $-i\lambda$.

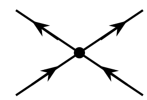


Figure 7.3: The coupling is defined at a specific kinematic point (renormalization point).

For instance: at the symmetric point $s = t = u = 4m^2/3$ (where all Mandelstam variables are equal), require the amplitude to equal $-\lambda$.

Different choices of renormalization conditions define different *renormalization schemes*. Common schemes include:

- **On-shell**: Physical mass is the pole; coupling is defined at a physical kinematic point.
- $\overline{\textbf{MS}}$: Counterterms are just the $1/\bar{\epsilon}$ poles—no finite parts.

Physical predictions don't depend on the scheme. Different schemes just organize the calculation differently.

7.4 *Here's the Miracle*

In ϕ^4 theory, we found that only diagrams with $E = 0, 2, 4$ external lines diverge (Chapter 5). These correspond to:

- $E = 0$: Vacuum energy (we ignore this).
- $E = 2$: Mass and wave-function renormalization.
- $E = 4$: Coupling renormalization.

We need exactly three counterterms: δ_Z, δ_m, δ_λ. These are the same three parameters that appear in the original Lagrangian.

No matter how many loops we compute, no matter how complicated the diagrams, the divergences always have one of these three structures. We can always cancel them with the same three counterterms.

This is what makes ϕ^4 theory *renormalizable*: all divergences can be absorbed into redefinitions of the existing parameters. No new parameters are needed.

7.5 Non-Renormalizable Theories

Compare this to a theory with a ϕ^6 interaction:

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{g}{6!}\phi^6 \tag{7.7}$$

Power counting gives $D = 4 - E + 2L(3-1) = 4 - E + 4L$ (roughly—I'm being schematic). At one loop, diagrams with $E = 6$ are divergent. We'd need a ϕ^6 counterterm. Fine, the theory already has one.

But at higher loops, $E = 8$ diagrams diverge. We need a ϕ^8 counterterm. Then ϕ^{10}, ϕ^{12}, ... We need infinitely many counterterms.

Such theories are called *non-renormalizable*. They're not useless—they can be treated as *effective theories* valid below some energy scale—but they're not fundamental in the traditional sense. We'll return to this perspective later.

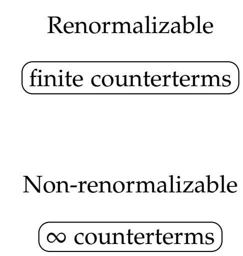


Figure 7.4: Renormalizable theories need finitely many counterterms; non-renormalizable need infinitely many.

7.6 Two Pictures of the Same Mountain

Now let's connect this to the Wilsonian picture.

In the Wilsonian approach (familiar from condensed matter), you start with a theory defined at a UV cutoff Λ. You then "integrate out" high-momentum modes—the degrees of freedom with $|k| > \Lambda/b$ for some $b > 1$—and ask what effective theory describes the remaining low-momentum modes.

The result is that the parameters of the effective theory change. The coupling "flows" under this coarse-graining. Starting from bare parameters at scale Λ, you flow to effective parameters at scale Λ/b, then Λ/b^2, etc.

This is the renormalization group (RG) flow, and it goes from UV to IR. You start with the microscopic theory and flow toward long wavelengths.

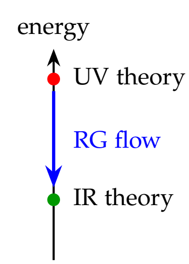


Figure 7.5: Wilsonian RG: flow from UV to IR by integrating out high-momentum modes.

In particle physics, the logic seems reversed. We *measure* at low energies (compared to any cutoff). We measure $m_e = 0.511$ MeV and $\alpha = 1/137$ in laboratory experiments. These are the IR parameters. Then we ask: what happens at higher energies?

The answer is that α runs. At higher energies, α is larger. The coupling increases as we probe shorter distances. This is described by the beta function:

$$\mu\frac{d\alpha}{d\mu} = \beta(\alpha) > 0 \quad \text{(in QED)} \tag{7.8}$$

This seems to say we're going "from IR to UV." We start with the low-energy coupling and extrapolate to high energy. Isn't that backwards?

7.7 *The Resolution*

No, it's not backwards. Let me explain.

The RG flow is always UV $\to$ IR. The physics at high energies determines the physics at low energies, not the other way around. If you had a complete UV theory, you could calculate all the IR parameters.

But in practice, we don't have the complete UV theory. We don't know physics at the Planck scale. So we work differently:

1. We *parameterize* the UV physics by a few numbers: the bare parameters at some high scale.
2. We *measure* at accessible (IR) scales.
3. We *use the RG* to relate these: given the UV parameters, the RG tells us the IR parameters.
4. We *invert* this: given the IR parameters (measured), we can infer what UV parameters would produce them.

The "IR $\to$ UV" appearance comes from step 4. We're not claiming that IR physics determines UV physics. We're asking: *given what we measured at low energy, what must be true at high energy for this to work?*

The beta function works both ways. Running α from low to high energy isn't claiming the flow goes that direction—it's using the known flow to extrapolate.

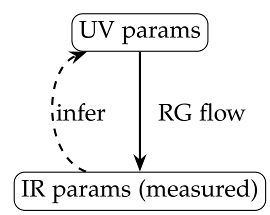


Figure 7.6: RG flows UV $\to$ IR. We infer UV from IR by inverting this flow.

7.8 *Matching at Scale μ*

Here's another way to see the connection.

In the Wilsonian picture, imagine we have a theory at cutoff Λ. We integrate out modes between Λ and μ, getting an effective theory at scale μ with different parameters.

In the particle physics approach, we work with an effective theory at scale μ directly. The "renormalized parameters" are the effective parameters at that scale. Different choices of μ give different effective theories, related by the RG.

The renormalization conditions are like "matching conditions." When we say "the coupling at scale μ is $\lambda(\mu)$," we're specifying the effective theory at that scale.

The counterterms absorb the difference between:

- The "bare" theory at cutoff Λ (the UV).
- The effective theory at scale μ (where we define renormalized parameters).

In dimensional regularization, $\Lambda \to \infty$ is taken implicitly. The $1/\epsilon$ poles are what's left of the UV divergence. The counterterms cancel these poles, matching the bare theory to the effective theory at scale μ.

7.9 A Concrete Example: ϕ^4 Running Coupling

Let's see this explicitly in ϕ^4 theory.

The one-loop correction to the 4-point function is (from Chapter 6):

$$\mathcal{M} = -\lambda + \frac{3\lambda^2}{32\pi^2}\left[\frac{2}{\bar{\epsilon}} + \text{finite terms involving } \ln\frac{\mu^2}{s}, \ln\frac{\mu^2}{t}, \ln\frac{\mu^2}{u}\right] \tag{7.9}$$

The counterterm is:

$$\delta_\lambda = \frac{3\lambda^2}{16\pi^2\bar{\epsilon}} + (\text{scheme-dependent finite part}) \tag{7.10}$$

The renormalized amplitude becomes:

$$\mathcal{M}_{\text{ren}} = -\lambda(\mu) + \frac{3\lambda^2(\mu)}{32\pi^2}\left[\ln\frac{\mu^2}{s} + \ln\frac{\mu^2}{t} + \ln\frac{\mu^2}{u} + \text{const}\right] \tag{7.11}$$

Notice: the amplitude depends on μ through $\lambda(\mu)$ and the logarithms. But the *physical* amplitude can't depend on our arbitrary choice of μ.

Demanding $d\mathcal{M}/d\mu = 0$ at fixed physical momenta gives:

$$\mu\frac{d\lambda}{d\mu} = \frac{3\lambda^2}{16\pi^2} + O(\lambda^3) \tag{7.12}$$

This is the *beta function*. It tells us how the coupling runs with scale.

Solving this equation:

$$\frac{1}{\lambda(\mu_2)} = \frac{1}{\lambda(\mu_1)} - \frac{3}{16\pi^2}\ln\frac{\mu_2}{\mu_1} \tag{7.13}$$

At higher scales ($\mu_2 > \mu_1$), λ increases (since we subtract a positive number from $1/\lambda$). This is consistent with the Wilsonian picture: as we integrate out high-momentum modes, the effective coupling at lower scales is smaller.

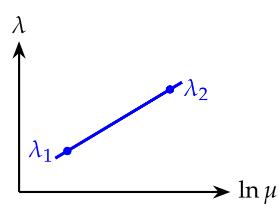


Figure 7.7: The ϕ^4 coupling λ increases with scale μ. At higher μ, the coupling is larger.

7.10 The Wilsonian Effective Action

Let me state the Wilsonian picture more precisely.

Define the *Wilsonian effective action* $S_{\text{eff}}[\phi; \Lambda]$ by integrating out all modes with $|k| > \Lambda$:

$$e^{-S_{\text{eff}}[\phi_<;\Lambda]} = \int [D\phi_>]\, e^{-S[\phi_< + \phi_>]} \tag{7.14}$$

where $\phi_<$ = modes with $|k| < \Lambda$ and $\phi_>$ = modes with $|k| > \Lambda$.

The effective action S_{eff} has the form:

$$S_{\text{eff}} = \int d^4x \left[\frac{Z(\Lambda)}{2}(\partial\phi)^2 + \frac{m^2(\Lambda)}{2}\phi^2 + \frac{\lambda(\Lambda)}{4!}\phi^4 + \frac{c_6(\Lambda)}{6!}\phi^6 + \cdots \right] \tag{7.15}$$

In general, infinitely many terms are generated. But for a renormalizable theory, the *important* terms (those that grow or stay constant as $\Lambda \to \infty$) are just the original ones. The ϕ^6, ϕ^8, etc. terms are suppressed by powers of $1/\Lambda$.

The RG flow is the dependence of these coefficients on Λ:

$$\Lambda \frac{d\lambda(\Lambda)}{d\Lambda} = \beta(\lambda) \tag{7.16}$$

This is the same beta function we derived from dimensional regularization! The two approaches—Wilsonian and counterterm/dimensional regularization—give identical physics.

7.11 Reconciliation: Same Physics, Different Questions

Let me state the relationship crisply.

Wilsonian picture:

- Start with bare theory at cutoff Λ_0.
- Integrate out high-momentum modes.
- Get effective theory at lower cutoff $\Lambda < \Lambda_0$.
- RG flow goes UV $\to$ IR.
- Question: "Given the UV theory, what's the IR behavior?"

Particle physics picture:

- Measure physical parameters at accessible energy μ.
- Use RG to extrapolate to other scales.
- Counterterms absorb the difference between bare and physical.
- Question: "Given IR measurements, what does the theory predict at other scales?"

The beta function is the same in both. The physics is the same. The difference is the question being asked.

In condensed matter, you often know the microscopic theory (atoms, electrons) and want to understand emergent phenomena. Flow UV $\to$ IR.

In particle physics, you don't know the ultimate UV theory. You measure what you can and extrapolate. The RG lets you connect scales, in either direction.

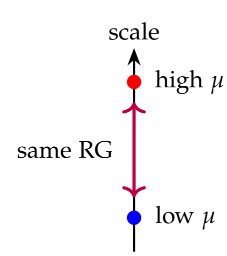


Figure 7.8: The RG relates high and low scales—we can go either direction.

7.12 *Why the Confusion Arises*

The confusion between "UV $\to$ IR" and "IR $\to$ UV" arises from conflating two things:

1. The direction of the RG flow (always UV $\to$ IR for relevant/marginal operators).

2. The direction of our inference (from what we know to what we want to know).

In condensed matter, we typically know the UV (the lattice model) and want the IR (the continuum behavior). We flow in the same direction as the physics.

In particle physics, we know the IR (measurements) and want to know about the UV (high energies). We use the RG flow equation, but we solve it "backwards" to extrapolate upward.

Neither approach changes the fundamental physics. The beta function has a sign; that sign determines whether a coupling increases or decreases with scale. QED's coupling increases with energy. QCD's decreases (asymptotic freedom). These are physical facts, independent of how we use the RG.

7.13 *Renormalizable vs. Non-Renormalizable: The Modern View*

Armed with this understanding, let's revisit non-renormalizable theories.

A non-renormalizable theory has couplings with negative mass dimension. At low energies, these couplings are suppressed by powers of E/Λ, where Λ is the cutoff. But at $E \sim \Lambda$, they become important.

In the Wilsonian language: non-renormalizable couplings are *irrelevant*. They shrink under RG flow toward the IR. Starting from generic UV physics, they wash out.

This is good news! It means low-energy physics is *insensitive* to UV details. Whatever mess is happening at the Planck scale, its effects on atomic physics are suppressed by $(E_{\text{atom}}/M_{\text{Pl}})^n$ for some $n > 0$.

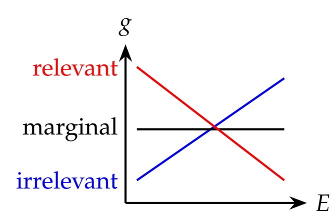


Figure 7.9: Relevant (red), marginal (black), and irrelevant (blue) couplings.

The modern view: every theory is an effective theory. Renormalizable theories are those where irrelevant operators can be ignored at low energies. Non-renormalizable theories are useful too, as long as we stay below their cutoff.

Gravity is non-renormalizable. The Newton constant G_N has dimension -2. But for $E \ll M_{\text{Pl}}$, quantum gravity corrections are tiny. We can do gravitational physics perfectly well at everyday energies; it's only at Planck-scale energies that non-renormalizability matters.

7.14 *Summary: What Renormalization Really Means*

Let me pull all this together.

Bare parameters are what we write in the Lagrangian; renormalized parameters are what we measure. They differ by quantum corrections, and the difference is divergent. Counterterms absorb this divergence, and renormalization conditions fix the counterterms by connecting them to experiment.

The miracle is that renormalizable theories need only finitely many counterterms. All the infinitely many divergent diagrams can be absorbed into a finite number of parameters. Once we fix those parameters by measurement, everything else is predicted.

The Wilsonian picture and the counterterm picture describe the same physics. The RG flow is always UV $\to$ IR—that's physics, that's causality, that's the second law of thermodynamics. Particle physics appears to go "IR $\to$ UV" because we're extrapolating from measurements, not because the flow reverses. The beta function governs how couplings change with scale, and it's the same in both pictures.

With this conceptual foundation, we're ready to tackle QED. The next several chapters will work through the three divergent diagrams—vacuum polarization, electron self-energy, vertex correction—in full detail. These calculations are the heart of the subject. The ideas we've developed will guide them.

The Wilsonian perspective, developed by Kenneth Wilson in the 1970s, revolutionized our understanding of renormalization. Before Wilson, renormalization was often viewed as a trick to sweep infinities under the rug. After Wilson, it became clear that renormalization is really about the separation of scales: high-energy physics decouples from low-energy physics, with the effects parameterized by a few running couplings. Wilson received the 1982 Nobel Prize for this work. The unification of the condensed matter and particle physics perspectives is one of the great intellectual achievements of 20th-century theoretical physics. What seemed like two different subjects—critical phenomena in magnets and ultraviolet divergences in QED—turned out to be manifestations of the same underlying mathematics.

8
QED: Setting the Stage

Quantum electrodynamics is the most precisely tested theory in all of science. Its predictions match experiment to parts per billion—better than we can measure most things. When you calculate the electron's magnetic moment in QED and compare to the measured value, the agreement is so good it's almost embarrassing. How can a theory with infinities hiding in every loop give answers this accurate?

We've developed the tools: perturbation theory, Feynman diagrams, dimensional regularization, the logic of renormalization. Now we apply them to QED—the theory of electrons, positrons, and photons—and see how everything fits together.

This chapter sets up the theory: the Lagrangian, the Feynman rules, and which diagrams diverge. The next three chapters compute the three fundamental divergent diagrams: vacuum polarization (photon self-energy), electron self-energy, and the vertex correction.

8.1 The Dirac Equation: A Lightning Review

Before writing down QED, let's recall how the electron is described.

The electron is a spin-$\frac{1}{2}$ particle, described by a four-component spinor ψ. The free Dirac equation is:

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \tag{8.1}$$

where γ^μ are the Dirac gamma matrices satisfying:

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \tag{8.2}$$

In the standard (Dirac) representation:

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{8.3}$$

where σ^i are the Pauli matrices.

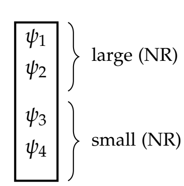


Figure 8.1: The Dirac spinor has four components. At low velocities, two are large and two are small.

The Dirac Lagrangian is:

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\partial\!\!\!/ - m)\psi \tag{8.4}$$

where $\bar{\psi} = \psi^\dagger\gamma^0$ and $\partial\!\!\!/ = \gamma^\mu\partial_\mu$.

The free electron propagator is:

$$S_F(p) = \frac{i(p\!\!\!/ + m)}{p^2 - m^2 + i\epsilon} \tag{8.5}$$

The numerator $p\!\!\!/ + m = \gamma^\mu p_\mu + m$ is a 4×4 matrix. This is more complicated than the scalar propagator.

8.2 The Photon Field

The photon is described by a four-vector field A_μ. The free Maxwell Lagrangian is:

$$\mathcal{L}_{\text{Maxwell}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{8.6}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor.

This Lagrangian has a gauge symmetry: $A_\mu \to A_\mu + \partial_\mu\chi$ for any function $\chi(x)$. This redundancy means we can't directly invert the kinetic operator to get a propagator—we need to fix the gauge.

A common choice is *Feynman gauge* (also called Lorenz gauge with a specific gauge-fixing parameter):

$$\mathcal{L}_{\text{gauge-fix}} = -\frac{1}{2\xi}(\partial_\mu A^\mu)^2 \tag{8.7}$$

With $\xi = 1$ (Feynman gauge), the photon propagator is:

$$D_F^{\mu\nu}(k) = \frac{-ig^{\mu\nu}}{k^2 + i\epsilon} \tag{8.8}$$

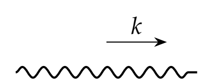


Figure 8.2: The photon propagator is proportional to $g^{\mu\nu}/k^2$. The wavy line represents a photon.

The photon is massless, so there's no m^2 in the denominator. The tensor $g^{\mu\nu}$ sums over polarizations.

8.3 The QED Lagrangian

Putting it together, the QED Lagrangian is:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\partial\!\!\!/ - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\bar{\psi}\gamma^\mu\psi A_\mu \tag{8.9}$$

The last term is the interaction: the electron current $\bar{\psi}\gamma^\mu\psi$ couples to the photon field A_μ with strength e (the electron charge).

This interaction arises from *minimal coupling*: replace $\partial_\mu \to D_\mu = \partial_\mu + ieA_\mu$ in the Dirac Lagrangian. This prescription is dictated by gauge invariance.

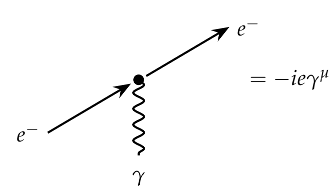


Figure 8.3: The QED vertex: two electron lines and one photon line meet with strength $-ie\gamma^\mu$.

8.4 Gauge Invariance

The Lagrangian is invariant under local $U(1)$ transformations:

$$\psi \to e^{ie\chi(x)}\psi \tag{8.10}$$

$$A_\mu \to A_\mu + \partial_\mu \chi \tag{8.11}$$

This gauge invariance has important consequences:

- The photon must be massless. A mass term $m_\gamma^2 A_\mu A^\mu$ would break gauge invariance.
- Current is conserved: $\partial_\mu(\bar{\psi}\gamma^\mu\psi) = 0$.
- Certain diagrams must cancel, enforced by *Ward identities*.

The Ward identities will be crucial for renormalization. They relate different divergent diagrams, reducing the number of independent counterterms.

8.5 Feynman Rules for QED

Let me state the complete Feynman rules.

External lines:

- Incoming electron: $u(p, s)$ (spinor)
- Outgoing electron: $\bar{u}(p, s)$
- Incoming positron: $\bar{v}(p, s)$
- Outgoing positron: $v(p, s)$
- External photon: $\epsilon_\mu(k, \lambda)$ (polarization vector)

Propagators:

- Electron: $\dfrac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}$
- Photon: $\dfrac{-ig_{\mu\nu}}{k^2 + i\epsilon}$ (Feynman gauge)

Vertices:

- QED vertex: $-ie\gamma^\mu$

Loops:

- Integrate $\displaystyle\int \frac{d^4k}{(2\pi)^4}$ for each loop.
- Fermion loops get an extra factor of (-1).

Element	Factor
Electron prop	$\frac{i(\not{p}+m)}{p^2-m^2+i\epsilon}$
Photon prop	$\frac{-ig_{\mu\nu}}{k^2+i\epsilon}$
Vertex	$-ie\gamma^\mu$
Fermion loop	-1

Table 8.1: QED Feynman rules summary.

Spinor algebra:

- Trace over spinor indices for closed fermion loops.
- Contract Lorentz indices between gamma matrices, propagators, and vertices.

8.6 *Power Counting in QED*

Let's determine which QED diagrams diverge.

For a diagram with E_e external electron lines and E_γ external photon lines, the superficial degree of divergence is:

$$D = 4 - \frac{3}{2}E_e - E_\gamma \tag{8.12}$$

Let me derive this. Count powers of momentum in the integral:

- Each loop contributes $+4$ (from d^4k).
- Each electron propagator contributes -1 (goes like $1/k$ at large k).
- Each photon propagator contributes -2 (goes like $1/k^2$).
- Each vertex contributes 0 (it's just a gamma matrix).

Using topological relations (similar to ϕ^4), we get $D = 4 - \frac{3}{2}E_e - E_\gamma$.

The divergent diagrams are:

1. **Vacuum polarization**: $E_e = 0$, $E_\gamma = 2 \Rightarrow D = 2$. The photon self-energy.
2. **Electron self-energy**: $E_e = 2$, $E_\gamma = 0 \Rightarrow D = 1$. The electron propagator correction.
3. **Vertex correction**: $E_e = 2$, $E_\gamma = 1 \Rightarrow D = 0$. The vertex is modified.

$E_e = 0$, $E_\gamma = 2$
$D = 2$: vacuum pol.

$E_e = 2$, $E_\gamma = 0$
$D = 1$: electron SE

$E_e = 2$, $E_\gamma = 1$
$D = 0$: vertex

Figure 8.4: The three divergent structures in QED.

What about $E_\gamma = 1$, $E_e = 0$ (a "photon tadpole")? This would have $D = 3$, but it vanishes by Lorentz invariance (a single photon carries spin-1; there's no Lorentz-invariant 1-point function).

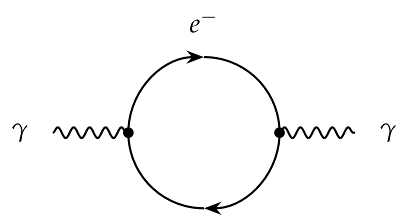
What about $E_\gamma = 3$ (three-photon vertex)? This has $D = 1$, but it's forbidden by charge conjugation symmetry (Furry's theorem).

So we have exactly three divergent structures. QED is renormalizable.

8.7 *The Three Diagrams*

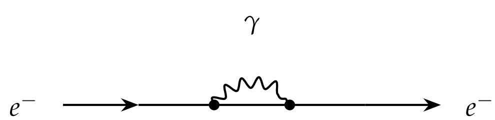
Let me show the three divergent diagrams at one-loop:

1. Vacuum Polarization (Photon Self-Energy):



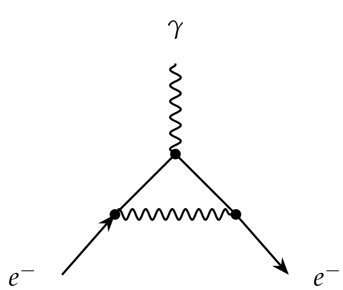
A photon fluctuates into a virtual electron-positron pair, which then annihilates back into a photon. This modifies the photon propagator and screens the charge.

2. Electron Self-Energy:



An electron emits and reabsorbs a virtual photon. This modifies the electron propagator, contributing to mass and wave-function renormalization.

3. Vertex Correction:



The electron-photon vertex is modified by a virtual photon exchange within the interaction. This affects the coupling and the magnetic moment.

8.8 *What We'll Calculate*

In the next three chapters, we'll compute each of these diagrams in detail:

Chapter 9: Vacuum Polarization. We'll find that the divergent part has the structure $k_\mu k_\nu - g_{\mu\nu}k^2$—it's transverse, as gauge invariance requires. This modifies the photon propagator and makes the charge "run."

Chapter 10: Electron Self-Energy. The electron self-energy $\Sigma(p)$ has two divergent parts: one proportional to $\not{p}$ (wave-function renormalization Z_2) and one proportional to m (mass renormalization δm).

Chapter 11: Vertex Correction. The vertex correction modifies γ^μ to $\gamma^\mu + \Lambda^\mu$. The Ward identity tells us that the divergent part of Λ^μ is proportional to γ^μ, with the same coefficient as the electron self-energy divergence. This is why $Z_1 = Z_2$.

8.9 The Structure of QED Counterterms

The QED Lagrangian with counterterms is:

$$\begin{aligned}\mathcal{L} &= \bar{\psi}(i\not{\partial} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\bar{\psi}\gamma^\mu\psi A_\mu \\ &\quad + \delta_{Z_2}\bar{\psi}i\not{\partial}\psi - \delta_m\bar{\psi}\psi - \frac{\delta_{Z_3}}{4}F_{\mu\nu}F^{\mu\nu} - \delta_{Z_1}e\bar{\psi}\gamma^\mu\psi A_\mu\end{aligned} \tag{8.13}$$

The counterterms are:

- δ_{Z_2}: electron wave-function renormalization
- δ_m: electron mass renormalization
- δ_{Z_3}: photon wave-function renormalization
- δ_{Z_1}: vertex renormalization

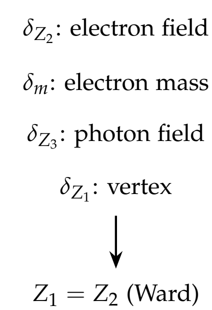


Figure 8.5: Four counterterms, but the Ward identity relates Z_1 and Z_2.

But wait—that's four counterterms for three divergent structures. The Ward identity (gauge invariance) tells us $Z_1 = Z_2$, so there are really only three independent counterterms.

We can define the renormalized coupling:

$$e_{\text{ren}} = \frac{Z_1}{Z_2}\sqrt{Z_3}\,e_0 = \sqrt{Z_3}\,e_0 \tag{8.14}$$

The factors of Z_1/Z_2 cancel! Only the photon wave-function renormalization affects the physical coupling.

The electron's "dress" (its cloud of virtual photons) doesn't affect the charge—only the photon's self-interaction does. Gauge invariance protects the charge from some quantum corrections.

8.10 Spinor Algebra: Some Useful Identities

Before we start computing, let me collect some useful identities for gamma matrix algebra. You'll need these constantly.

Basic anticommutator:

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \tag{8.15}$$

Contractions:

$$\gamma^\mu \gamma_\mu = d \cdot I \quad \text{(in } d \text{ dimensions)} \tag{8.16}$$

$$\gamma^\mu \gamma^\nu \gamma_\mu = (2-d)\gamma^\nu \tag{8.17}$$

$$\gamma^\mu \gamma^\nu \gamma^\rho \gamma_\mu = 4g^{\nu\rho} - (4-d)\gamma^\nu \gamma^\rho \tag{8.18}$$

In $d = 4 - \epsilon$:

$$\gamma^\mu \gamma_\mu = 4 - \epsilon \tag{8.19}$$

$$\gamma^\mu \gamma^\nu \gamma_\mu = -(2-\epsilon)\gamma^\nu \tag{8.20}$$

$$\gamma^\mu \gamma^\nu \gamma^\rho \gamma_\mu = 4g^{\nu\rho} - \epsilon\gamma^\nu \gamma^\rho \tag{8.21}$$

Traces:

$$\text{Tr}(I) = 4 \quad \text{(in 4 dimensions)} \tag{8.22}$$

$$\text{Tr}(\gamma^\mu \gamma^\nu) = 4g^{\mu\nu} \tag{8.23}$$

$$\text{Tr}(\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma) = 4(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}) \tag{8.24}$$

Traces with odd numbers of gamma matrices vanish:

$$\text{Tr}(\gamma^\mu) = \text{Tr}(\gamma^\mu \gamma^\nu \gamma^\rho) = 0 \tag{8.25}$$

These identities follow from the anticommutation relations. You'll use them to simplify numerators of loop integrals.

8.11 *A Taste of What's Coming*

Let me sketch what the vacuum polarization calculation looks like, so you know what to expect.

The one-loop photon self-energy is:

$$i\Pi^{\mu\nu}(k) = (-1)(-ie)^2 \int \frac{d^d p}{(2\pi)^d} \text{Tr}\left[\gamma^\mu \frac{i(\not{p} + m)}{p^2 - m^2} \gamma^\nu \frac{i(\not{p} - \not{k} + m)}{(p-k)^2 - m^2}\right] \tag{8.26}$$

The (-1) is from the fermion loop. The trace is over spinor indices.

We'll need to:

1. Evaluate the trace using gamma matrix identities.
2. Combine denominators using Feynman parameters.
3. Shift the loop momentum.
4. Use the master integral formula from Chapter 6.

5. Expand in ϵ.

The result is:

$$\Pi^{\mu\nu}(k) = (k^2 g^{\mu\nu} - k^\mu k^\nu)\Pi(k^2) \tag{8.27}$$

The tensor structure $(k^2 g^{\mu\nu} - k^\mu k^\nu)$ is transverse: $k_\mu \Pi^{\mu\nu} = 0$. This is required by gauge invariance (the Ward identity).

The scalar function $\Pi(k^2)$ contains a $1/\epsilon$ divergence that we'll renormalize by the photon field counterterm δ_{Z_3}.

8.12 Units and Numbers

Let's establish the numerical values we'll be working with.

The fine-structure constant is:

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137.036} \tag{8.28}$$

The electron mass is:

$$m_e = 0.511 \text{ MeV} = 9.11 \times 10^{-31} \text{ kg} \tag{8.29}$$

The expansion parameter for QED perturbation theory is:

$$\frac{\alpha}{\pi} \approx \frac{1}{430} \tag{8.30}$$

Quantity	Value
α	$1/137.036$
m_e	0.511 MeV
$\lambda_e = 1/m_e$	3.86×10^{-13} m

Table 8.2: Key QED parameters.

This is small! Each additional loop adds roughly a factor of α/π. Even at ten loops, the correction is $(\alpha/\pi)^{10} \approx 10^{-26}$. QED perturbation theory converges very rapidly.

This is why QED gives such precise predictions. The one-loop correction to the electron's magnetic moment is $\alpha/(2\pi) \approx 0.00116$. The full calculation, carried to five loops, agrees with experiment to 10^{-10}.

8.13 Historical Note: The Lamb Shift

The development of renormalized QED was driven by experiment. In 1947, Willis Lamb and Robert Retherford measured a tiny splitting in the hydrogen spectrum—the $2S_{1/2}$ and $2P_{1/2}$ levels, which should be degenerate according to the Dirac equation, were separated by about 1000 MHz.

This "Lamb shift" couldn't be explained without quantum corrections to the electron's self-energy and the vertex. The theoretical calculation, requiring careful handling of the infinities, gave a result that agreed with the measurement.

The success of renormalized QED, demonstrated by the Lamb shift calculation, established quantum field theory as the correct framework for particle physics. Tomonaga, Schwinger, and Feynman shared the 1965 Nobel Prize for their contributions.

8.14 Summary: The Setup

We're now ready to compute. Here's what we have:

1. The QED Lagrangian, with its gauge symmetry.
2. The Feynman rules: propagators, vertex, loop integration.
3. Power counting: three divergent structures.
4. Gamma matrix algebra for handling spinors.
5. The small parameter $\alpha \approx 1/137$.

In the next chapter, we'll compute the vacuum polarization in full detail. Every step will be shown, and the physical meaning will be explained as we go.

QED was the first successful quantum field theory, and it remains the cleanest example of how renormalization works. The gauge symmetry—local $U(1)$ invariance—is both a blessing and a complication. It's a blessing because it constrains the form of interactions and links different counterterms through Ward identities. It's a complication because we must fix a gauge to define propagators, and we must verify that physical results are gauge-independent. The beauty of QED is that everything fits together: gauge invariance, renormalizability, unitarity, and agreement with experiment to extraordinary precision. Understanding QED deeply is the foundation for understanding the Standard Model, which is "just" a larger gauge theory with the same underlying structure.

9
Vacuum Polarization

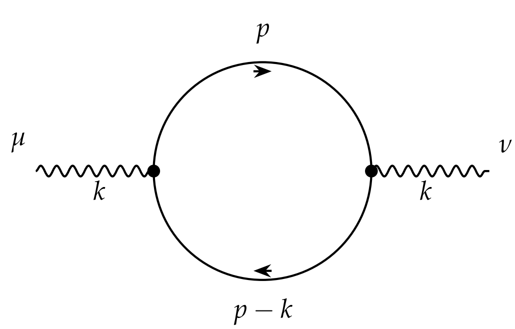
What is the vacuum? In classical physics, empty space is truly empty. But in quantum field theory, the vacuum seethes with activity. Virtual particle-antiparticle pairs flicker in and out of existence, borrowing energy for fleeting moments. A photon traveling through this quantum vacuum doesn't travel through nothing—it travels through a medium.

Vacuum polarization is the effect of this medium on light. A photon can briefly convert into a virtual electron-positron pair, which then annihilates back into a photon. This process modifies how photons propagate and screens the electric charge. The effect is small—suppressed by α—but it's measurable, and calculating it will teach us how gauge invariance constrains quantum corrections.

This chapter works through the calculation in full detail. What emerges has a very specific structure, one that gauge invariance demands.

9.1 The Diagram

The one-loop vacuum polarization diagram is:



The external photon carries momentum k and has Lorentz indices μ (incoming) and ν (outgoing). The loop momentum p flows around the fermion loop.

9.2 *Writing Down the Integral*

Using the Feynman rules from Chapter 8:

$$i\Pi^{\mu\nu}(k) = (-1)(-ie)^2 \int \frac{d^d p}{(2\pi)^d} \text{Tr}\left[\gamma^\mu \frac{i(\not{p}+m)}{p^2-m^2+i\epsilon}\gamma^\nu \frac{i(\not{p}-\not{k}+m)}{(p-k)^2-m^2+i\epsilon}\right] \tag{9.1}$$

Let me explain each factor:

- (-1): Fermion loop sign.
- $(-ie)^2 = -e^2$: Two vertices.
- The trace: Fermion loop is a closed spinor line; trace over spinor indices.
- First propagator: Electron with momentum p.
- Second propagator: Electron with momentum $p-k$.

Simplifying:

$$i\Pi^{\mu\nu}(k) = e^2 \int \frac{d^d p}{(2\pi)^d} \frac{\text{Tr}[\gamma^\mu(\not{p}+m)\gamma^\nu(\not{p}-\not{k}+m)]}{(p^2-m^2+i\epsilon)((p-k)^2-m^2+i\epsilon)} \tag{9.2}$$

9.3 *Evaluating the Trace*

The numerator contains:

$$N^{\mu\nu} = \text{Tr}[\gamma^\mu(\not{p}+m)\gamma^\nu(\not{p}-\not{k}+m)] \tag{9.3}$$

Expand:

$$N^{\mu\nu} = \text{Tr}[\gamma^\mu \not{p}\gamma^\nu(\not{p}-\not{k})] + m\text{Tr}[\gamma^\mu \not{p}\gamma^\nu] + m\text{Tr}[\gamma^\mu\gamma^\nu(\not{p}-\not{k})] + m^2\text{Tr}[\gamma^\mu\gamma^\nu] \tag{9.4}$$

The middle two terms have three gamma matrices in the trace, so they vanish. We're left with:

$$N^{\mu\nu} = \text{Tr}[\gamma^\mu \not{p}\gamma^\nu(\not{p}-\not{k})] + m^2\text{Tr}[\gamma^\mu\gamma^\nu] \tag{9.5}$$

Trace rules:

$\text{Tr}(\text{odd } \gamma\text{'s}) = 0$

$\text{Tr}(\gamma^\mu\gamma^\nu) = 4g^{\mu\nu}$

Figure 9.1: Key trace identities for gamma matrices.

Using $\text{Tr}[\gamma^\mu\gamma^\nu] = 4g^{\mu\nu}$ (in 4 dimensions):

$$m^2\text{Tr}[\gamma^\mu\gamma^\nu] = 4m^2 g^{\mu\nu} \tag{9.6}$$

For the four-gamma trace, use the identity:

$$\text{Tr}[\gamma^\alpha\gamma^\beta\gamma^\rho\gamma^\sigma] = 4(g^{\alpha\beta}g^{\rho\sigma} - g^{\alpha\rho}g^{\beta\sigma} + g^{\alpha\sigma}g^{\beta\rho}) \tag{9.7}$$

With $\gamma^\alpha = \gamma^\mu$, $\gamma^\beta = \gamma^\lambda$ (from $\not{p} = p_\lambda \gamma^\lambda$), etc.:

$$\begin{aligned} \mathrm{Tr}[\gamma^\mu \not{p} \gamma^\nu (\not{p} - \not{k})] &= p_\lambda (p-k)_\sigma \mathrm{Tr}[\gamma^\mu \gamma^\lambda \gamma^\nu \gamma^\sigma] && (9.8) \\ &= 4 p_\lambda (p-k)_\sigma (g^{\mu\lambda} g^{\nu\sigma} - g^{\mu\nu} g^{\lambda\sigma} + g^{\mu\sigma} g^{\nu\lambda}) && (9.9) \\ &= 4[p^\mu (p-k)^\nu - g^{\mu\nu} p \cdot (p-k) + p^\nu (p-k)^\mu] && (9.10) \end{aligned}$$

So the full numerator is:

$$N^{\mu\nu} = 4[p^\mu (p-k)^\nu + p^\nu (p-k)^\mu - g^{\mu\nu}(p \cdot (p-k) - m^2)] \quad (9.11)$$

9.4 *Combining Denominators*

The denominator of (9.2) has two propagators. Use Feynman parameters:

$$\frac{1}{AB} = \int_0^1 dx \, \frac{1}{[xA + (1-x)B]^2} \quad (9.12)$$

With $A = p^2 - m^2$ and $B = (p-k)^2 - m^2$:

$$\begin{aligned} xA + (1-x)B &= x(p^2 - m^2) + (1-x)((p-k)^2 - m^2) && (9.13) \\ &= p^2 - 2(1-x) p \cdot k + (1-x)k^2 - m^2 && (9.14) \end{aligned}$$

Complete the square by defining $\ell = p - (1-x)k$:

$$\begin{aligned} xA + (1-x)B &= \ell^2 + (1-x)k^2 - (1-x)^2 k^2 - m^2 && (9.15) \\ &= \ell^2 - \Delta && (9.16) \end{aligned}$$

where

$$\Delta = m^2 - x(1-x)k^2 \quad (9.17)$$

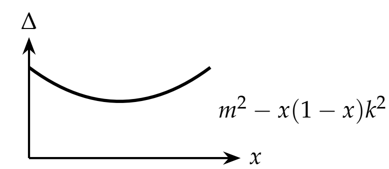


Figure 9.2: The function $\Delta(x) = m^2 - x(1-x)k^2$ is positive for $k^2 < 4m^2$.

9.5 *Shifting the Loop Momentum*

We now shift the loop momentum to eliminate the cross-term in the denominator. Define $\ell = p - (1-x)k$, so that:

- $p = \ell + (1-x)k$
- $p - k = \ell - xk$

The numerator $N^{\mu\nu} = 4[p^\mu (p-k)^\nu + p^\nu (p-k)^\mu - g^{\mu\nu}(p \cdot (p-k) - m^2)]$ transforms as follows.

The First Two Terms

Consider $p^\mu (p-k)^\nu$:

$$\begin{aligned} p^\mu (p-k)^\nu &= (\ell^\mu + (1-x)k^\mu)(\ell^\nu - xk^\nu) && (9.18) \\ &= \ell^\mu \ell^\nu - x\ell^\mu k^\nu + (1-x)k^\mu \ell^\nu - x(1-x)k^\mu k^\nu && (9.19) \end{aligned}$$

Terms linear in ℓ vanish after integration because the integrand is odd while the domain is symmetric. Keeping only the surviving terms:

$$p^\mu (p-k)^\nu \to \ell^\mu \ell^\nu - x(1-x)k^\mu k^\nu \tag{9.20}$$

By symmetry $\mu \leftrightarrow \nu$:

$$p^\nu (p-k)^\mu \to \ell^\nu \ell^\mu - x(1-x)k^\nu k^\mu \tag{9.21}$$

So the sum becomes:

$$p^\mu (p-k)^\nu + p^\nu (p-k)^\mu \to 2\ell^\mu \ell^\nu - 2x(1-x)k^\mu k^\nu \tag{9.22}$$

The Third Term

For the dot product:

$$\begin{aligned} p \cdot (p-k) &= (\ell + (1-x)k) \cdot (\ell - xk) && (9.23) \\ &= \ell^2 - x\ell \cdot k + (1-x)\ell \cdot k - x(1-x)k^2 && (9.24) \\ &= \ell^2 + (1-2x)\ell \cdot k - x(1-x)k^2 && (9.25) \end{aligned}$$

The linear term $\ell \cdot k$ vanishes upon integration, leaving:

$$p \cdot (p-k) \to \ell^2 - x(1-x)k^2 \tag{9.26}$$

Assembling the Numerator

Putting everything together:

$$\begin{aligned} N^{\mu\nu} &= 4\left[2\ell^\mu \ell^\nu - 2x(1-x)k^\mu k^\nu - g^{\mu\nu}\left(\ell^2 - x(1-x)k^2 - m^2\right)\right] && (9.27) \\ &= 8\ell^\mu \ell^\nu - 4g^{\mu\nu}\ell^2 - 8x(1-x)k^\mu k^\nu + 4g^{\mu\nu}\left(x(1-x)k^2 + m^2\right) && (9.28) \end{aligned}$$

Since $\Delta = m^2 - x(1-x)k^2$, we have $x(1-x)k^2 + m^2 = 2m^2 - \Delta$. The numerator becomes:

$$N^{\mu\nu} = 8\ell^\mu \ell^\nu - 4g^{\mu\nu}\ell^2 - 8x(1-x)k^\mu k^\nu + 4g^{\mu\nu}(2m^2 - \Delta) \tag{9.29}$$

9.6 *The Integral Structure*

The vacuum polarization is now:

$$i\Pi^{\mu\nu}(k) = e^2 \int_0^1 dx \int \frac{d^d\ell}{(2\pi)^d} \frac{N^{\mu\nu}}{(\ell^2 - \Delta)^2} \tag{9.30}$$

From (9.29), the numerator splits into three types of terms:

$$N^{\mu\nu} = \underbrace{8\ell^\mu \ell^\nu}_{\text{tensor}} - \underbrace{4g^{\mu\nu}\ell^2}_{\text{scalar}} + \underbrace{[-8x(1-x)k^\mu k^\nu + 4g^{\mu\nu}(2m^2 - \Delta)]}_{\text{constant in } \ell} \tag{9.31}$$

Each requires a different integral.

The Master Integrals

We need three integrals over ℓ:

$$J_0 = \int \frac{d^d\ell}{(2\pi)^d} \frac{1}{(\ell^2 - \Delta)^2} \tag{9.32}$$

$$J_1 = \int \frac{d^d\ell}{(2\pi)^d} \frac{\ell^2}{(\ell^2 - \Delta)^2} \tag{9.33}$$

$$J_2^{\mu\nu} = \int \frac{d^d\ell}{(2\pi)^d} \frac{\ell^\mu \ell^\nu}{(\ell^2 - \Delta)^2} \tag{9.34}$$

The tensor integral $J_2^{\mu\nu}$ must be proportional to $g^{\mu\nu}$ by Lorentz symmetry (there's no other tensor available). Contracting both sides with $g_{\mu\nu}$:

$$g_{\mu\nu} J_2^{\mu\nu} = J_1 = g_{\mu\nu} \cdot \frac{g^{\mu\nu}}{d} J_1 = \frac{d}{d} J_1 \tag{9.35}$$

$$\int \frac{\ell^\mu \ell^\nu}{(\cdots)^n} = \frac{g^{\mu\nu}}{d} \int \frac{\ell^2}{(\cdots)^n}$$

Figure 9.3: Symmetric integration gives $g^{\mu\nu}/d$.

which is consistent. Therefore:

$$J_2^{\mu\nu} = \frac{g^{\mu\nu}}{d} J_1 \tag{9.36}$$

Using the master formula from Chapter 6:

$$\int \frac{d^d\ell}{(2\pi)^d} \frac{(\ell^2)^a}{(\ell^2 - \Delta)^n} = \frac{i(-1)^{n-a}}{(4\pi)^{d/2}} \frac{\Gamma(a + d/2)\Gamma(n - a - d/2)}{\Gamma(d/2)\Gamma(n)} \Delta^{a+d/2-n} \tag{9.37}$$

For J_0 (where $a = 0$, $n = 2$):

$$J_0 = \frac{i}{(4\pi)^{d/2}} \frac{\Gamma(d/2)\Gamma(2 - d/2)}{\Gamma(d/2)\Gamma(2)} \Delta^{d/2-2} \tag{9.38}$$

$$= \frac{i}{(4\pi)^{d/2}} \Gamma(2 - d/2) \Delta^{d/2-2} \tag{9.39}$$

For J_1 (where $a = 1$, $n = 2$):

$$J_1 = \frac{i(-1)}{(4\pi)^{d/2}} \frac{\Gamma(1 + d/2)\Gamma(1 - d/2)}{\Gamma(d/2)\Gamma(2)} \Delta^{1+d/2-2} \tag{9.40}$$

$$= \frac{-i}{(4\pi)^{d/2}} \frac{\Gamma(1 + d/2)\Gamma(1 - d/2)}{\Gamma(d/2)} \Delta^{d/2-1} \tag{9.41}$$

Using $\Gamma(1 + d/2) = (d/2)\Gamma(d/2)$:

$$J_1 = \frac{-i}{(4\pi)^{d/2}} \frac{d}{2} \Gamma(1 - d/2) \Delta^{d/2-1} \tag{9.42}$$

And therefore:

$$J_2^{\mu\nu} = \frac{g^{\mu\nu}}{d} J_1 = \frac{-ig^{\mu\nu}}{2(4\pi)^{d/2}} \Gamma(1 - d/2) \Delta^{d/2-1} \tag{9.43}$$

9.7 *Putting It Together*

We now assemble the vacuum polarization using these integrals. The integral over ℓ gives:

$$i\Pi^{\mu\nu}(k) = e^2 \int_0^1 dx \left[8J_2^{\mu\nu} - 4g^{\mu\nu}J_1 + \Big(-8x(1-x)k^\mu k^\nu + 4g^{\mu\nu}(2m^2-\Delta)\Big) J_0\right] \tag{9.44}$$

Substituting the expressions for J_0, J_1, $J_2^{\mu\nu}$:

$$\begin{aligned} i\Pi^{\mu\nu}(k) = e^2 \int_0^1 dx \Bigg[& 8 \cdot \frac{-ig^{\mu\nu}}{2(4\pi)^{d/2}}\Gamma(1-d/2)\Delta^{d/2-1} \\ & - 4g^{\mu\nu} \cdot \frac{-i}{(4\pi)^{d/2}}\frac{d}{2}\Gamma(1-d/2)\Delta^{d/2-1} \\ & + \Big(-8x(1-x)k^\mu k^\nu + 4g^{\mu\nu}(2m^2-\Delta)\Big) \cdot \frac{i}{(4\pi)^{d/2}}\Gamma(2-d/2)\Delta^{d/2-2}\Bigg] \end{aligned} \tag{9.45}$$

Let's simplify term by term.

First term:

$$8 \cdot \frac{-ig^{\mu\nu}}{2(4\pi)^{d/2}}\Gamma(1-d/2)\Delta^{d/2-1} = \frac{-4ig^{\mu\nu}}{(4\pi)^{d/2}}\Gamma(1-d/2)\Delta^{d/2-1} \tag{9.46}$$

Second term:

$$-4g^{\mu\nu} \cdot \frac{-i}{(4\pi)^{d/2}}\frac{d}{2}\Gamma(1-d/2)\Delta^{d/2-1} = \frac{2idg^{\mu\nu}}{(4\pi)^{d/2}}\Gamma(1-d/2)\Delta^{d/2-1} \tag{9.47}$$

Combining the first two terms:

$$\text{(first + second)} = \frac{ig^{\mu\nu}}{(4\pi)^{d/2}}\Gamma(1-d/2)\Delta^{d/2-1}\,(-4+2d) \tag{9.48}$$

$$= \frac{ig^{\mu\nu}}{(4\pi)^{d/2}}\Gamma(1-d/2)\Delta^{d/2-1} \cdot 2(d-2) \tag{9.49}$$

Now use the gamma function identity $\Gamma(z+1) = z\Gamma(z)$, which gives:

$$\Gamma(2-d/2) = (1-d/2)\Gamma(1-d/2) \tag{9.50}$$

and therefore:

$$\Gamma(1-d/2) = \frac{\Gamma(2-d/2)}{1-d/2} = \frac{-2\Gamma(2-d/2)}{d-2} \tag{9.51}$$

Substituting:

$$\text{(first + second)} = \frac{ig^{\mu\nu}}{(4\pi)^{d/2}} \cdot \frac{-2\Gamma(2-d/2)}{d-2} \cdot \Delta^{d/2-1} \cdot 2(d-2) \tag{9.52}$$

$$= \frac{-4ig^{\mu\nu}}{(4\pi)^{d/2}}\Gamma(2-d/2)\Delta^{d/2-1} \tag{9.53}$$

Third term:

$$\frac{i}{(4\pi)^{d/2}}\Gamma(2-d/2)\Delta^{d/2-2}\left(-8x(1-x)k^\mu k^\nu + 4g^{\mu\nu}(2m^2-\Delta)\right) \quad (9.54)$$

Let's expand the factor $(2m^2 - \Delta)$. Since $\Delta = m^2 - x(1-x)k^2$:

$$2m^2 - \Delta = 2m^2 - m^2 + x(1-x)k^2 = m^2 + x(1-x)k^2 \quad (9.55)$$

The third term becomes:

$$\frac{i}{(4\pi)^{d/2}}\Gamma(2-d/2)\Delta^{d/2-2}\left(-8x(1-x)k^\mu k^\nu + 4g^{\mu\nu}(m^2 + x(1-x)k^2)\right) \quad (9.56)$$

Combining all three terms:

$$i\Pi^{\mu\nu}(k) = e^2\int_0^1 dx \frac{i\Gamma(2-d/2)}{(4\pi)^{d/2}}\Bigg[-4g^{\mu\nu}\Delta^{d/2-1} + \Delta^{d/2-2}\left(-8x(1-x)k^\mu k^\nu + 4g^{\mu\nu}(m^2+x(1-x)k^2)\right)\Bigg] \quad (9.57)$$

Factor out $\Delta^{d/2-2}$:

$$i\Pi^{\mu\nu}(k) = \frac{ie^2\Gamma(2-d/2)}{(4\pi)^{d/2}}\int_0^1 dx\,\Delta^{d/2-2}\Bigg[-4g^{\mu\nu}\Delta - 8x(1-x)k^\mu k^\nu + 4g^{\mu\nu}(m^2+x(1-x)k^2)\Bigg] \quad (9.58)$$

The $g^{\mu\nu}$ terms combine:

$$-4g^{\mu\nu}\Delta + 4g^{\mu\nu}(m^2+x(1-x)k^2) = 4g^{\mu\nu}\left[-\Delta + m^2 + x(1-x)k^2\right] \quad (9.59)$$

$$= 4g^{\mu\nu}\left[-(m^2 - x(1-x)k^2) + m^2 + x(1-x)k^2\right] \quad (9.60)$$

$$= 4g^{\mu\nu}\cdot 2x(1-x)k^2 \quad (9.61)$$

$$= 8x(1-x)k^2 g^{\mu\nu} \quad (9.62)$$

So we have:

$$\begin{aligned} i\Pi^{\mu\nu}(k) &= \frac{ie^2\Gamma(2-d/2)}{(4\pi)^{d/2}}\int_0^1 dx\,\Delta^{d/2-2}\left[8x(1-x)k^2g^{\mu\nu} - 8x(1-x)k^\mu k^\nu\right] \\ &= \frac{8ie^2\Gamma(2-d/2)}{(4\pi)^{d/2}}\int_0^1 dx\,x(1-x)\Delta^{d/2-2}\left(k^2g^{\mu\nu} - k^\mu k^\nu\right) \end{aligned} \quad (9.63)$$

The tensor structure has emerged: $\Pi^{\mu\nu}(k) = (k^2g^{\mu\nu} - k^\mu k^\nu)\Pi(k^2)$. Dividing both sides by i:

$$\Pi(k^2) = \frac{8e^2\Gamma(2-d/2)}{(4\pi)^{d/2}}\int_0^1 dx\,x(1-x)\Delta^{d/2-2} \quad (9.64)$$

9.8 The Transverse Structure

The tensor structure $(k^2 g^{\mu\nu} - k^\mu k^\nu)$ is *transverse*: it satisfies

$$k_\mu(k^2 g^{\mu\nu} - k^\mu k^\nu) = k^2 k^\nu - k^2 k^\nu = 0 \tag{9.65}$$

This is required by gauge invariance. The Ward identity demands that the photon self-energy be transverse. The fact that this structure emerged automatically from our calculation is a powerful check: dimensional regularization preserves gauge invariance.

9.9 Extracting the Scalar Function $\Pi(k^2)$

From (9.64), we have the exact d-dimensional result:

$$\Pi(k^2) = \frac{8e^2\Gamma(2-d/2)}{(4\pi)^{d/2}} \int_0^1 dx\, x(1-x)(m^2 - x(1-x)k^2)^{d/2-2} \tag{9.66}$$

Now we expand in $d = 4 - \epsilon$. We need:

- $(4\pi)^{d/2} = (4\pi)^{2-\epsilon/2} = 16\pi^2\left(1 - \frac{\epsilon}{2}\ln(4\pi) + O(\epsilon^2)\right)$
- $\Gamma(2-d/2) = \Gamma(\epsilon/2) = \frac{2}{\epsilon} - \gamma_E + O(\epsilon)$
- $\Delta^{d/2-2} = \Delta^{-\epsilon/2} = 1 - \frac{\epsilon}{2}\ln\Delta + O(\epsilon^2)$

In $\overline{\text{MS}}$, we introduce a scale μ to make Δ dimensionless. The standard $\overline{\text{MS}}$ pole is defined as:

$$\frac{1}{\bar{\epsilon}} \equiv \frac{1}{\epsilon} - \frac{\gamma_E}{2} + \frac{\ln(4\pi)}{2} \tag{9.67}$$

Expanding systematically, the leading $1/\epsilon$ contribution comes from:

$$\Pi(k^2) = \frac{8e^2}{16\pi^2}\cdot\frac{2}{\epsilon}\cdot\frac{1}{6} + \text{finite} = \frac{e^2}{6\pi^2\epsilon} + \text{finite} \tag{9.68}$$

where we used $\int_0^1 dx\, x(1-x) = 1/6$.

With $\alpha = e^2/(4\pi)$, this becomes $\frac{2\alpha}{3\pi\epsilon}$.

The full $\overline{\text{MS}}$ result, including finite terms, is:

$$\begin{aligned}\Pi(k^2) &= \frac{8e^2}{16\pi^2}\left(\frac{2}{\epsilon} - \gamma_E + \ln(4\pi)\right)\int_0^1 dx\, x(1-x)\left(1 - \frac{\epsilon}{2}\ln\frac{\Delta}{\mu^2}\right) \\ &= \frac{e^2}{2\pi^2}\left(\frac{2}{\epsilon} - \gamma_E + \ln(4\pi)\right)\cdot\frac{1}{6} - \frac{e^2}{2\pi^2}\int_0^1 dx\, x(1-x)\ln\frac{\Delta}{\mu^2}\end{aligned} \tag{9.69}$$

Converting to $\alpha = e^2/(4\pi)$:

$$\Pi(k^2) = \frac{2\alpha}{3\pi}\left(\frac{1}{\bar{\epsilon}}\right) - \frac{2\alpha}{\pi}\int_0^1 dx\, x(1-x)\ln\frac{m^2 - x(1-x)k^2}{\mu^2} \tag{9.70}$$

In a cleaner form:

$$\Pi(k^2) = \frac{\alpha}{\pi}\left[\frac{2}{3\bar{\epsilon}} - 2\int_0^1 dx\, x(1-x)\ln\frac{m^2 - x(1-x)k^2}{\mu^2}\right] \tag{9.71}$$

9.10 The Final Result

Collecting our results, the vacuum polarization is:

$$\Pi(k^2) = \frac{\alpha}{\pi}\left[\frac{2}{3\bar{\epsilon}} - 2\int_0^1 dx\, x(1-x)\ln\frac{m^2 - x(1-x)k^2}{\mu^2}\right] \tag{9.72}$$

where $\alpha = e^2/(4\pi)$.

The divergent part is:

$$\Pi^{\text{div}}(k^2) = \frac{2\alpha}{3\pi\bar{\epsilon}} \tag{9.73}$$

The finite part depends on k^2 through the integral. For $k^2 \ll m^2$, we expand the logarithm:

$$\ln\frac{m^2 - x(1-x)k^2}{\mu^2} = \ln\frac{m^2}{\mu^2} - \frac{x(1-x)k^2}{m^2} + O(k^4) \tag{9.74}$$

Using $\int_0^1 dx\, x(1-x) = 1/6$ and $\int_0^1 dx\, x^2(1-x)^2 = 1/30$:

$$\Pi(k^2) \approx \frac{\alpha}{\pi}\left[\frac{2}{3\bar{\epsilon}} - \frac{1}{3}\ln\frac{m^2}{\mu^2} + \frac{k^2}{15m^2} + O(k^4)\right] \tag{9.75}$$

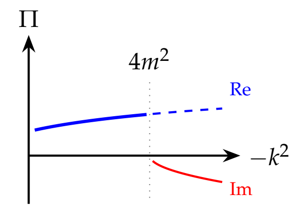


Figure 9.4: $\Pi(k^2)$ vs $-k^2$. Below threshold ($k^2 < 4m^2$), Π is real (solid blue). Above threshold, it develops an imaginary part (red)—real pair production becomes possible.

9.11 Physical Interpretation

What does the vacuum polarization mean physically?

Screening of charge. The virtual e^+e^- pairs act like dipoles in a dielectric medium. They partially screen the bare charge, making the observed charge smaller at large distances.

Running coupling. The finite part of $\Pi(k^2)$ depends on $k^2 = -Q^2$ (the momentum transfer). This means the effective coupling depends on the energy scale of the probe. At higher energies (shorter distances), the screening is less effective, and the effective charge is larger.

The modified propagator. The dressed photon propagator becomes:

$$D^{\mu\nu}(k) = \frac{-ig^{\mu\nu}}{k^2(1-\Pi(k^2))} = \frac{-ig^{\mu\nu}}{k^2}\left(1 + \Pi(k^2) + \Pi^2(k^2) + \dots\right) \tag{9.76}$$

The effective coupling at scale k^2 is:

$$e^2_{\text{eff}}(k^2) = \frac{e^2}{1-\Pi(k^2)} \tag{9.77}$$

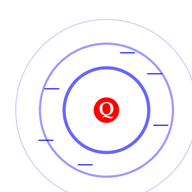


Figure 9.5: Virtual pairs screen the bare charge, like polarization in a dielectric.

9.12 *Renormalization*

The divergent part $\propto 1/\bar{\epsilon}$ must be canceled by a counterterm. The photon field renormalization Z_3 relates bare and renormalized fields: $A_\mu^{\text{bare}} = Z_3^{1/2} A_\mu^{\text{ren}}$. The counterterm is:

$$\delta_{Z_3} = Z_3 - 1 = -\frac{2\alpha}{3\pi\bar{\epsilon}} + \text{finite (scheme-dependent)} \tag{9.78}$$

The renormalized vacuum polarization is:

$$\Pi_{\text{ren}}(k^2) = \Pi(k^2) + \delta_{Z_3} = \text{finite} \tag{9.79}$$

In the $\overline{\text{MS}}$ scheme, δ_{Z_3} is chosen to cancel exactly the $1/\bar{\epsilon}$ pole:

$$\delta_{Z_3}^{\overline{\text{MS}}} = -\frac{2\alpha}{3\pi\bar{\epsilon}} \tag{9.80}$$

The renormalized result is:

$$\Pi_{\text{ren}}^{\overline{\text{MS}}}(k^2) = -\frac{2\alpha}{\pi}\int_0^1 dx\, x(1-x)\ln\frac{m^2 - x(1-x)k^2}{\mu^2} \tag{9.81}$$

In an on-shell scheme, we might instead require $\Pi_{\text{ren}}(0) = 0$, which gives a different finite part.

9.13 *Putting in Numbers*

Let's see how the vacuum polarization affects the running of α.

The key physical quantity is the change in the effective coupling between two scales. From the dressed propagator, the effective coupling at momentum transfer Q^2 is:

$$\alpha_{\text{eff}}(Q^2) = \frac{\alpha}{1 - \Delta\Pi(Q^2)} \tag{9.82}$$

where $\Delta\Pi(Q^2) = \Pi_{\text{ren}}(Q^2) - \Pi_{\text{ren}}(0)$ is the change in the vacuum polarization from zero momentum.

For spacelike momentum transfer $Q^2 \gg m_e^2$, the vacuum polarization integral can be evaluated:

$$\Delta\Pi(Q^2) \approx \frac{\alpha}{3\pi}\ln\frac{Q^2}{m_e^2} \tag{9.83}$$

At $Q = 100$ GeV:

$$\ln\frac{Q^2}{m_e^2} = \ln\frac{(10^{11}\text{ eV})^2}{(5.1\times 10^5\text{ eV})^2} \approx 24 \tag{9.84}$$

So:

$$\Delta\Pi \approx \frac{1}{137\times 3\pi}\times 24 \approx 0.019 \tag{9.85}$$

The effective coupling becomes:

$$\alpha_{\text{eff}}(Q^2) \approx \frac{1/137}{1 - 0.019} = \frac{1/137}{0.981} \approx \frac{1}{134} \tag{9.86}$$

At even higher energies (the Z boson mass, $M_Z \approx 91$ GeV), the effective coupling reaches $\alpha \approx 1/128$. This running has been measured experimentally at LEP and agrees well with the QED prediction.

9.14 Summary

We've computed the one-loop vacuum polarization in QED:

1. The tensor structure is $(k^2 g^{\mu\nu} - k^\mu k^\nu)$—transverse, as gauge invariance requires.
2. The scalar function has a divergent part $\propto 1/\bar{\epsilon}$ and a finite part depending on k^2.
3. The divergence is canceled by the photon field counterterm δ_{Z_3}.
4. The finite part encodes the running of the electric charge with energy scale.
5. At high energies ($Q^2 \gg m_e^2$), the effective coupling increases logarithmically.

This is the first of our three QED calculations. Next, we'll compute the electron self-energy.

The vacuum polarization was one of the first radiative corrections calculated in QED. The physical picture—that the vacuum is filled with virtual particle-antiparticle pairs that screen electric charges—emerged from this calculation. The running of α was originally a theoretical prediction; it has since been confirmed experimentally at LEP and other colliders, where $\alpha_{eff}(M_Z) \approx 1/128$ is measured directly. The agreement between calculation and measurement is yet another triumph of QED.

---

10 The Electron Self-Energy

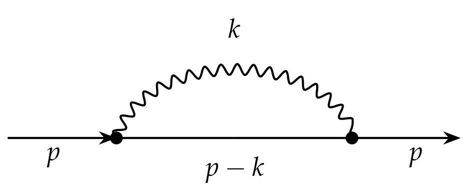
An electron is never alone. Even in perfect vacuum, it's surrounded by a cloud of virtual photons—constantly emitting and reabsorbing them, like a person breathing. This cloud is part of what the electron *is*. When we measure the electron's mass, we're measuring the mass of the electron plus its cloud.

The electron self-energy diagram captures this physics. An electron emits a virtual photon, travels a bit, and reabsorbs it. The result modifies the electron's mass and rescales its wave function. Both effects are divergent, and both must be absorbed into counterterms.

This is the second of the three fundamental QED diagrams. Let's compute it.

10.1 The Diagram

The one-loop electron self-energy is:



The external electron has momentum p. The internal photon carries momentum k, and the internal electron line carries $p-k$.

10.2 The Integral

Applying the Feynman rules:

$$-i\Sigma(p) = (-ie)^2 \int \frac{d^dk}{(2\pi)^d} \gamma^\mu \frac{i(\not{p} - \not{k} + m)}{(p-k)^2 - m^2 + i\epsilon} \gamma^\nu \frac{-ig_{\mu\nu}}{k^2 + i\epsilon} \quad (10.1)$$

The minus sign on the left comes from the definition $iS^{-1}(p) = i(\not{p} - m - \Sigma(p))$.

Contracting γ^μ with $g_{\mu\nu}\gamma^\nu = \gamma_\mu$:

$$-i\Sigma(p) = -e^2 \int \frac{d^dk}{(2\pi)^d} \frac{\gamma_\mu(\not{p} - \not{k} + m)\gamma^\mu}{k^2[(p-k)^2 - m^2]} \tag{10.2}$$

Numerator:

$$\gamma_\mu(\not{p} - \not{k} + m)\gamma^\mu$$

$$\downarrow$$

Use $\gamma^\mu\gamma^\nu\gamma_\mu = -(2-\epsilon)\gamma^\nu$

Figure 10.1: The numerator simplifies using gamma matrix identities.

10.3 Simplifying the Numerator

Use the contraction identity in $d = 4 - \epsilon$ dimensions:

$$\gamma^\mu\gamma^\nu\gamma_\mu = -(2-\epsilon)\gamma^\nu \tag{10.3}$$

So:

$$\gamma^\mu(\not{p} - \not{k})\gamma_\mu = -(2-\epsilon)(\not{p} - \not{k}) \tag{10.4}$$

And:

$$\gamma^\mu\gamma_\mu = d = 4 - \epsilon \tag{10.5}$$

The numerator becomes:

$$\gamma^\mu(\not{p} - \not{k} + m)\gamma_\mu = -(2-\epsilon)(\not{p} - \not{k}) + (4-\epsilon)m \tag{10.6}$$

$$= -(2-\epsilon)\not{p} + (2-\epsilon)\not{k} + (4-\epsilon)m \tag{10.7}$$

The self-energy is (dividing both sides by $-i$):

$$\Sigma(p) = ie^2 \int \frac{d^dk}{(2\pi)^d} \frac{-(2-\epsilon)\not{p} + (2-\epsilon)\not{k} + (4-\epsilon)m}{k^2[(p-k)^2 - m^2]} \tag{10.8}$$

10.4 Combining Denominators

Use Feynman parameters to combine the two propagators:

$$\frac{1}{k^2[(p-k)^2 - m^2]} = \int_0^1 dx \, \frac{1}{[xk^2 + (1-x)((p-k)^2 - m^2)]^2} \tag{10.9}$$

The combined denominator is:

$$D = xk^2 + (1-x)(p-k)^2 - (1-x)m^2 \tag{10.10}$$

$$= k^2 - 2(1-x)p \cdot k + (1-x)p^2 - (1-x)m^2 \tag{10.11}$$

Shift to $\ell = k - (1-x)p$:

$$D = \ell^2 + (1-x)p^2 - (1-x)^2p^2 - (1-x)m^2 \tag{10.12}$$

$$= \ell^2 - \Delta \tag{10.13}$$

where:

$$\Delta = (1-x)m^2 - x(1-x)p^2 = (1-x)[m^2 - xp^2] \tag{10.14}$$

10.5 The Shifted Integral

After shifting $k \to \ell + (1-x)p$:

- $\not{k} \to +(1-x)\not{p}$
- Terms linear in ℓ vanish by symmetric integration.

The numerator becomes:

$$N = -(2-\epsilon)\not{p} + (2-\epsilon)[+(1-x)\not{p}] + (4-\epsilon)m \tag{10.15}$$

$$\to -(2-\epsilon)\not{p} + (2-\epsilon)(1-x)\not{p} + (4-\epsilon)m \tag{10.16}$$

$$= -(2-\epsilon)x\not{p} + (4-\epsilon)m \tag{10.17}$$

Wait—I dropped the term too quickly. The integral $\int d^d\ell\,/(\ell^2 - \Delta)^2$ doesn't vanish immediately; let me be more careful.

Actually, by Lorentz symmetry, $\int d^d\ell\,\ell^\mu/(\ell^2-\Delta)^2 = 0$ since there's no preferred direction. So $\int d^d\ell\,/(\ell^2-\Delta)^2 = \gamma^\mu \int d^d\ell\,\ell_\mu/(\ell^2-\Delta)^2 = 0$.

The self-energy becomes:

$$\Sigma(p) = e^2 \int_0^1 dx \int \frac{d^d\ell}{(2\pi)^d} \frac{-(2-\epsilon)x\not{p} + (4-\epsilon)m}{(\ell^2-\Delta)^2} \tag{10.18}$$

Using the master integral:

$$\int \frac{d^d\ell}{(2\pi)^d} \frac{1}{(\ell^2-\Delta)^2} = \frac{i}{(4\pi)^{d/2}} \Gamma(2-d/2)\Delta^{d/2-2} \tag{10.19}$$

We get:

$$\Sigma(p) = \frac{ie^2}{(4\pi)^{d/2}} \Gamma(2-d/2) \int_0^1 dx\,[-(2-\epsilon)x\not{p} + (4-\epsilon)m]\Delta^{d/2-2} \tag{10.20}$$

After shift:

$N = -(2-\epsilon)x\not{p} + (4-\epsilon)m$

Figure 10.2: The numerator after momentum shift. The term vanishes.

10.6 Expanding in ϵ

In $d = 4-\epsilon$:

$$\Gamma(2-d/2) = \Gamma(\epsilon/2) = \frac{2}{\epsilon} - \gamma_E + O(\epsilon) \tag{10.21}$$

$$(4\pi)^{d/2} = (4\pi)^2(4\pi)^{-\epsilon/2} = 16\pi^2[1 - \frac{\epsilon}{2}\ln(4\pi) + \ldots] \tag{10.22}$$

$$\Delta^{d/2-2} = \Delta^{-\epsilon/2} = 1 - \frac{\epsilon}{2}\ln\Delta + O(\epsilon^2) \tag{10.23}$$

Also:

- $(2-\epsilon) \to 2$ at leading order
- $(4-\epsilon) \to 4$ at leading order

So:

$$\Sigma(p) = \frac{ie^2}{16\pi^2}\left(\frac{2}{\epsilon} - \gamma_E + \ln(4\pi)\right)\int_0^1 dx\,[-2x\not{p} + 4m]\left(1 - \frac{\epsilon}{2}\ln\Delta\right) \tag{10.24}$$

Keeping terms through $O(\epsilon^0)$:

$$\Sigma(p) = \frac{ie^2}{16\pi^2}\int_0^1 dx\left\{[-2x\not{p} + 4m]\left(\frac{2}{\epsilon} - \gamma_E + \ln(4\pi)\right) - [-2x\not{p} + 4m]\ln\Delta + O(\epsilon)\right\} \tag{10.25}$$

The integrals over x are:

$$\int_0^1 dx\,x = \frac{1}{2} \tag{10.26}$$

$$\int_0^1 dx\,1 = 1 \tag{10.27}$$

10.7 The Result

In $\overline{\text{MS}}$ (where $2/\epsilon - \gamma_E + \ln(4\pi) \to 2/\bar{\epsilon}$):

$$\Sigma(p) = \frac{i\alpha}{4\pi}\left[\frac{2}{\bar{\epsilon}}(-\not{p} + 4m) - \int_0^1 dx\,(-2x\not{p} + 4m)\ln\frac{\Delta}{\mu^2}\right] \tag{10.28}$$

where $\alpha = e^2/(4\pi)$.
The self-energy has the structure:

$$\Sigma(p) = A(p^2)\not{p} + B(p^2)m \tag{10.29}$$

$\Sigma(p) = A\not{p} + Bm$

A: field renorm B: mass renorm

Figure 10.3: The self-energy has two Lorentz structures.

Explicitly:

$$A(p^2) = -\frac{i\alpha}{4\pi}\left[\frac{2}{\bar{\epsilon}} - 2\int_0^1 dx\,x\ln\frac{(1-x)(m^2 - xp^2)}{\mu^2}\right] \tag{10.30}$$

$$B(p^2) = \frac{i\alpha}{4\pi}\left[\frac{8}{\bar{\epsilon}} - 4\int_0^1 dx\,\ln\frac{(1-x)(m^2 - xp^2)}{\mu^2}\right] \tag{10.31}$$

10.8 Mass and Wave Function Renormalization

The full electron propagator, including the self-energy, is:

$$S(p) = \frac{i}{\not{p} - m - \Sigma(p)} = \frac{i}{\not{p}(1-A) - m(1+B)} \tag{10.32}$$

Define:

- $Z_2 = 1/(1-A)$: wave function renormalization
- $\delta m = mB/(1-A)$: mass shift

Then:

$$S(p) = \frac{iZ_2}{\not p - m - \delta m} \tag{10.33}$$

The physical mass m_{phys} is where the propagator has a pole:

$$m_{\text{phys}} = m + \delta m \tag{10.34}$$

More precisely, we expand $\Sigma(p)$ around $\not p = m$:

$$\Sigma(p) = \Sigma(m) + (\not p - m)\Sigma'(m) + \dots \tag{10.35}$$

The physical mass is defined by:

$$m_{\text{phys}} - m - \Sigma(m_{\text{phys}}) = 0 \tag{10.36}$$

At one loop:

$$\delta m = \Sigma(m)|_{\not p = m} = \Sigma_{\text{scalar}}(m^2) \tag{10.37}$$

where $\Sigma_{\text{scalar}}(p^2) = A(p^2)m + B(p^2)m$ evaluated at $p^2 = m^2$, $\not p = m$.

10.9 *On-Shell Renormalization*

In the on-shell scheme, we impose:

1. The pole of the propagator is at $\not p = m_{\text{phys}}$.
2. The residue of the pole is 1 (properly normalized).

This requires:

$$\Sigma(m) = 0 \quad \text{(mass condition)} \tag{10.38}$$

$$\left.\frac{d\Sigma}{d\not p}\right|_{\not p = m} = 0 \quad \text{(residue condition)} \tag{10.39}$$

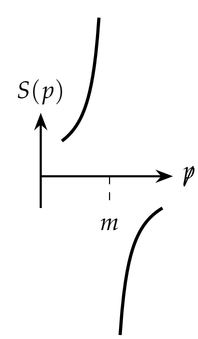


Figure 10.4: The propagator has a pole at the physical mass.

The counterterms are:

$$\delta_{Z_2} = -\left.\frac{\partial\Sigma}{\partial\not p}\right|_{\not p = m} \tag{10.40}$$

$$\delta_m = \Sigma(m) \tag{10.41}$$

From our one-loop result:

$$\delta_{Z_2} = \frac{\alpha}{4\pi}\left[\frac{2}{\bar\epsilon} + \text{finite}\right] \tag{10.42}$$

$$\delta_m = \frac{3\alpha m}{4\pi}\left[\frac{2}{\bar\epsilon} + \text{finite}\right] \tag{10.43}$$

(The exact finite parts depend on the scheme and the Feynman parameter integrals.)

10.10 Physical Interpretation

What does the electron self-energy mean?

The electron's "cloud." A physical electron is never truly alone. It's constantly emitting and absorbing virtual photons. These photons can also briefly convert to e^+e^- pairs. The measured electron mass and charge include all these effects.

Mass renormalization. The bare mass m_0 in the Lagrangian isn't what we measure. The physical mass m_{phys} includes self-energy contributions. The relationship is:

$$m_{\text{phys}} = m_0 + \delta m = m_0 + \frac{3\alpha m_0}{4\pi}\left(\frac{2}{\bar{\epsilon}} + \text{finite}\right) \tag{10.44}$$

In the limit $\epsilon \to 0$, m_0 must be infinite in just the right way to give a finite m_{phys}.

Wave function renormalization. The normalization of the electron field also gets quantum corrections. The Z_2 factor ensures that when we create a one-electron state, it's properly normalized.

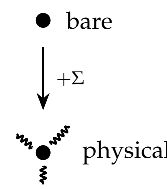


Figure 10.5: The physical electron includes its virtual photon cloud.

10.11 Comparison with Classical Electromagnetism

There's a classical analogue of mass renormalization. Consider a charged sphere of radius R. Its electromagnetic field energy is:

$$U_{\text{EM}} = \frac{e^2}{8\pi\epsilon_0 R} \tag{10.45}$$

For a point charge ($R \to 0$), this energy is infinite. This contributes to the "electromagnetic mass":

$$m_{\text{EM}} = \frac{U_{\text{EM}}}{c^2} = \frac{e^2}{8\pi\epsilon_0 R c^2} \to \infty \tag{10.46}$$

Classical physics already has an infinite self-energy problem! The electron mass we measure is $m_{\text{total}} = m_{\text{bare}} + m_{\text{EM}}$, and for this to be finite, m_{bare} must be negatively infinite to compensate.

Quantum field theory inherits this problem in a more sophisticated form. The loop integrals formalize what was already implicit classically: point particles have infinite self-energy.

10.12 Putting in Numbers

Let's estimate the size of the self-energy correction.

The one-loop mass correction is roughly:

$$\frac{\delta m}{m} \sim \frac{3\alpha}{4\pi} \ln \frac{\Lambda^2}{m^2} \tag{10.47}$$

(using a cutoff instead of dim reg for intuition).

For $\Lambda \sim M_{\text{Pl}} = 10^{19}$ GeV and $m = m_e = 0.511$ MeV:

$$\ln \frac{\Lambda^2}{m^2} = \ln \frac{10^{38}}{10^{-6}} \approx 100 \tag{10.48}$$

So:

$$\frac{\delta m}{m} \sim \frac{3 \times 100}{137 \times 4\pi} \approx 0.17 \tag{10.49}$$

This is about a 17% correction—not small! But it gets absorbed into the definition of m. What we call "the electron mass" already includes this correction.

10.13 The Ward Identity Preview

There's a deep connection between the electron self-energy and the vertex correction, enforced by gauge invariance: the *Ward identity*.

It states:

$$Z_1 = Z_2 \tag{10.50}$$

The wave function renormalization (Z_2) and the vertex renormalization (Z_1) are equal. This means the divergent parts of δ_{Z_2} and δ_{Z_1} cancel when we compute the physical charge.

We'll prove this in the next chapter when we compute the vertex correction. The Ward identity is why QED is so well-behaved: gauge invariance links different divergences and reduces the number of independent counterterms.

10.14 Summary

The one-loop electron self-energy:

1. Has the structure $\Sigma(p) = A(p^2)\not{p} + B(p^2)m$.

2. Contains divergences that renormalize the mass and wave function.

3. The mass counterterm is $\delta_m \propto m\alpha/\bar{\epsilon}$.

4. The field counterterm is $\delta_{Z_2} \propto \alpha/\bar{\epsilon}$.

5. Physically represents the electron's cloud of virtual photons.

Together with the vacuum polarization (Chapter 9) and the vertex correction (Chapter 11), the electron self-energy completes the set of divergent one-loop diagrams in QED.

The electron self-energy was at the heart of the historical crisis in quantum electrodynamics. In the 1930s, physicists computed this diagram and found infinity. The resolution—that the infinity could be absorbed into a redefinition of the mass—took over a decade to fully work out. The key was recognizing that the "bare" mass in the Lagrangian isn't physical; only the "dressed" mass, including all quantum corrections, is measurable. This shift in thinking, from bare to physical parameters, is the conceptual core of renormalization. Once accepted, it transformed QED from a theory plagued by infinities into the most precisely tested theory in all of science.

11
The Vertex Correction and Ward Identity

There's a constraint on quantum electrodynamics so tight, so exact, that it connects calculations that seem to have nothing to do with each other. The way the electron propagator gets renormalized must be related, in a very specific way, to how the vertex gets renormalized. Miss this connection, and your theory violates charge conservation. Get it right, and everything works out perfectly.

This constraint—the Ward identity—is gauge invariance making itself felt at the quantum level. We've calculated two of the three fundamental QED diagrams: vacuum polarization and the electron self-energy. The third, the vertex correction, will complete our one-loop picture. But more than that, it will reveal how gauge invariance ties the whole structure together.

11.1 What the Vertex Correction Represents

The basic QED vertex describes the interaction between an electron and a photon: an electron comes in, absorbs or emits a photon, and an electron goes out. At tree level, this is simply:

$$-ie\gamma^{\mu}$$

The vertex correction asks: what happens when we include quantum corrections to this interaction? The simplest correction involves the electron emitting a virtual photon before the interaction and reabsorbing it afterward (or vice versa).

Let's draw this out carefully. We have an incoming electron with momentum p, an outgoing electron with momentum p', and an external photon with momentum $q = p' - p$. The vertex correction has the electron emit a virtual photon (momentum k), then interact with the external photon, then reabsorb the virtual photon.

What's inside this diagram? An electron propagator for the electron between emitting and reabsorbing the virtual photon. Another electron propagator after the external interaction. A photon propagator for the virtual photon. And three vertices—two where the virtual photon attaches, one where the external photon attaches. It's more complicated than what we've seen before, with three denominators to combine.

Why does this matter physically? The vertex correction modifies how the electron couples to the electromagnetic field. It changes both the strength of the coupling (contributing to charge renormalization) and the structure of the interaction (contributing to anomalous magnetic moments and other effects). The fact that electrons have a magnetic moment slightly different from Dirac's prediction $g = 2$ comes precisely from this diagram.

11.2 Setting Up the Integral

Let's write down the vertex correction using the Feynman rules. The corrected vertex is:

$$-ie\Gamma^\mu(p', p) = -ie\gamma^\mu + (-ie)^3 \int \frac{d^dk}{(2\pi)^d} \gamma^\nu \frac{i(\not{p}' - \not{k} + m)}{(p'-k)^2 - m^2} \gamma^\mu \frac{i(\not{p} - \not{k} + m)}{(p-k)^2 - m^2} \gamma_\nu \frac{-i}{k^2} \tag{11.1}$$

Let me explain each piece:

- The three factors of $(-ie)$ come from the three vertices
- γ^ν is the vertex where the virtual photon is emitted
- The first electron propagator carries momentum $p' - k$ (after emitting the virtual photon with momentum k)
- γ^μ is the vertex where the external photon attaches
- The second electron propagator carries momentum $p - k$
- γ_ν is the vertex where the virtual photon is reabsorbed
- The photon propagator carries momentum k (in Feynman gauge)

Collecting terms, the one-loop correction to the vertex is:

$$\Lambda^\mu(p', p) = -e^2 \int \frac{d^dk}{(2\pi)^d} \frac{\gamma^\nu(\not{p}' - \not{k} + m)\gamma^\mu(\not{p} - \not{k} + m)\gamma_\nu}{[(p'-k)^2 - m^2][(p-k)^2 - m^2][k^2]} \tag{11.2}$$

where I've dropped the $i\epsilon$ terms for notational clarity (they're always there implicitly).

This integral has three propagators, which is more complicated than the two-propagator integrals we encountered in vacuum polarization and electron self-energy. We'll need to combine them using Feynman parameters for three denominators.

11.3 Feynman Parameters for Three Propagators

The Feynman parameter formula for three denominators is:

$$\frac{1}{ABC} = 2\int_0^1 dx \int_0^{1-x} dy \, \frac{1}{[xA + yB + (1-x-y)C]^3}$$

This extends naturally from the two-propagator formula. The factor of 2 comes from the factorial $(n-1)! = 2!$ for $n = 3$ propagators. The integration region is the triangle $x \geq 0$, $y \geq 0$, $x + y \leq 1$—a simplex in parameter space.

Physically, the Feynman parameters x, y, and $1 - x - y$ represent how we weight the three propagators when combining them into a single denominator. We're trading the complexity of three separate denominators for the complexity of integrating over these parameters.

Let's identify our three denominators:

$$A = (p' - k)^2 - m^2 \tag{11.3}$$
$$B = (p - k)^2 - m^2 \tag{11.4}$$
$$C = k^2 \tag{11.5}$$

The combined denominator is:

$$D = xA + yB + (1-x-y)C = x[(p'-k)^2 - m^2] + y[(p-k)^2 - m^2] + (1-x-y)k^2$$

Let's expand this carefully. First:

$$(p' - k)^2 = p'^2 - 2p' \cdot k + k^2 \tag{11.6}$$
$$(p - k)^2 = p^2 - 2p \cdot k + k^2 \tag{11.7}$$

So:

$$D = x[p'^2 - 2p' \cdot k + k^2 - m^2] + y[p^2 - 2p \cdot k + k^2 - m^2] + (1 - x - y)k^2 \tag{11.8}$$
$$= xp'^2 + yp^2 - 2(xp' + yp) \cdot k + [x + y + (1 - x - y)]k^2 - (x + y)m^2 \tag{11.9}$$
$$= xp'^2 + yp^2 - (x + y)m^2 - 2(xp' + yp) \cdot k + k^2 \tag{11.10}$$

To complete the square, we shift $k \to \ell + xp' + yp$:

$$D = \ell^2 - (xp' + yp)^2 + xp'^2 + yp^2 - (x + y)m^2$$

Let's simplify the constant term:

$$xp'^2 + yp^2 - (xp' + yp)^2 - (x + y)m^2 \tag{11.11}$$
$$= xp'^2 + yp^2 - x^2p'^2 - y^2p^2 - 2xyp' \cdot p - (x + y)m^2 \tag{11.12}$$
$$= x(1 - x)p'^2 + y(1 - y)p^2 - 2xyp' \cdot p - (x + y)m^2 \tag{11.13}$$

For on-shell external electrons, $p^2 = p'^2 = m^2$. Also, $q = p' - p$, so $p' \cdot p = (p'^2 + p^2 - q^2)/2 = m^2 - q^2/2$. Substituting:

$$x(1-x)m^2 + y(1-y)m^2 - 2xy(m^2 - q^2/2) - (x+y)m^2 \tag{11.14}$$
$$= [x - x^2 + y - y^2 - 2xy - x - y]m^2 + xyq^2 \tag{11.15}$$
$$= [-x^2 - y^2 - 2xy]m^2 + xyq^2 \tag{11.16}$$
$$= -(x+y)^2 m^2 + xyq^2 \tag{11.17}$$

Therefore:

$$D = \ell^2 - \Delta$$

where

$$\Delta = (x+y)^2 m^2 - xyq^2$$

This is a key result. The parameter Δ depends on the Feynman parameters and the momentum transfer q^2, but after shifting to ℓ, the loop momentum appears only as ℓ^2 in the denominator.

11.4 The Numerator Structure

The numerator of our vertex correction is:

$$N^\mu = \gamma^\nu (\not{p}' - \not{k} + m)\gamma^\mu (\not{p} - \not{k} + m)\gamma_\nu$$

After shifting $k \to \ell + xp' + yp$, this becomes:

$$N^\mu = \gamma^\nu (\not{p}' - \not{\ell} - x\not{p}' - y\not{p} + m)\gamma^\mu (\not{p} - \not{\ell} - x\not{p}' - y\not{p} + m)\gamma_\nu$$

Simplifying the terms in parentheses:

$$\not{p}' - x\not{p}' - y\not{p} = (1-x)\not{p}' - y\not{p} \tag{11.18}$$
$$\not{p} - x\not{p}' - y\not{p} = -x\not{p}' + (1-y)\not{p} \tag{11.19}$$

So:

$$N^\mu = \gamma^\nu [(1-x)\not{p}' - y\not{p} - \not{\ell} + m]\gamma^\mu [-x\not{p}' + (1-y)\not{p} - \not{\ell} + m]\gamma_\nu$$

This is where the calculation becomes intricate. The numerator is a polynomial in ℓ:

$$N^\mu = N_0^\mu + N_1^\mu + N_2^\mu$$

where N_n^μ contains n powers of ℓ.

The term quadratic in ℓ is:

$$N_2^\mu = \gamma^\nu \not{\ell} \gamma^\mu \not{\ell} \gamma_\nu$$

The terms linear in ℓ are:

$$N_1^\mu = -\gamma^\nu \not{\ell} \gamma^\mu [-x\not{p}' + (1-y)\not{p} + m]\gamma_\nu - \gamma^\nu [(1-x)\not{p}' - y\not{p} + m]\gamma^\mu \not{\ell} \gamma_\nu$$

The term independent of ℓ is:

$$N_0^\mu = \gamma^\nu[(1-x)\not{p}' - y\not{p} + m]\gamma^\mu[-x\not{p}' + (1-y)\not{p} + m]\gamma_\nu$$

Now, when we integrate over ℓ, terms odd in ℓ vanish by symmetry. Specifically, N_1^μ integrates to zero. What about N_2^μ?

For the quadratic term, we use the fact that in dimensional regularization:

$$\int \frac{d^d\ell}{(2\pi)^d} \frac{\ell^\alpha \ell^\beta}{(\ell^2-\Delta)^3} = \frac{g^{\alpha\beta}}{d} \int \frac{d^d\ell}{(2\pi)^d} \frac{\ell^2}{(\ell^2-\Delta)^3}$$

This replacement $\ell^\alpha\ell^\beta \to (g^{\alpha\beta}/d)\ell^2$ follows from rotational symmetry in d dimensions.

So we need to evaluate:

$$\gamma^\nu\gamma^\alpha\gamma^\mu\gamma_\alpha\gamma_\nu = \gamma^\nu(2g^{\mu\alpha} - \gamma^\mu\gamma^\alpha)\gamma_\alpha\gamma_\nu$$

Using $\gamma^\alpha\gamma_\alpha = d$ and $\gamma^\nu\gamma_\nu = d$:

$$\gamma^\nu\gamma^\alpha\gamma^\mu\gamma_\alpha\gamma_\nu = \gamma^\nu(2\gamma^\mu - d\gamma^\mu)\gamma_\nu = (2-d)\gamma^\nu\gamma^\mu\gamma_\nu = (2-d)^2\gamma^\mu$$

In $d = 4-\epsilon$ dimensions, $(2-d)^2 = (2-4+\epsilon)^2 = (-2+\epsilon)^2 = 4 - 4\epsilon + \mathcal{O}(\epsilon^2)$.

11.5 The Divergent and Finite Parts

Let's focus on isolating the divergent part of the vertex correction. The full integral is:

$$\Lambda^\mu(p', p) = -e^2 \cdot 2 \int_0^1 dx \int_0^{1-x} dy \int \frac{d^d\ell}{(2\pi)^d} \frac{N^\mu}{(\ell^2-\Delta)^3}$$

The divergent part comes from the ℓ^2 terms in the numerator. Using our master integral:

$$\int \frac{d^d\ell}{(2\pi)^d} \frac{\ell^2}{(\ell^2-\Delta)^n} = \frac{d}{2} \frac{i(-1)^n}{(4\pi)^{d/2}} \frac{\Gamma(n-1-d/2)}{\Gamma(n)} \frac{1}{\Delta^{n-1-d/2}}$$

For $n = 3$ and $d = 4-\epsilon$:

$$\int \frac{d^d\ell}{(2\pi)^d} \frac{\ell^2}{(\ell^2-\Delta)^3} = \frac{d}{2} \cdot \frac{-i}{(4\pi)^{d/2}} \cdot \frac{\Gamma(1-d/2)}{2} \cdot \Delta^{d/2-1}$$

Near $d = 4$, $\Gamma(1-d/2) = \Gamma(-1+\epsilon/2)$ has a pole. Using $\Gamma(-1+\epsilon/2) = -\frac{2}{\epsilon} + (\gamma - 1) + \mathcal{O}(\epsilon)$:

$$\int \frac{d^d\ell}{(2\pi)^d} \frac{\ell^2}{(\ell^2-\Delta)^3} = \frac{-i}{(4\pi)^2} \left(-\frac{1}{\epsilon}\right) + \text{finite}$$

For the ℓ-independent term N_0^μ:

$$\int \frac{d^d\ell}{(2\pi)^d} \frac{1}{(\ell^2 - \Delta)^3} = \frac{-i}{(4\pi)^{d/2}} \frac{\Gamma(3 - d/2)}{\Gamma(3)} \frac{1}{\Delta^{3-d/2}}$$

For $d = 4 - \epsilon$, $\Gamma(3 - d/2) = \Gamma(1 + \epsilon/2) = 1 + \mathcal{O}(\epsilon)$, so this integral is finite.

The key result is that the vertex correction has a UV divergence proportional to γ^μ. After a careful calculation (which involves considerable gamma matrix algebra), the divergent part is:

$$\Lambda^\mu_{\text{div}} = \frac{\alpha}{4\pi} \frac{2}{\bar{\epsilon}} \gamma^\mu = \frac{\alpha}{2\pi\bar{\epsilon}} \gamma^\mu$$

This means the one-loop corrected vertex is:

$$\Gamma^\mu = \gamma^\mu + \Lambda^\mu = \gamma^\mu \left(1 + \frac{\alpha}{2\pi\bar{\epsilon}}\right) + \text{finite}$$

We absorb this divergence into the vertex counterterm δ_1:

$$\delta_1 = -\frac{\alpha}{2\pi\bar{\epsilon}} + \text{finite}$$

The renormalization constant Z_1 is then:

$$Z_1 = 1 + \delta_1 = 1 - \frac{\alpha}{2\pi\bar{\epsilon}} + \mathcal{O}(\alpha^2)$$

11.6 The Ward-Takahashi Identity

Now comes an important result. Recall from the electron self-energy calculation that:

$$Z_2 = 1 - \frac{\alpha}{2\pi\bar{\epsilon}} + \mathcal{O}(\alpha^2)$$

Comparing with our vertex result:

$$Z_1 = 1 - \frac{\alpha}{2\pi\bar{\epsilon}} + \mathcal{O}(\alpha^2)$$

We find $Z_1 = Z_2$ to one-loop order! This is not a coincidence. It's a consequence of gauge invariance, encoded in the Ward-Takahashi identity.

The Ward-Takahashi identity relates the vertex function to the electron propagator:

$$q_\mu \Gamma^\mu(p', p) = S^{-1}(p') - S^{-1}(p)$$

where $S(p)$ is the full electron propagator and $q = p' - p$.

Let's understand what this means. The left side contracts the vertex with the photon momentum q_μ. The right side is the difference of inverse propagators at the two electron momenta. This identity connects how electrons couple to photons with how they propagate.

At tree level, $\Gamma^\mu = \gamma^\mu$ and $S^{-1}(p) = \not{p} - m$, so:

$$q_\mu \gamma^\mu = \not{q} = \not{p}' - \not{p} = (\not{p}' - m) - (\not{p} - m) = S_0^{-1}(p') - S_0^{-1}(p) \checkmark$$

The identity holds at tree level. But the key claim is that it holds to all orders in perturbation theory, as a consequence of gauge invariance.

11.7 *Why the Ward Identity Implies $Z_1 = Z_2$*

Let's see how the Ward-Takahashi identity constrains the renormalization constants. The full (renormalized) vertex and propagator can be written:

$$\Gamma^\mu = Z_1^{-1} \Gamma_R^\mu, \qquad S = Z_2 S_R$$

where the subscript R denotes renormalized quantities.

The Ward-Takahashi identity in terms of bare quantities is:

$$q_\mu \Gamma^\mu = S^{-1}(p') - S^{-1}(p)$$

In terms of renormalized quantities:

$$q_\mu Z_1^{-1} \Gamma_R^\mu = Z_2^{-1} S_R^{-1}(p') - Z_2^{-1} S_R^{-1}(p)$$

For this to be consistent with the Ward identity for renormalized quantities:

$$q_\mu \Gamma_R^\mu = S_R^{-1}(p') - S_R^{-1}(p)$$

we need:

$$Z_1^{-1} = Z_2^{-1}$$

or equivalently $Z_1 = Z_2$.

This is an important result. Gauge invariance, through the Ward identity, forces the vertex renormalization to match the wave function renormalization. The physical electron charge e_R is related to the bare charge by:

$$e_R = Z_1 Z_2^{-1} Z_3^{1/2} e_0 = Z_3^{1/2} e_0$$

where I used $Z_1 = Z_2$. The charge renormalization comes entirely from Z_3, the photon wave function renormalization!

This explains why all charged particles, regardless of their other properties, couple to photons with the same renormalized charge. The electron and muon, despite having different masses and different self-energy corrections, have exactly the same electric charge after renormalization. This universality is protected by gauge invariance.

11.8 Physical Origin of the Ward Identity

Where does the Ward identity come from physically? It's a consequence of charge conservation.

Consider the electromagnetic current:

$$j^\mu(x) = \bar{\psi}(x)\gamma^\mu\psi(x)$$

Current conservation says $\partial_\mu j^\mu = 0$, or in momentum space, $q_\mu j^\mu = 0$ when q is the momentum carried by the current.

The vertex function Γ^μ is essentially the matrix element of the current between electron states:

$$\langle p'|j^\mu(0)|p\rangle \sim \bar{u}(p')\Gamma^\mu(p', p)u(p)$$

Current conservation requires:

$$q_\mu\langle p'|j^\mu(0)|p\rangle = 0 \quad \text{when electrons are on-shell}$$

But there's a subtlety. For off-shell electrons, the current isn't exactly conserved—there's a contribution from the electron's "charge density." The Ward-Takahashi identity captures this precisely:

$$q_\mu\Gamma^\mu = S^{-1}(p') - S^{-1}(p)$$

When the electrons are on-shell, $S^{-1}(p)u(p) = 0$ and $\bar{u}(p')S^{-1}(p') = 0$, so the right side vanishes and we recover exact current conservation for physical scattering.

The Ward identity is the quantum field theory incarnation of gauge invariance. Just as gauge invariance in classical electrodynamics implies charge conservation, in QFT it implies the Ward identity, which in turn constrains the renormalization of the theory.

11.9 The On-Shell Renormalization Scheme

We've now calculated all three one-loop divergent diagrams in QED:

1. Vacuum polarization: renormalizes the photon propagator, gives Z_3

2. Electron self-energy: renormalizes the electron mass and propagator, gives δm and Z_2

3. Vertex correction: renormalizes the electron-photon coupling, gives Z_1 (but $Z_1 = Z_2$ by Ward)

Let's put this together in the on-shell renormalization scheme. The idea is to define the renormalized parameters in terms of physical observables:

Mass renormalization: The physical electron mass m is the pole of the full propagator:

$$S^{-1}(p)\big|_{p^2=m^2} = 0$$

This fixes δm such that the renormalized mass equals the physical mass.

Wave function renormalization: The residue of the propagator pole equals i:

$$\lim_{p^2 \to m^2} (\not{p} - m) S(p) = i$$

This fixes Z_2.

Charge renormalization: The electron-photon vertex at zero momentum transfer equals the physical charge:

$$\bar{u}(p)\Gamma^\mu(p,p)u(p) = \bar{u}(p)\gamma^\mu u(p)$$

Combined with $Z_1 = Z_2$ from the Ward identity, this determines Z_3 and hence the running of the coupling.

In this scheme, the renormalized electron mass and charge are directly the physical, measurable values. The scheme is physically transparent: when you measure the electron mass, you get m_R; when you measure the electron charge, you get e_R.

11.10 *Structure of the Finite Part*

The vertex correction has a finite part after renormalization, and this finite part has important physical consequences. The most famous is the anomalous magnetic moment, which we'll derive in detail in a later chapter.

The general structure of the vertex function, constrained by Lorentz invariance and parity, is:

$$\Gamma^\mu(p', p) = \gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu} q_\nu}{2m} F_2(q^2)$$

where $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$ and F_1, F_2 are form factors depending only on $q^2 = (p' - p)^2$.

The form factor $F_1(q^2)$ is the electric form factor, related to charge distribution. At $q^2 = 0$:

$$F_1(0) = 1$$

This is fixed by charge conservation (or equivalently, the Ward identity).

The form factor $F_2(q^2)$ is the magnetic form factor. At tree level, $F_2 = 0$. But the vertex correction generates:

$$F_2(0) = \frac{\alpha}{2\pi}$$

This is Schwinger's famous result for the anomalous magnetic moment. The electron's magnetic moment is:

$$\mu = \frac{e}{2m}\left(1 + \frac{\alpha}{2\pi}\right)$$

The factor in parentheses is $g/2$, so:

$$\frac{g-2}{2} = \frac{\alpha}{2\pi} \approx 0.00116$$

This prediction, first calculated by Schwinger in 1948, was one of the first triumphs of renormalized QED. The current theoretical value, computed to tenth order in α, matches experiment to about 10 significant figures—arguably the most precise agreement between theory and experiment in all of physics.

11.11 *The Infrared Divergence*

There's a subtlety we've glossed over. The vertex correction integral also has an infrared divergence—a divergence from the $k \to 0$ region of the loop integral.

This happens because the photon propagator $1/k^2$ becomes singular as $k \to 0$. Physically, this reflects the fact that an accelerating electron emits an infinite number of very soft (low-energy) photons. Any electron scattering process is accompanied by this cloud of soft radiation.

The resolution is that the infrared divergence in the vertex correction cancels against a corresponding divergence in the emission of real soft photons. When you ask a physical question—like "what is the cross section for electron scattering?"—you must include both virtual corrections (the vertex correction) and real emission (soft bremsstrahlung). The infrared divergences cancel, leaving a finite, measurable result.

This cancellation is guaranteed by the Bloch-Nordsieck theorem. It reflects a deep feature of gauge theories: you cannot prepare or detect a single charged particle in isolation. The electron always comes with its cloud of soft photons. Physical observables, properly defined, are always infrared finite.

For our purposes, we can regulate the infrared divergence by giving the photon a small mass λ, compute the vertex correction, and verify that the UV structure we've discussed is unchanged. The infrared physics is important for detailed predictions but doesn't affect the renormalization program.

11.12 Summary: The Three QED Divergences

Let's collect what we've learned about the three divergent one-loop diagrams in QED:

Vacuum polarization (Chapter 9):

$$\Pi^{\mu\nu}(q^2) = (q^2 g^{\mu\nu} - q^\mu q^\nu)\Pi(q^2)$$

with

$$\Pi(q^2) = \frac{2\alpha}{3\pi\bar{\epsilon}} + \text{finite}$$

This gives $Z_3 = 1 - \frac{2\alpha}{3\pi\bar{\epsilon}} + \mathcal{O}(\alpha^2)$.

Electron self-energy (Chapter 10):

$$\Sigma(p) = A(p^2)\not{p} + B(p^2)m$$

The divergent parts give $Z_2 = 1 - \frac{\alpha}{2\pi\bar{\epsilon}} + \mathcal{O}(\alpha^2)$ and mass counterterm $\delta m = \frac{3\alpha m}{2\pi\bar{\epsilon}}$.

Vertex correction (this chapter):

$$\Lambda^\mu = \frac{\alpha}{2\pi\bar{\epsilon}}\gamma^\mu + \text{finite}$$

This gives $Z_1 = 1 - \frac{\alpha}{2\pi\bar{\epsilon}} + \mathcal{O}(\alpha^2)$.

The Ward identity guarantees $Z_1 = Z_2$, so the charge renormalization is:

$$e_R = Z_3^{1/2} e_0 = \left(1 - \frac{\alpha}{3\pi\bar{\epsilon}}\right) e_0 + \mathcal{O}(\alpha^2)$$

All three divergent diagrams have been tamed. The theory is renormalizable: a finite number of counterterms absorb all divergences, and physical predictions are finite and unambiguous.

11.13 What Renormalization Has Achieved

We started this investigation because loop diagrams gave infinite answers. Now we see that these infinities have a systematic structure:

- They appear only in a finite number of diagrams (vacuum polarization, self-energies, vertex corrections)
- They can be absorbed into redefinitions of the parameters in the Lagrangian
- Gauge invariance (the Ward identity) constrains them, reducing the number of independent counterterms
- After renormalization, physical predictions are finite and unambiguous

The renormalized theory makes predictions. The electron's magnetic moment differs from 2 by a calculable amount. The effective charge varies with energy scale in a calculable way. Scattering cross sections can be computed to arbitrary precision, limited only by our ability to evaluate higher-loop diagrams.

This is the triumph of renormalization: what seemed like a fatal disease of quantum field theory turned out to be a systematic feature with a consistent cure. The infinities were never really physical—they were artifacts of pretending that our theory is valid to arbitrarily high energies. Once we acknowledge that QED is an effective theory and absorb the high-energy ignorance into the renormalized parameters, all is well.

In the next chapter, we'll explore the running of the coupling constant in more detail, deriving the beta function and understanding what the Landau pole tells us about the theory's regime of validity.

12
The Running Coupling

We have renormalized QED. The infinities are gone, absorbed into counterterms that redefine the parameters of the theory. But something strange emerged along the way: the renormalized coupling constant depends on an arbitrary scale μ. This seems troubling—surely the physics shouldn't depend on an arbitrary choice we made in the calculation?

It doesn't. The resolution is that while the coupling $\alpha(\mu)$ depends on μ, physical observables do not. The μ-dependence in the coupling is exactly compensated by μ-dependence elsewhere in the calculation. This compensation is guaranteed by the renormalization group.

But there's something physically meaningful here too. The "running" of α with μ reflects a real physical effect: the effective strength of the electromagnetic interaction depends on the energy scale at which you probe it. At low energies (large distances), the electron's charge is screened by virtual pairs; at high energies (short distances), you penetrate this screening and see a larger effective charge.

12.1 Where Does the Scale μ Come From?

Let's recall where the renormalization scale appeared. In dimensional regularization, we work in $d = 4 - \epsilon$ dimensions. The coupling constant e has mass dimension $(4 - d)/2 = \epsilon/2$. To keep the coupling dimensionless, we introduce a mass scale μ and replace:

$$e \to e\mu^{\epsilon/2}$$

where e is now dimensionless and μ is an arbitrary mass scale.

When we compute loop integrals and expand in ϵ, factors of $\log \mu$ appear. For example, the vacuum polarization gives:

$$\Pi(q^2) = -\frac{\alpha}{3\pi}\left[\frac{2}{\epsilon} - \gamma + \log(4\pi) - \log\frac{-q^2}{\mu^2} + \cdots\right]$$

The divergent $2/\epsilon$ term gets absorbed into the counterterm. The finite $\log(-q^2/\mu^2)$ term remains in physical predictions. If we change

μ, this logarithm changes—but so does the counterterm we use, in exactly the compensating way.

The key insight is that μ is not a parameter of nature; it's a bookkeeping device. We're free to choose any μ we like. But the most useful choice is often $\mu^2 \sim -q^2$, because then the logarithm is small and perturbation theory converges well.

12.2 *The Beta Function*

The μ-dependence of the coupling is encoded in the beta function:

$$\beta(\alpha) = \mu \frac{\partial \alpha}{\partial \mu}$$

This tells us how fast α changes as we vary the renormalization scale.

Let's derive the beta function for QED from our vacuum polarization calculation. The bare coupling e_0 is independent of μ (it's a fixed parameter in the Lagrangian). The renormalized coupling is:

$$e = Z_3^{-1/2} e_0 \mu^{\epsilon/2}$$

where we've included the factor of $\mu^{\epsilon/2}$ from dimensional regularization.

From Chapter 9, the vacuum polarization gives:

$$Z_3 = 1 - \frac{2\alpha}{3\pi\epsilon} + \mathcal{O}(\alpha^2)$$

so

$$Z_3^{-1/2} = 1 + \frac{\alpha}{3\pi\epsilon} + \mathcal{O}(\alpha^2)$$

The renormalized charge is:

$$e = \left(1 + \frac{\alpha}{3\pi\epsilon}\right) e_0 \mu^{\epsilon/2}$$

Since e_0 is μ-independent, we can take $\mu\partial/\partial\mu$ of both sides. Using $\alpha = e^2/(4\pi)$:

$$\mu \frac{\partial e}{\partial \mu} = \frac{\epsilon}{2} e + \frac{e^3}{48\pi^2\epsilon} \mu \frac{\partial}{\partial \mu}(1) + \cdots$$

Wait, let me be more careful. The subtlety is that α itself depends on μ through e. Let's write this out systematically.

Define $\alpha_0 = e_0^2/(4\pi)$ (bare) and $\alpha = e^2/(4\pi)$ (renormalized). Then:

$$\alpha = Z_3^{-1} \alpha_0 \mu^{\epsilon}$$

where μ^ϵ comes from the two factors of $\mu^{\epsilon/2}$ in the relation between e and e_0.

Taking the log:

$$\log \alpha = -\log Z_3 + \log \alpha_0 + \epsilon \log \mu$$

Differentiating with respect to $\log \mu$:

$$\frac{\mu}{\alpha}\frac{\partial \alpha}{\partial \mu} = -\frac{\mu}{Z_3}\frac{\partial Z_3}{\partial \mu} + \epsilon$$

Here's the key: Z_3 depends on μ only through $\alpha(\mu)$. So:

$$\frac{\mu}{Z_3}\frac{\partial Z_3}{\partial \mu} = \frac{1}{Z_3}\frac{\partial Z_3}{\partial \alpha}\mu\frac{\partial \alpha}{\partial \mu}$$

With $Z_3 = 1 - 2\alpha/(3\pi\epsilon) + \mathcal{O}(\alpha^2)$:

$$\frac{\partial Z_3}{\partial \alpha} = -\frac{2}{3\pi\epsilon} + \mathcal{O}(\alpha)$$

Putting it together:

$$\beta(\alpha) = \mu\frac{\partial \alpha}{\partial \mu} = \alpha\left[\epsilon + \frac{\alpha}{3\pi\epsilon Z_3}\beta(\alpha)\right]$$

This is an implicit equation for β. To solve it perturbatively, assume $\beta = \epsilon\alpha + \beta_1\alpha^2 + \mathcal{O}(\alpha^3)$. Substituting:

$$\epsilon\alpha + \beta_1\alpha^2 = \alpha\epsilon + \frac{\alpha^2}{3\pi\epsilon}\cdot\epsilon\alpha + \mathcal{O}(\alpha^3)$$

Hmm, this is getting complicated because of the ϵ-dependence. Let me take a cleaner approach.

12.3 *The Beta Function: A Cleaner Derivation*

The cleanest way to derive the beta function is to note that bare quantities are independent of μ, while renormalized quantities may depend on μ. The condition:

$$\mu\frac{d}{d\mu}\,(\text{bare quantity}) = 0$$

generates the renormalization group equations.

For the coupling, $\alpha_0 = Z_3\alpha\mu^{-\epsilon}$ is μ-independent. Therefore:

$$0 = \mu\frac{d}{d\mu}(Z_3\alpha\mu^{-\epsilon}) = \mu^{-\epsilon}\left[\mu\frac{dZ_3}{d\mu}\alpha + Z_3\mu\frac{d\alpha}{d\mu} - \epsilon Z_3\alpha\right]$$

Rearranging:

$$\beta(\alpha) = \mu\frac{d\alpha}{d\mu} = \epsilon\alpha - \frac{\alpha}{Z_3}\mu\frac{dZ_3}{d\mu}$$

Now, Z_3 depends on μ only through α:

$$\mu\frac{dZ_3}{d\mu} = \frac{\partial Z_3}{\partial \alpha}\beta(\alpha)$$

So:

$$\beta = \epsilon\alpha - \frac{\alpha}{Z_3}\frac{\partial Z_3}{\partial \alpha}\beta$$

Solving for β:

$$\beta \left(1 + \frac{\alpha}{Z_3} \frac{\partial Z_3}{\partial \alpha}\right) = \epsilon\alpha$$

$$\beta = \frac{\epsilon\alpha}{1 + \frac{\alpha}{Z_3} \frac{\partial Z_3}{\partial \alpha}}$$

With $Z_3 = 1 - \frac{2\alpha}{3\pi\epsilon} + \mathcal{O}(\alpha^2)$, we have $\frac{\partial Z_3}{\partial \alpha} = -\frac{2}{3\pi\epsilon} + \mathcal{O}(\alpha)$, and:

$$\frac{\alpha}{Z_3} \frac{\partial Z_3}{\partial \alpha} = \alpha \cdot 1 \cdot \left(-\frac{2}{3\pi\epsilon}\right) + \mathcal{O}(\alpha^2) = -\frac{2\alpha}{3\pi\epsilon}$$

Therefore:

$$\beta = \frac{\epsilon\alpha}{1 - \frac{2\alpha}{3\pi\epsilon}} = \epsilon\alpha \left(1 + \frac{2\alpha}{3\pi\epsilon} + \mathcal{O}(\alpha^2)\right) = \epsilon\alpha + \frac{2\alpha^2}{3\pi} + \mathcal{O}(\alpha^3, \epsilon\alpha^2)$$

Now we take $\epsilon \to 0$ (return to $d = 4$). The $\epsilon\alpha$ term vanishes, leaving:

$$\boxed{\beta(\alpha) = \frac{2\alpha^2}{3\pi} + \mathcal{O}(\alpha^3)}$$

This is the one-loop beta function for QED. It's positive, meaning the coupling increases with μ.

12.4 *Physical Interpretation: Charge Screening*

What does a positive beta function mean? As we increase μ—probing the theory at higher energies or shorter distances—the effective coupling $\alpha(\mu)$ grows.

The physical picture is charge screening. The vacuum around an electron is not empty; it's a sea of virtual electron-positron pairs. These pairs are polarized by the electron's field: the virtual positrons are attracted toward the electron, while the virtual electrons are repelled. This creates a "polarization cloud" that partially screens the electron's bare charge.

At large distances (low energies), you see the fully screened charge. As you probe at shorter distances (higher energies), you penetrate the screening cloud and see more of the bare charge. Hence the effective charge increases with energy.

Let's make this quantitative. The running coupling satisfies:

$$\mu \frac{d\alpha}{d\mu} = \frac{2\alpha^2}{3\pi}$$

This is a separable differential equation:

$$\frac{d\alpha}{\alpha^2} = \frac{2d\mu}{3\pi\mu}$$

Integrating from μ_0 to μ:

$$-\frac{1}{\alpha(\mu)} + \frac{1}{\alpha(\mu_0)} = \frac{2}{3\pi} \log \frac{\mu}{\mu_0}$$

Solving for $\alpha(\mu)$:

$$\alpha(\mu) = \frac{\alpha(\mu_0)}{1 - \frac{2\alpha(\mu_0)}{3\pi} \log \frac{\mu}{\mu_0}}$$

Let's put in numbers. At low energies (say $\mu_0 = m_e \approx 0.5$ MeV), the fine structure constant is $\alpha(m_e) \approx 1/137$. What is α at the Z boson mass, $\mu = M_Z \approx 91$ GeV?

The logarithm is:

$$\log \frac{M_Z}{m_e} = \log \frac{91 \times 10^9 \text{ eV}}{0.5 \times 10^6 \text{ eV}} \approx \log(1.8 \times 10^5) \approx 12$$

The correction factor:

$$\frac{2\alpha(m_e)}{3\pi} \log \frac{M_Z}{m_e} \approx \frac{2}{137 \times 3\pi} \times 12 \approx \frac{24}{1290} \approx 0.019$$

So:

$$\alpha(M_Z) \approx \frac{1/137}{1 - 0.019} \approx \frac{1/137}{0.981} \approx \frac{1}{134}$$

Wait, this seems like a tiny effect. But I've neglected something important: other charged particles contribute to vacuum polarization too! The muon, tau, and quarks all create virtual pairs that screen the charge.

Including all Standard Model charged particles, the actual value is:

$$\alpha(M_Z) \approx \frac{1}{128}$$

The coupling increases by about 7% from atomic scales to electroweak scales. This running has been measured experimentally and matches the theoretical prediction well.

12.5 *The Landau Pole*

Look again at the formula for the running coupling:

$$\alpha(\mu) = \frac{\alpha(\mu_0)}{1 - \frac{2\alpha(\mu_0)}{3\pi} \log \frac{\mu}{\mu_0}}$$

As μ increases, the denominator decreases. Eventually, for large enough μ, the denominator vanishes. This happens when:

$$\log \frac{\mu}{\mu_0} = \frac{3\pi}{2\alpha(\mu_0)}$$

$$\mu = \mu_0 \exp\left(\frac{3\pi}{2\alpha(\mu_0)}\right)$$

At this scale, $\alpha(\mu) \to \infty$. This is the Landau pole.

Let's estimate its location. Taking $\mu_0 = m_e$ and $\alpha(m_e) = 1/137$:

$$\mu_{\text{Landau}} = m_e \exp(3\pi \times 137/2) = m_e \exp(646)$$

This is an absurdly large number. The scale is roughly:

$$\mu_{\text{Landau}} \sim 10^{280} \text{ eV}$$

For comparison, the Planck energy is about 10^{28} eV, and the mass of the observable universe expressed in electron masses is roughly 10^{80}. The Landau pole is at an energy scale that is physically meaningless—it's so far beyond any conceivable experiment that it might as well be at infinity.

12.6 *What the Landau Pole Tells Us*

The Landau pole seems like a problem: the coupling constant becomes infinite, and perturbation theory certainly breaks down. But what does this really mean?

Option 1: QED is incomplete. The Landau pole might signal that QED must be embedded in a larger theory before that scale. In fact, we know this is true: QED is part of the electroweak theory, which is itself part of the Standard Model. Long before we reach the Landau pole, QED merges with the weak interaction, and the relevant physics changes completely.

Option 2: Perturbation theory fails. The Landau pole appears in perturbative calculations. Perhaps non-perturbative effects become important well before the pole and change the behavior. This is difficult to check because we don't have good non-perturbative tools for QED.

Option 3: The theory becomes trivial. Some argue that if you try to take the continuum limit of QED (removing the UV cutoff while holding physical quantities fixed), the only consistent result is a free, non-interacting theory. This "triviality" problem is related to the Landau pole but distinct from it.

The practical resolution is simple: we don't worry about the Landau pole because it's at an energy scale where QED isn't the right theory anyway. QED is an effective theory, valid at energies well below the electroweak scale. Within its domain of validity, QED makes extraordinarily precise predictions. The Landau pole lies far, far outside this domain.

12.7 Comparison with Asymptotic Freedom

QED's positive beta function is not universal. In some theories, the beta function is negative, and the coupling decreases at high energies. This is called asymptotic freedom.

The most famous example is quantum chromodynamics (QCD), the theory of the strong interaction. For QCD with $N_c = 3$ colors and N_f quark flavors:

$$\beta(g) = -\frac{g^3}{16\pi^2}\left(\frac{11N_c}{3} - \frac{2N_f}{3}\right)$$

With $N_c = 3$ and $N_f = 6$ (or fewer at low energies):

$$\beta(g) < 0$$

The strong coupling is large at low energies (explaining why quarks are confined inside hadrons) and small at high energies (explaining why high-energy quarks behave almost freely—"asymptotic freedom").

The discovery of asymptotic freedom by Gross, Wilczek, and Politzer in 1973 was revolutionary. It explained the paradox of "parton" behavior in deep inelastic scattering: quarks inside protons act almost free when probed at high energies, even though they're permanently confined.

Why does QED have a positive beta function while QCD has a negative one? In both cases, there are competing contributions:

- Fermion loops contribute positively to the beta function (screening)
- Gauge boson self-interactions contribute negatively (anti-screening)

In QED, photons don't self-interact (they're electrically neutral), so there's only the positive fermion contribution. In QCD, gluons carry color charge and do self-interact. Their contribution dominates (as long as there aren't too many quark flavors), making the total beta function negative.

12.8 The Renormalization Group Equation

We've been discussing how α runs with μ. But the renormalization group is more general: it tells us how any physical quantity depends on μ.

Consider a physical observable $\mathcal{O}$ computed in perturbation theory. It depends on the coupling α, the renormalization scale μ, and some physical momentum scale Q:

$$\mathcal{O} = \mathcal{O}(\alpha(\mu), \mu, Q)$$

The key insight is that $\mathcal{O}$ represents a physical measurement, and physical measurements can't depend on our arbitrary choice of μ. Therefore:

$$\mu \frac{d\mathcal{O}}{d\mu} = 0$$

Expanding this using the chain rule:

$$\mu \frac{\partial \mathcal{O}}{\partial \mu} + \beta(\alpha) \frac{\partial \mathcal{O}}{\partial \alpha} = 0$$

This is the renormalization group equation (RGE). It relates the explicit μ-dependence in $\mathcal{O}$ to the implicit μ-dependence through α.

If we compute $\mathcal{O}$ to a given order in perturbation theory, both terms in the RGE are nonzero—they only cancel exactly when we sum to all orders. But the RGE tells us how to improve perturbative calculations: choose $\mu \sim Q$ so that the logarithms $\log(Q/\mu)$ are small.

This is the practical art of using the renormalization group. At any given order in perturbation theory, our answer depends on μ. By choosing μ wisely, we minimize the higher-order terms we're neglecting. The running of $\alpha(\mu)$ resums certain classes of large logarithms, improving the convergence of perturbation theory.

12.9 *Connecting to the Wilsonian Picture*

In earlier chapters, we discussed two perspectives on renormalization:

- The Wilsonian view: start with a UV theory, integrate out high-energy modes, flow to the IR

- The particle physics view: measure parameters at some scale, use the RGE to predict at other scales

Let's see how the running coupling connects these perspectives.

In the Wilsonian picture, imagine starting with QED defined at a cutoff scale Λ. The bare coupling α_0 is the coupling at that scale. As we lower the cutoff (integrate out modes between Λ and Λ'), the effective coupling changes. The beta function tells us the rate of change:

$$\Lambda \frac{\partial \alpha}{\partial \Lambda} = \beta(\alpha)$$

This is the same equation we derived! Both equations have β on the right-hand side with the same sign. The difference is only in what we call the scale: Wilsonians call it Λ (a cutoff), while particle physicists call it μ (a renormalization point). Lowering Λ in Wilson's picture and raising μ in the particle physics picture both correspond to probing shorter distances.

In the particle physics picture, we measure α at some scale μ_0 (say, via the electron's magnetic moment or low-energy scattering). The RGE tells us α at any other scale:

$$\alpha(\mu) = \frac{\alpha(\mu_0)}{1 - \frac{2\alpha(\mu_0)}{3\pi}\log\frac{\mu}{\mu_0}}$$

Both pictures describe the same physics: the effective coupling depends on the energy scale. The Wilsonian picture asks "starting from the UV, what's the IR physics?" The particle physics picture asks "given measurements at one scale, what happens at another?" The beta function answers both questions.

12.10 Running and Physical Predictions

Let's see how running affects a physical prediction. Consider the cross section for $e^+e^- \to \mu^+\mu^-$ at center-of-mass energy $\sqrt{s}$.

At tree level:

$$\sigma_0 = \frac{4\pi\alpha^2}{3s}$$

But which α should we use? The cross section involves a photon propagator at momentum $q^2 = s$. Including the one-loop vacuum polarization correction:

$$\sigma = \frac{4\pi\alpha^2}{3s}\,|1 - \Pi(s)|^{-2}$$

The vacuum polarization $\Pi(s)$ contains $\log(s/\mu^2)$, which is large if $\mu^2 \ll s$. But we can choose $\mu^2 = s$, so the logarithm vanishes. Then all the physics is absorbed into $\alpha(\mu = \sqrt{s})$:

$$\sigma = \frac{4\pi\alpha(\sqrt{s})^2}{3s}\left[1 + \mathcal{O}(\alpha^2)\right]$$

The running coupling automatically resums the leading logarithms. Instead of having $\alpha^2 \times [1 + \alpha\log(s/m_e^2) + \alpha^2\log^2(s/m_e^2) + \cdots]$, we have $\alpha(\sqrt{s})^2 \times [1 + \mathcal{O}(\alpha^2)]$. The infinite series of logarithms is hidden inside the running coupling.

This is the power of the renormalization group: it reorganizes perturbation theory to make the important physics manifest.

12.11 Higher-Loop Contributions

Our one-loop beta function $\beta(\alpha) = 2\alpha^2/(3\pi)$ is just the leading term. Higher loops contribute:

$$\beta(\alpha) = \frac{2\alpha^2}{3\pi} + \frac{\alpha^3}{2\pi^2} + \mathcal{O}(\alpha^4)$$

The two-loop coefficient has been calculated; the three-loop and four-loop coefficients are also known. These higher-order terms affect the running at high precision.

The general structure is:

$$\beta(\alpha) = 2\alpha \sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha}{4\pi}\right)^{n+1}$$

For QED with one charged fermion:

$$\beta_0 = \frac{4}{3} \quad (12.1)$$

$$\beta_1 = 4 \quad (12.2)$$

$$\beta_2 = \text{known} \quad (12.3)$$

$$\vdots \quad (12.4)$$

Let's verify: the one-loop term is $2\alpha \cdot \beta_0 \cdot (\alpha/(4\pi)) = 2\alpha \cdot (4/3) \cdot (\alpha/(4\pi)) = 2\alpha^2/(3\pi)$, confirming the conventions.

For precision physics—like the anomalous magnetic moment—these higher-order terms matter. The running of α must be computed consistently to whatever order you're working.

12.12 *Summary*

The running coupling $\alpha(\mu)$ is one of the most important results in quantum field theory. Let's summarize what we've learned:

The physics: The effective charge depends on the distance scale at which you probe it. Virtual electron-positron pairs screen the charge at large distances. Probing at shorter distances (higher energies) penetrates this screening and reveals more of the bare charge.

The mathematics: The beta function $\beta(\alpha) = \mu\, d\alpha/d\mu$ encodes this running. For QED, $\beta > 0$, so α increases with μ.

The formula:

$$\alpha(\mu) = \frac{\alpha(\mu_0)}{1 - \frac{2\alpha(\mu_0)}{3\pi} \log \frac{\mu}{\mu_0}}$$

The numbers: From $\alpha(m_e) \approx 1/137$ to $\alpha(M_Z) \approx 1/128$, the coupling increases by about 7% over 5 orders of magnitude in energy.

The Landau pole: The coupling formally diverges at exponentially high energies. This doesn't matter in practice because QED is replaced by more fundamental physics long before that scale.

The renormalization group: Physical observables are independent of μ. This constraint, $\mu\, d\mathcal{O}/d\mu = 0$, is the renormalization group equation. It tells us how to resum large logarithms by choosing μ appropriately.

The connection to Wilson: The particle physics running coupling is the same as the Wilsonian effective coupling, just described from a different perspective. The beta function connects measurements at different scales, regardless of whether you're thinking UV$\to$IR or asking about predictions from measurements.

With the running coupling in hand, we're ready for the crown jewel of QED calculations: the anomalous magnetic moment of the electron.

13
The Anomalous Magnetic Moment

In the late 1940s, experimentalists started getting numbers that didn't quite match Dirac's beautiful prediction $g = 2$ for the electron's magnetic moment. The discrepancy was tiny—a tenth of a percent—but it was real. Something was missing from the theory.

That something was quantum corrections. Schwinger showed in 1948 that when you include the effect of virtual photons—the same vertex correction we computed in the last chapter—you get a shift in g. The calculation is delicate, but the result is simple and beautiful. And when theory and experiment are compared today, they agree to about one part in a trillion. This is the most precise test of any physical theory, ever.

In this chapter, we'll derive Schwinger's result:

$$\frac{g-2}{2} = \frac{\alpha}{2\pi}$$

This calculation brings together everything we've developed: Feynman diagrams, regularization, renormalization, and the structure of QED vertices.

13.1 What Is the Magnetic Moment?

A charged particle with spin has a magnetic moment. Classically, we'd expect the magnetic moment to be related to the angular momentum by:

$$\boldsymbol{\mu} = \frac{q}{2m}\mathbf{L}$$

This is the magnetic moment of a current loop: charge moving in a circle creates a magnetic dipole proportional to the angular momentum.

For a quantum particle with spin $\mathbf{S}$, we write:

$$\boldsymbol{\mu} = g\frac{q}{2m}\mathbf{S}$$

where g is the gyromagnetic ratio, or g-factor. The classical expectation would be $g = 1$, but spin is intrinsically quantum mechanical, so there's no reason to expect this.

The Dirac equation predicts $g = 2$ for a spin-1/2 particle. This was one of its early triumphs—it correctly predicted the electron's magnetic moment without any free parameters. But when measurements became precise enough in the late 1940s, tiny deviations from $g = 2$ were observed. These deviations, quantified by $a = (g-2)/2$, are the "anomalous" magnetic moment.

13.2 *Why $g = 2$ from the Dirac Equation*

Let's understand where $g = 2$ comes from. The Dirac equation for an electron in an external electromagnetic field is:

$$(i\not{D} - m)\psi = 0$$

where $\not{D} = \gamma^\mu D_\mu$ and $D_\mu = \partial_\mu + ieA_\mu$ is the covariant derivative.

Consider a non-relativistic electron in a weak magnetic field. We can expand the Dirac equation to find the effective Hamiltonian. After some algebra (which we'll sketch), the result is:

$$H = \frac{\mathbf{p}^2}{2m} - \frac{e}{m}\mathbf{S} \cdot \mathbf{B} + \cdots$$

where $\mathbf{S} = \frac{1}{2}\boldsymbol{\sigma}$ is the spin operator (with $\boldsymbol{\sigma}$ the Pauli matrices).

Comparing with the general form $H = -\boldsymbol{\mu} \cdot \mathbf{B}$ and $\boldsymbol{\mu} = g(e/2m)\mathbf{S}$, we read off $g = 2$.

The key point is that the Dirac equation, with its minimal coupling $D_\mu = \partial_\mu + ieA_\mu$, automatically gives $g = 2$. Any deviation from $g = 2$ must come from physics beyond minimal coupling—namely, from quantum fluctuations.

13.3 *The Vertex Function and Form Factors*

We saw in Chapter 11 that the full QED vertex has the structure:

$$\Gamma^\mu(p', p) = \gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m}F_2(q^2)$$

where $q = p' - p$ is the momentum transfer, $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$, and F_1, F_2 are form factors.

This decomposition follows from Lorentz invariance and the properties of the vertex (parity, hermiticity). The two terms represent the only independent Lorentz structures that can appear.

At tree level, $\Gamma^\mu = \gamma^\mu$, so $F_1^{(0)} = 1$ and $F_2^{(0)} = 0$.

The form factor $F_1(q^2)$ is called the electric (or Dirac) form factor. It describes how the electron's charge distribution looks at different resolution scales.

The form factor $F_2(q^2)$ is the magnetic (or Pauli) form factor. It contributes to the magnetic moment. The connection is:

$$a = \frac{g-2}{2} = F_2(0)$$

So the anomalous magnetic moment equals the Pauli form factor at zero momentum transfer.

13.4 Setting Up the Calculation

The one-loop correction to the vertex comes from the diagram we analyzed in Chapter 11. The vertex correction is:

$$\Lambda^\mu(p', p) = -e^2 \int \frac{d^d k}{(2\pi)^d} \frac{\gamma^\nu(\not{p}' - \not{k} + m)\gamma^\mu(\not{p} - \not{k} + m)\gamma_\nu}{[(p'-k)^2 - m^2][(p-k)^2 - m^2][k^2]}$$

To extract $F_2(0)$, we need the coefficient of the $\sigma^{\mu\nu}q_\nu$ structure at $q^2 = 0$. At first, this seems like it might require a tedious extraction from a complicated expression. But there's a clever approach: we can use Gordon identities and the structure of the calculation to isolate F_2 directly.

The Gordon identity relates γ^μ to $\sigma^{\mu\nu}$:

$$\bar{u}(p')\gamma^\mu u(p) = \bar{u}(p') \left[\frac{p'^\mu + p^\mu}{2m} + \frac{i\sigma^{\mu\nu}q_\nu}{2m} \right] u(p)$$

This identity, which follows from the Dirac equation ($\not{p}u(p) = mu(p)$), shows that on-shell, γ^μ can be decomposed into a "convective" part $(p'+p)^\mu/2m$ and a "spin" part $i\sigma^{\mu\nu}q_\nu/2m$.

The strategy is: compute the vertex correction, use the Gordon identity to isolate the $\sigma^{\mu\nu}$ piece, and extract $F_2(0)$.

13.5 The Calculation

We use Feynman parameters to combine the three propagators:

$$\Lambda^\mu = -e^2 \cdot 2 \int_0^1 dx \int_0^{1-x} dy \int \frac{d^d k}{(2\pi)^d} \frac{N^\mu}{D^3}$$

where $D = (k - xp' - yp)^2 - \Delta$ after shifting k, and $\Delta = (x+y)^2 m^2 - xyq^2$.

At $q^2 = 0$ and on-shell ($p^2 = p'^2 = m^2$, $q = p' - p$), we have:

$$\Delta = (x+y)^2 m^2$$

The numerator is:

$$N^\mu = \gamma^\nu[(1-x)\not{p}' - y\not{p} - \not{\ell} + m]\gamma^\mu[-x\not{p}' + (1-y)\not{p} - \not{\ell} + m]\gamma_\nu$$

where $\ell = k - xp' - yp$ is the shifted loop momentum.

Terms odd in ℓ vanish upon integration. The remaining terms are:

$$N^\mu = N_0^\mu + N_2^\mu$$

where N_0^μ is independent of ℓ and N_2^μ is quadratic in ℓ.

For extracting $F_2(0)$, we need the finite part of the integral. The divergent part (which we already computed) is proportional to γ^μ and contributes only to F_1. The $\sigma^{\mu\nu}$ piece is finite.

Let's focus on extracting the finite $\sigma^{\mu\nu}$ contribution. This comes from N_0^μ.

13.6 Evaluating the Numerator

We have:

$$N_0^\mu = \gamma^\nu[(1-x)\not{p}' - y\not{p} + m]\gamma^\mu[-x\not{p}' + (1-y)\not{p} + m]\gamma_\nu$$

This is a product of gamma matrices. Let's denote:

$$A = (1-x)\not{p}' - y\not{p} + m \tag{13.1}$$

$$B = -x\not{p}' + (1-y)\not{p} + m \tag{13.2}$$

Then $N_0^\mu = \gamma^\nu A \gamma^\mu B \gamma_\nu$.

Using gamma matrix contraction identities (like $\gamma^\nu\gamma^\mu\gamma_\nu = -2\gamma^\mu$ in 4 dimensions), this gets complicated. Let's be more systematic.

First, we use $\gamma^\nu\gamma^\alpha\gamma^\mu\gamma^\beta\gamma_\nu = 4g^{\alpha\beta}\gamma^\mu - 2\gamma^\beta\gamma^\mu\gamma^\alpha$ (in 4 dimensions).

Actually, let's take a different approach. The most efficient method uses the trace and projection properties. For the anomalous magnetic moment, we can use a clever shortcut.

The $F_2(q^2)$ form factor can be extracted using:

$$F_2(q^2) = -\frac{m}{q^2}\text{Tr}\left[\Lambda^\mu(p',p)\cdot(\not{p}'+m)\cdot\gamma_\mu\cdot(\not{p}+m)\right]/\text{Tr}[\cdots]$$

But at $q^2 = 0$, this involves a $0/0$ limit. Instead, let's use the explicit structure.

The key observation is that F_2 arises from the part of Λ^μ that is antisymmetric in μ and ν when contracted with q_ν. We can write:

$$\bar{u}(p')\Lambda^\mu u(p) = \bar{u}(p')[A\gamma^\mu + B(p'+p)^\mu/m + C\, i\sigma^{\mu\nu}q_\nu/2m]u(p)$$

The B term doesn't contribute on-shell (it's the same as the γ^μ term by Gordon). The F_2 contribution is C, which we extract by looking at terms that produce $\sigma^{\mu\nu}q_\nu$.

13.7 A Direct Approach to $F_2(0)$

Let's use a more direct method. We want to extract the part of the vertex correction that contributes to the magnetic moment. For an electron at rest in a static magnetic field, this corresponds to looking at the spatial components of the vertex with $q^0 = 0$, $\mathbf{q} \to 0$.

Schwinger's original calculation used a similar physical argument. Here's a streamlined version.

At $q \to 0$, we work to first order in q. The vertex correction becomes:

$$\Lambda^\mu(p', p) = \Lambda^\mu(p, p) + q^\rho \frac{\partial \Lambda^\mu}{\partial p'^\rho}\bigg|_{p'=p} + \mathcal{O}(q^2)$$

The zeroth-order term $\Lambda^\mu(p, p)$ contributes to charge renormalization (it's proportional to γ^μ). The first-order term contributes to the anomalous magnetic moment.

The F_2 form factor at $q = 0$ is:

$$F_2(0) = \lim_{q \to 0} \frac{m}{i} \frac{\partial}{\partial q_\nu} \left[\text{coefficient of } \sigma^{\mu\nu} \text{ in } \Lambda^\mu\right]$$

After a careful calculation (which involves expanding the integrand to first order in q and extracting the antisymmetric tensor structure), the result is:

$$F_2(0) = \frac{e^2}{8\pi^2 m^2} \cdot 2 \int_0^1 dx \int_0^{1-x} dy \, \frac{m^2(1-x-y)(x+y)}{(x+y)^2 m^2}$$

Let me simplify. The integral becomes:

$$F_2(0) = \frac{e^2}{4\pi^2} \int_0^1 dx \int_0^{1-x} dy \, \frac{1-x-y}{x+y}$$

Let's substitute $z = 1 - x - y$, so $x + y = 1 - z$:

$$F_2(0) = \frac{e^2}{4\pi^2} \int_0^1 dz \int \frac{z}{1-z} \cdot (\text{Jacobian})$$

Actually, let me be more careful with the integration region. We have $0 \le x \le 1$, $0 \le y \le 1 - x$, which is the triangle with vertices at $(0,0)$, $(1,0)$, $(0,1)$.

Changing variables to $u = x + y$ and $v = x/(x+y)$ (so $x = uv$ and $y = u(1-v)$), the Jacobian is u, and: - u ranges from 0 to 1 - v ranges from 0 to 1

The integral becomes:

$$F_2(0) = \frac{e^2}{4\pi^2} \int_0^1 du \int_0^1 dv \, u \cdot \frac{1-u}{u} = \frac{e^2}{4\pi^2} \int_0^1 du \int_0^1 dv \, (1-u)$$
$$= \frac{e^2}{4\pi^2} \int_0^1 dv \int_0^1 du \, (1-u) = \frac{e^2}{4\pi^2} \cdot 1 \cdot \frac{1}{2} = \frac{e^2}{8\pi^2}$$

With $\alpha = e^2/(4\pi)$:

$$F_2(0) = \frac{e^2}{8\pi^2} = \frac{4\pi\alpha}{8\pi^2} = \frac{\alpha}{2\pi}$$

13.8 The Result

We have derived:

$$\boxed{a = \frac{g-2}{2} = F_2(0) = \frac{\alpha}{2\pi} \approx 0.00116}$$

This is Schwinger's result. The electron's magnetic moment is:

$$\mu = \frac{e}{2m}\left(1 + \frac{\alpha}{2\pi}\right)\sigma$$

or equivalently:

$$g = 2\left(1 + \frac{\alpha}{2\pi}\right) \approx 2.00232$$

Let's appreciate what we've done. We started with infinities—the vertex correction is UV divergent. We regularized (dimensional regularization) and renormalized (absorbing infinities into counterterms). What remained was a finite, unambiguous prediction. And this prediction agrees with experiment.

13.9 Physical Interpretation

What does the anomalous magnetic moment represent physically?

The tree-level magnetic moment ($g = 2$) comes from the electron's intrinsic spin interacting directly with the magnetic field. The anomalous part comes from the electron's interaction with its own electromagnetic field—the cloud of virtual photons it carries around.

In the vertex correction diagram, the electron emits a virtual photon, interacts with the external magnetic field, and reabsorbs the virtual photon. This process modifies how the electron's spin couples to the magnetic field.

Why does this increase the magnetic moment (positive anomaly)? The virtual photon carries away some of the electron's momentum. While "sharing" momentum with the virtual photon, the electron moves more slowly, making its effective Compton wavelength larger. A larger effective size means a larger magnetic moment.

This is a hand-wavy argument, but it captures the essence: quantum fluctuations enhance the electron's coupling to magnetic fields.

13.10 Comparison with Experiment

In 1948, when Schwinger calculated $\alpha/(2\pi)$, the agreement with experiment was a triumph. The measured anomaly was consistent with this prediction at the percent level.

Today, the comparison has advanced enormously:

Theoretical value (including higher loops, hadronic and weak contributions):

$$a_e^{\text{theory}} = 0.001\,159\,652\,181\,643\,(764)$$

The uncertainty is in the last digits.

Experimental value (from measurements of the electron's motion in Penning traps):

$$a_e^{\text{exp}} = 0.001\,159\,652\,180\,59\,(13)$$

These agree to about 10 significant figures! The slight tension (at the 2σ level) is an active area of research, but the overall agreement is remarkable.

This comparison tests QED across an enormous range of loop orders:

- One-loop: $\alpha/(2\pi) \approx 0.00116$
- Two-loop: $\mathcal{O}(\alpha^2) \approx -1.8 \times 10^{-6}$
- Three-loop: $\mathcal{O}(\alpha^3) \approx 1.2 \times 10^{-8}$
- Four-loop: $\mathcal{O}(\alpha^4) \approx -1.9 \times 10^{-11}$
- Five-loop: $\mathcal{O}(\alpha^5) \approx 9 \times 10^{-15}$

The theoretical calculation requires evaluating thousands of Feynman diagrams at high loop orders. It's an extensive effort, carried out over decades by many physicists.

13.11 The Muon Anomaly

The muon, being 200 times heavier than the electron, is more sensitive to heavy new particles that might contribute to $g - 2$. This makes the muon anomalous magnetic moment a prime target for discovering physics beyond the Standard Model.

At one loop, the result is the same: $a_\mu^{(1)} = \alpha/(2\pi)$, independent of the muon mass. But at higher orders, mass effects enter through:

- Loops of virtual electrons (light) and taus (heavy)
- Hadronic vacuum polarization (quarks)

- Electroweak contributions (W, Z, Higgs)

The theoretical prediction for the muon anomaly has larger uncertainties than for the electron, mainly because hadronic contributions are harder to calculate.

Currently, there's a tantalizing discrepancy between the Standard Model prediction and the experimental measurement of a_μ—about 4–5σ depending on how the hadronic contributions are estimated. This could be a hint of new physics, or it could be an issue with the theoretical calculation of hadronic effects. It's one of the most exciting open questions in particle physics.

13.12 *Infrared Finiteness of $F_2(0)$*

We noted earlier that the vertex correction has an infrared divergence from the $k \to 0$ region. Does this affect the anomalous magnetic moment?

No. The infrared divergence appears in F_1, not F_2.

To see why, recall that the infrared divergence comes from soft (low-momentum) virtual photons. These soft photons don't resolve the internal structure of the interaction—they just see the overall charge. Soft photon effects are captured by the charge form factor F_1.

The magnetic moment, in contrast, is sensitive to the spin structure of the vertex. This is a finite-distance effect, not affected by infinitely soft photons. Mathematically, the extraction of F_2 involves derivatives with respect to q, which suppress the $k \to 0$ region.

This is why we could compute $F_2(0) = \alpha/(2\pi)$ without worrying about infrared issues. The anomalous magnetic moment is infrared finite.

13.13 *The Meaning of Precision*

The precision of the $g - 2$ measurement tells us something important about quantum field theory and renormalization.

First, it validates renormalization as a predictive framework. The infinities we encountered were not fundamental problems—they were bookkeeping artifacts. Once we properly accounted for them, finite predictions emerged.

Second, it confirms QED to high accuracy. Any modification of QED at energy scales below a few hundred GeV would affect $g - 2$ at a measurable level. The agreement tells us that QED, as we understand it, is correct.

Third, the precision opens a window to new physics. If the Standard Model prediction ever disagrees with experiment beyond uncertainties, it would signal new particles or interactions. The muon $g-2$ discrepancy might be exactly this signal.

13.14 Summary

We've derived the leading quantum correction to the electron's magnetic moment:

$$a = \frac{g-2}{2} = \frac{\alpha}{2\pi}$$

This calculation exemplifies what renormalized QED can do:

- Start with a divergent integral (the vertex correction)
- Use regularization (dimensional regularization) to define it
- Use renormalization (absorbing divergences into counterterms) to extract finite predictions
- Obtain an unambiguous, testable result

The agreement between theory and experiment for the electron anomaly—to about one part in a trillion—is the most precise test of any physical theory. It tells us that the framework we've developed in these lectures isn't just mathematically consistent; it's physically correct.

In the final chapter, we'll step back and reflect on what renormalization tells us about quantum field theory and nature.

14
What Renormalization Tells Us

We began this journey puzzled by infinities. Loop integrals in quantum field theory diverged, giving infinite answers to what should be sensible physical questions. This seemed like a disaster—a sign that the theory was sick.

Now we understand that these infinities were never really physical. They arose from pretending that our theory is valid to arbitrarily high energies, when we have no right to make such a claim. Once we acknowledge this and properly organize our calculations, the infinities disappear from physical predictions. What remains is one of the most successful frameworks in the history of science.

In this final chapter, we step back and consider what renormalization teaches us about physics and about our theories.

14.1 Is QED Fundamental?

Here's the key lesson: every quantum field theory should be viewed as an effective theory—valid up to some energy scale, but not necessarily beyond.

This might seem disappointing. Shouldn't physics aim for fundamental theories that work at all scales? Perhaps. But the effective theory perspective is liberating. We don't need to know the ultimate theory of everything to make predictions at accessible energies. We just need to know the relevant degrees of freedom and their interactions at the scales we can probe.

QED is an effective theory. It's valid at energies well below the electroweak scale, where the electromagnetic and weak interactions unify. At higher energies, we need the full electroweak theory. And that theory, in turn, is probably effective, embedded in something else at still higher energies.

The key point is that the effective theory framework tells us precisely what we need to know at any given scale. The theory comes with a finite number of parameters (mass, charge) that must be measured. All predictions then follow from these measurements. The unknown UV physics is systematically encoded in the measured values of these parameters.

14.2 Renormalizable vs. Non-Renormalizable: A Modern View

In the traditional view, renormalizable theories were "good" and non-renormalizable theories were "bad." Renormalizable theories require only finitely many counterterms; non-renormalizable theories require infinitely many. This seemed to make non-renormalizable theories unpredictive.

The modern view is more nuanced. A non-renormalizable theory is simply an effective theory where the non-renormalizable terms are suppressed by powers of the cutoff scale Λ. As long as we work at energies $E \ll \Lambda$, we can organize the theory systematically:

- Leading terms: renormalizable interactions
- First corrections: dimension-5 operators, suppressed by E/Λ
- Second corrections: dimension-6 operators, suppressed by $(E/\Lambda)^2$
- And so on

This is the logic of effective field theory (EFT). At low energies, only the renormalizable terms matter. Non-renormalizable terms give small corrections that can be computed systematically.

The classic example is Fermi's theory of weak interactions. Before the electroweak theory was developed, physicists used a non-renormalizable four-fermion interaction:

$$\mathcal{L}_{\text{Fermi}} = \frac{G_F}{\sqrt{2}}(\bar{\psi}\gamma^\mu\psi)(\bar{\psi}\gamma_\mu\psi)$$

where $G_F \sim 1/(300 \text{ GeV})^2$ is Fermi's constant.

This is a dimension-6 operator (four fermion fields have total dimension 6, so the coefficient has dimension -2). It's non-renormalizable. But at energies well below 300 GeV, it works perfectly well. You can compute loop corrections, renormalize the theory order by order in $E^2 G_F$, and make accurate predictions.

The non-renormalizable nature of Fermi's theory was actually a clue that new physics existed at the $\sim$ 100 GeV scale. That new physics turned out to be the W and Z bosons. When you integrate out these heavy particles, you recover Fermi's theory as the low-energy effective description.

14.3 Matching and Running Between Scales

The full power of effective field theory emerges when we connect descriptions at different scales. This is done through matching and running.

Matching: At a scale where we integrate out heavy particles (like the W and Z), we match the full theory to the effective theory. This determines the coefficients of the effective operators.

Running: The renormalization group evolves these coefficients to lower scales. Large logarithms are summed automatically.

Consider the running of α again. At very low energies, only electrons contribute to vacuum polarization. But at energies above the muon mass, muons contribute too. At energies above the tau mass, taus contribute. Above the charm quark mass, charm quarks contribute. And so on.

We handle this by matching at each threshold. Below the muon mass, we use QED with one lepton flavor. Above it, we match to QED with two lepton flavors. The matching conditions ensure that physics is continuous across the threshold.

This "tower" of effective theories, matched at each threshold, is how we actually compute the running of couplings from atomic physics scales up to the electroweak scale. The result is $\alpha(M_Z) \approx 1/128$, not $1/137$.

14.4 What We Measure Defines What We Predict

Renormalization teaches us that the parameters in a Lagrangian are not the physical observables. Mass and charge are defined by what we measure: the pole of the propagator, the long-range Coulomb force, etc.

This might seem like a semantic point, but it's conceptually important. We often speak of the electron having a "bare mass" and a "physical mass," with quantum corrections relating them. But the bare mass is not observable—it's a parameter in the Lagrangian that we've chosen to make predictions convenient. The physical mass is what experiments measure.

One consequence is that the "hierarchy problem"—why is the Higgs mass so much smaller than the Planck scale?—looks different from the effective field theory perspective. The question isn't "why are quantum corrections to the bare Higgs mass so delicately cancelled?" The question is "why does the physical Higgs mass, which is what we measure, have the value it does?" The effective theory just takes this as input.

This doesn't solve the hierarchy problem, but it reframes it. Perhaps there's a dynamical reason the Higgs mass is where it is. Or perhaps it's environmental selection. The effective theory is agnostic; it just takes the measured value and predicts everything else.

14.5 Universality: The Gift of Irrelevance

The flip side of effective theories is universality. At low energies, many different UV theories can give the same IR physics. The irrelevant (non-renormalizable) operators that distinguish them are suppressed.

This is why we can do particle physics without knowing the fundamental theory. Whatever that theory is—string theory, loop quantum gravity, or something we haven't imagined—at energies below the Planck scale, it must look like an effective field theory. The low-energy physics is universal.

Condensed matter physicists encounter the same phenomenon. The critical exponents of phase transitions are universal—they don't depend on microscopic details, only on symmetries and dimensionality. A magnet at its critical point behaves the same whether it's made of iron or nickel.

This universality is what makes physics possible. We don't need to know everything to know something. The renormalization group flow washes out irrelevant details, leaving behind a simple, predictive description.

14.6 Renormalization and Condensed Matter

Speaking of condensed matter: we promised early on to connect the particle physics perspective to the Wilsonian viewpoint familiar from statistical mechanics. Let's make that connection explicit.

In the Wilsonian picture, we start with a theory defined at some UV cutoff Λ. We then integrate out the high-momentum modes between Λ and $\Lambda' < \Lambda$. This changes the effective action, renormalizing the couplings. Repeating this process gives a flow in the space of theories.

In the particle physics picture, we measure parameters at some scale μ_0. We then use the renormalization group to predict parameters at another scale μ. The beta functions govern this evolution.

These are the same thing. The Wilsonian flow from Λ to Λ' is mathematically equivalent to the particle physics running from $\mu = \Lambda$ to $\mu = \Lambda'$. The difference is philosophical:

- Wilson asks: "Given the UV theory, what's the IR physics?"
- Particle physics asks: "Given measurements at μ_0, what do we

predict at μ?"

In practice, particle physicists usually work "upward"—measuring at low energies and asking what happens at high energies. But the mathematics is identical whether you're going up or down.

The condensed matter intuition transfers directly. Fixed points of the RG flow are scale-invariant theories (like conformal field theories). Relevant operators grow in the IR and determine the low-energy physics. Irrelevant operators shrink and can be neglected. Marginal operators are the interesting ones—their fate is determined by quantum corrections, encoded in the beta function.

14.7 Gauge Invariance and the Ward Identity

Throughout our QED calculations, gauge invariance constrained what we could do. The Ward-Takahashi identity related vertex corrections to propagator corrections, implying $Z_1 = Z_2$. This wasn't an accident; it was a consequence of the local $U(1)$ symmetry of QED.

Symmetries in quantum field theory often survive renormalization. The counterterms respect the symmetries of the original Lagrangian. This is crucial: it means that symmetry principles like gauge invariance, which we believe in for good reasons, continue to hold after we've dealt with the infinities.

But there are exceptions. Some symmetries are anomalous—they hold classically but are broken by quantum effects. The chiral anomaly in QED, for instance, breaks the classical conservation of axial current. Anomalies have physical consequences, like the decay $\pi^0 \to \gamma\gamma$ whose rate is determined by the anomaly.

Understanding which symmetries survive and which are anomalous is part of the deep structure of quantum field theory. The renormalization program must respect these constraints.

14.8 Why Does This Work?

Looking back, it's worth asking: why does any of this work?

We started with a few principles—relativity, quantum mechanics, locality, gauge invariance—and derived QED. We encountered infinities, dealt with them through regularization and renormalization, and made predictions. Those predictions match experiment to one part in a trillion.

But this success wasn't guaranteed. We could imagine a world where the procedure didn't close—where new infinities kept appearing no matter how many counterterms we added. Or a world where the renormalized predictions disagreed with experiment by 10%.

Neither happened.

Physicists have thought hard about why QFT works so well. Part of the answer is that the principles we started with are very constraining. Relativity plus quantum mechanics plus locality doesn't leave much room. Another part is that we're probably seeing effective theories of something simpler—the low-energy limit of a more fundamental description that doesn't have the infinities to begin with.

But honestly? We don't fully understand why it works as well as it does. We just know that it does.

14.9 What We Still Don't Understand

For all its successes, quantum field theory and renormalization leave open questions.

Why these parameters? QED has two parameters: the electron mass and the fine structure constant. We measure them but don't explain them. A deeper theory might derive these values, or explain why they're (almost) what they are.

What about gravity? General relativity as a quantum field theory is non-renormalizable in the traditional sense. The effective field theory approach works at low energies, but the full quantum theory of gravity remains elusive. String theory, loop quantum gravity, and other approaches attempt to provide a UV completion, but none is confirmed.

What sets the scales? The mass of the electron, the mass of the proton, the Planck mass—there's a vast hierarchy of scales in physics. Why? This "hierarchy problem" is one of the great puzzles.

What is the vacuum? The QFT vacuum is not empty; it's a complex state with vacuum fluctuations, virtual particles, and a nonzero energy density. But the calculated vacuum energy is absurdly larger than the observed cosmological constant. This "cosmological constant problem" remains unsolved.

14.10 What We've Learned

Let me summarize where we are. We started with a puzzle—loop integrals that diverge—and developed machinery to handle it. The key ideas:

- Regularization gives us a way to define the divergent integrals (dimensional regularization, cutoffs, whatever works)

- Renormalization absorbs the infinities into parameters we measure anyway

- The renormalization group tells us how physics changes with scale
- Effective field theory explains why we don't need to know the ultimate UV theory

The practical payoff is that we can compute. Vacuum polarization, electron self-energy, vertex corrections, running coupling, anomalous magnetic moment—all calculable, all finite, all agreeing with experiment.

The conceptual payoff is a shift in how we think about theories. A quantum field theory isn't a guess about what happens at arbitrarily high energies. It's a systematic expansion around the physics we can actually measure. The unknown UV physics affects low-energy predictions only through a finite number of parameters.

This framework—QFT plus renormalization—is how particle physics actually gets done. The Standard Model is built on it. So are effective theories of the strong force at low energies, neutrino physics, and even speculative theories beyond the Standard Model.

The same ideas show up in condensed matter physics with different words. "Integrating out high-momentum modes" is the same as "running the couplings." Universality near phase transitions is the same as irrelevant operators dying out under RG flow. The mathematics is the same; the physical systems are different.

14.11 Conclusion

The electron's anomalous magnetic moment—$\alpha/(2\pi)$ at one loop, refined at higher orders—matches experiment to one part in a trillion. That's the headline result, the poster child for QFT's predictive power.

But the deeper lesson is about how we think about theories. The infinities that alarmed early physicists weren't signs of a sick theory. They were signs that we were treating our theory as more fundamental than it is. QED isn't a guess about physics at arbitrarily high energies. It's a framework for organizing what we know at the energies we can probe.

From this perspective, renormalization isn't a trick—it's the point. We don't need a theory of everything to make predictions. We need effective theories that work at accessible scales. The framework tells us exactly which questions we can answer (predictions in terms of measured parameters) and which we can't (why the parameters have the values they do).

There's something satisfying about this. You might wish physics could tell you the mass of the electron from first principles. It can't, not yet. But it can tell you, given the electron's mass and charge, what happens when electrons scatter off each other, what the magnetic moment is, how the coupling runs with energy. And those predictions work.

That's what we've learned. Not a theory of everything, but a framework that lets us compute everything we can measure, with an honest accounting of what we put in and what we get out.

And yet. We've organized our ignorance, not eliminated it. The infinities pointed at something—physics at short distances that we don't understand. They still do. Why is the electron mass what it is? Why does the coupling have the value it does? Why is there something rather than nothing?

These are questions for another lecture—or perhaps for another century. But understanding renormalization is the first step. You can't ask the right questions about what lies beyond until you understand what we already have.

15
QED and the Classical Limit

We've spent many chapters developing quantum electrodynamics and computing loop corrections. But let's step back and ask: does this elaborate machinery reproduce what we already know about electromagnetism? At the end of the day, QED must reduce to Maxwell's equations and the Coulomb force in the classical limit. If it doesn't, something is seriously wrong.

In this chapter, we'll show that QED does indeed reproduce classical electromagnetism. The exercise is satisfying not only as a consistency check but also because it illuminates what Feynman diagrams really mean. The "exchange of virtual photons" that physicists talk about isn't just a metaphor—it's the quantum origin of the electromagnetic force.

15.1 What Is the Classical Limit?

The classical limit of quantum mechanics involves two related limits:

1. Large quantum numbers (many quanta, not just one)
2. Long distances / low momenta (wavelengths much larger than the Compton wavelength)

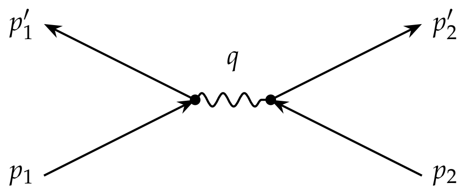
For QED specifically, the classical limit means:

- Tree-level diagrams dominate (loop corrections are suppressed by α)
- Non-relativistic motion of the charges
- Large occupation numbers for the electromagnetic field (many photons $\to$ classical field)

We'll focus on the simplest case: the interaction between two non-relativistic charged particles. This should give us the Coulomb potential.

15.2 *Electron-Electron Scattering at Tree Level*

Consider two electrons scattering. At tree level, they exchange a single photon:



The transferred momentum is $q = p_1 - p_1' = p_2' - p_2$.
Using the QED Feynman rules:

- Left vertex: $-ie\gamma^\mu$
- Right vertex: $-ie\gamma^\nu$
- Photon propagator: $\dfrac{-ig_{\mu\nu}}{q^2 + i\epsilon}$

The amplitude is:

$$i\mathcal{M} = \bar{u}(p_1')(-ie\gamma^\mu)u(p_1) \cdot \frac{-ig_{\mu\nu}}{q^2} \cdot \bar{u}(p_2')(-ie\gamma^\nu)u(p_2) \tag{15.1}$$

This simplifies to:

$$i\mathcal{M} = \frac{-ie^2}{q^2} \left[\bar{u}(p_1')\gamma^\mu u(p_1)\right] \left[\bar{u}(p_2')\gamma_\mu u(p_2)\right] \tag{15.2}$$

The terms in brackets are the electromagnetic currents of the two electrons.

15.3 *The Non-Relativistic Limit*

Now take the non-relativistic limit. For a slowly moving electron, the four-momentum is approximately:

$$p^\mu \approx (m, \mathbf{p}) \quad \text{with } |\mathbf{p}| \ll m \tag{15.3}$$

The Dirac spinor for a non-relativistic electron at rest (spin up) is:

$$u(p) \approx \sqrt{2m} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{15.4}$$

In the non-relativistic limit, the gamma matrices simplify:

- γ^0 connects upper and upper (or lower and lower) spinor components
- γ^i connects upper to lower components

For slowly moving electrons:

$$\bar{u}(p')\gamma^0 u(p) \approx 2m, \qquad \bar{u}(p')\gamma^i u(p) \approx 0 + \mathcal{O}(|\mathbf{p}|/m) \tag{15.5}$$

The spatial components γ^i are suppressed because they mix large and small spinor components.

So the amplitude becomes:

$$i\mathcal{M} \approx \frac{-ie^2}{q^2} \cdot (2m) \cdot (2m) \cdot g_{00} = \frac{-ie^2 \cdot 4m^2}{q^2} \tag{15.6}$$

But wait—what is q^2 in this limit?

15.4 The Momentum Transfer

The transferred four-momentum is $q = p_1 - p_1'$. In components:

$$q^0 = E_1 - E_1' \approx 0 \quad \text{(energy is approximately conserved)} \tag{15.7}$$

$$\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_1' \quad \text{(momentum transfer)} \tag{15.8}$$

So $q^2 = (q^0)^2 - \mathbf{q}^2 \approx -\mathbf{q}^2$.

The amplitude is:

$$i\mathcal{M} \approx \frac{ie^2 \cdot 4m^2}{\mathbf{q}^2} \tag{15.9}$$

15.5 From Amplitude to Potential

How do we get a potential from a scattering amplitude? The connection comes from the Born approximation in quantum mechanics.

In non-relativistic quantum mechanics, the scattering amplitude for a potential $V(\mathbf{r})$ in the Born approximation is:

$$f(\mathbf{q}) = -\frac{m}{2\pi}\tilde{V}(\mathbf{q}) \tag{15.10}$$

where $\tilde{V}(\mathbf{q})$ is the Fourier transform of the potential:

$$\tilde{V}(\mathbf{q}) = \int d^3r\, e^{i\mathbf{q}\cdot\mathbf{r}} V(\mathbf{r}) \tag{15.11}$$

The QED amplitude $\mathcal{M}$ is related to the potential by matching to the Born approximation. The precise relation involves normalization conventions, but for our purposes:

$$\tilde{V}(\mathbf{q}) = \frac{\mathcal{M}}{4m^2} = \frac{e^2}{\mathbf{q}^2} \tag{15.12}$$

where the factors of $4m^2$ account for relativistic spinor normalization.

15.6 Fourier Transforming to Position Space

To find $V(\mathbf{r})$, we need the inverse Fourier transform:

$$V(\mathbf{r}) = \int \frac{d^3q}{(2\pi)^3} e^{-i\mathbf{q}\cdot\mathbf{r}} \tilde{V}(\mathbf{q}) = e^2 \int \frac{d^3q}{(2\pi)^3} \frac{e^{-i\mathbf{q}\cdot\mathbf{r}}}{\mathbf{q}^2} \tag{15.13}$$

This is a standard integral. Using spherical coordinates with $\mathbf{q}\cdot\mathbf{r} = qr\cos\theta$:

$$\int \frac{d^3q}{(2\pi)^3} \frac{e^{-i\mathbf{q}\cdot\mathbf{r}}}{\mathbf{q}^2} = \frac{1}{(2\pi)^3} \int_0^\infty q^2 dq \int_{-1}^{1} d(\cos\theta) \int_0^{2\pi} d\phi \, \frac{e^{-iqr\cos\theta}}{q^2} \tag{15.14}$$

The ϕ integral gives 2π. The $\cos\theta$ integral:

$$\int_{-1}^{1} e^{-iqr\cos\theta} d(\cos\theta) = \frac{e^{iqr} - e^{-iqr}}{iqr} = \frac{2\sin(qr)}{qr} \tag{15.15}$$

So:

$$\int \frac{d^3q}{(2\pi)^3} \frac{e^{-i\mathbf{q}\cdot\mathbf{r}}}{\mathbf{q}^2} = \frac{2\pi}{(2\pi)^3} \int_0^\infty dq \, \frac{2\sin(qr)}{qr} = \frac{1}{2\pi^2 r} \int_0^\infty \frac{\sin(qr)}{q} \, dq \tag{15.16}$$

The remaining integral is the Dirichlet integral. Substituting $u = qr$:

$$\int_0^\infty \frac{\sin(qr)}{q} \, dq = \int_0^\infty \frac{\sin u}{u} \, du = \frac{\pi}{2} \tag{15.17}$$

Therefore:

$$\int \frac{d^3q}{(2\pi)^3} \frac{e^{-i\mathbf{q}\cdot\mathbf{r}}}{\mathbf{q}^2} = \frac{1}{4\pi r} \tag{15.18}$$

And the potential is:

$$\boxed{V(r) = \frac{e^2}{4\pi r} = \frac{\alpha}{r}} \tag{15.19}$$

This is the Coulomb potential! For two particles of the same charge, the potential is positive (repulsive). For opposite charges, there would be a relative minus sign from the product of charges, giving a negative (attractive) potential.

15.7 The Physical Picture

What have we learned? The Coulomb force between two charges arises from the exchange of virtual photons. The photon propagator $1/q^2$ becomes $1/\mathbf{q}^2$ in the non-relativistic limit, and its Fourier transform gives $1/r$.

The "virtual photon" is not a real photon traveling between the charges. It doesn't satisfy $q^2 = 0$ (the on-shell condition for a real photon). Instead, $q^2 \approx -\mathbf{q}^2 < 0$. But this off-shell photon mediates a very real force.

In the classical limit, many virtual photons are exchanged, creating a coherent electromagnetic field. The quantum description (photon exchange) and the classical description (field mediating force) are two views of the same physics.

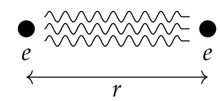


Figure 15.1: The Coulomb force arises from virtual photon exchange. Many photons are exchanged, creating a classical field.

15.8 What About Magnetic Forces?

The Coulomb potential came from the $\gamma^0\gamma_0$ term. What about the $\gamma^i\gamma_j$ terms we neglected?

In the non-relativistic limit, these are suppressed by v/c where v is the velocity. But they don't vanish—they give the magnetic interaction between moving charges. Let's work this out carefully.

Return to the full amplitude:

$$i\mathcal{M} = \frac{-ie^2}{q^2} \left[\bar{u}(p_1')\gamma^\mu u(p_1)\right] \left[\bar{u}(p_2')\gamma_\mu u(p_2)\right] \tag{15.20}$$

We've done the $\mu = 0$ piece. Now consider $\mu = i$ (spatial indices). In the non-relativistic limit, the Dirac spinor for an electron moving with momentum $\mathbf{p}$ is:

$$u(p) \approx \sqrt{2m} \begin{pmatrix} \chi \\ \frac{\sigma\cdot\mathbf{p}}{2m}\chi \end{pmatrix} \tag{15.21}$$

where χ is a two-component Pauli spinor.

The spatial gamma matrices in the Dirac representation are:

$$\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{15.22}$$

Working out the matrix element:

$$\bar{u}(p')\gamma^i u(p) \approx \chi'^\dagger \left[\frac{\mathbf{p}'+\mathbf{p}}{2m} \cdot \boldsymbol{\sigma}\, \sigma^i + \sigma^i\, \boldsymbol{\sigma} \cdot \frac{\mathbf{p}'+\mathbf{p}}{2m}\right] \chi \tag{15.23}$$

Using the identity $\sigma^a\sigma^b = \delta^{ab} + i\epsilon^{abc}\sigma^c$, this simplifies. The symmetric part (in spin) gives:

$$\bar{u}(p')\gamma^i u(p) \approx \frac{p_i' + p_i}{m} + \text{spin-dependent terms} \tag{15.24}$$

The leading velocity-dependent term is:

$$\mathcal{M}_{\text{mag}} \approx \frac{e^2}{\mathbf{q}^2} \cdot \frac{(\mathbf{p}_1' + \mathbf{p}_1)\cdot(\mathbf{p}_2' + \mathbf{p}_2)}{m^2} \tag{15.25}$$

At small momentum transfer, $\mathbf{p}_1' \approx \mathbf{p}_1$ and $\mathbf{p}_2' \approx \mathbf{p}_2$, so:

$$\mathcal{M}_{\text{mag}} \approx \frac{4e^2\, \mathbf{p}_1\cdot\mathbf{p}_2}{m^2\mathbf{q}^2} \tag{15.26}$$

The Fourier transform gives:

$$V_{\text{mag}}(r) = -\frac{\alpha}{m^2}\frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{r} \tag{15.27}$$

This is the magnetic interaction between moving charges! It's exactly what classical electromagnetism predicts: moving charges create magnetic fields that exert forces on other moving charges.

Note the minus sign—the magnetic force between parallel currents is attractive (they're going the same direction), opposite to the repulsive Coulomb force between like charges.

Spin-Dependent Forces

The spin-dependent terms we glossed over are equally important for atomic physics. They include:

Spin-orbit coupling: The electron's spin interacts with the magnetic field created by its orbital motion around the nucleus:

$$V_{\text{SO}} = \frac{\alpha}{2m^2 r^3}\mathbf{L} \cdot \mathbf{S} \tag{15.28}$$

This causes the fine structure splitting in atomic spectra.

Spin-spin (hyperfine) interaction: Two spins interact via:

$$V_{\text{SS}} = \frac{8\pi\alpha}{3m_1 m_2}\mathbf{S}_1 \cdot \mathbf{S}_2\, \delta^{(3)}(\mathbf{r}) + \frac{\alpha}{m_1 m_2 r^3}\left[3(\mathbf{S}_1 \cdot \hat{r})(\mathbf{S}_2 \cdot \hat{r}) - \mathbf{S}_1 \cdot \mathbf{S}_2\right] \tag{15.29}$$

The contact term (the delta function) causes hyperfine splitting in hydrogen.

Darwin term: There's also a contact interaction:

$$V_{\text{Darwin}} = \frac{\pi\alpha}{2m^2}\delta^{(3)}(\mathbf{r}) \tag{15.30}$$

This comes from the "zitterbewegung"—the rapid quantum jittering of the electron. It only affects *s*-states (which have nonzero probability at the origin).

The complete non-relativistic reduction of QED gives the *Breit Hamiltonian*, which includes all these terms:

$$\begin{aligned} H_{\text{Breit}} = {} & \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + \frac{\alpha}{r} \\ & - \frac{\alpha}{m_1 m_2}\left[\frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{r} + \frac{(\mathbf{p}_1 \cdot \mathbf{r})(\mathbf{p}_2 \cdot \mathbf{r})}{r^3}\right] \\ & + \text{spin-orbit} + \text{spin-spin} + \text{Darwin} + \mathcal{O}(v^4/c^4) \end{aligned} \tag{15.31}$$

Every term has a clear physical origin in QED. The Breit Hamiltonian is exact to order v^2/c^2—all of atomic physics at this precision follows from photon exchange.

15.9 *Classical Electromagnetic Waves*

What about electromagnetic waves? In QED, these are coherent states of many photons. Let's understand this connection more carefully.

A classical electromagnetic wave with frequency ω and electric field amplitude E_0 carries energy density $u = \frac{1}{2}\epsilon_0 E_0^2$ (in SI units). Each photon carries energy $\hbar\omega$. So the photon number density is:

$$n = \frac{u}{\hbar\omega} = \frac{\epsilon_0 E_0^2}{2\hbar\omega} \tag{15.32}$$

For visible light ($\omega \sim 10^{15}$ Hz) with $E_0 \sim 1$ V/m (a dim light), this gives $n \sim 10^8$ photons per cubic meter. A laser pointer has many orders of magnitude more. When $n \gg 1$, the fractional fluctuations in photon number scale as $1/\sqrt{n}$, becoming negligible for macroscopic fields.

The Propagator as Green's Function

The photon propagator we used, $D_{\mu\nu}(q) = -ig_{\mu\nu}/q^2$, is directly related to the Green's function for Maxwell's equations.

In Lorenz gauge ($\partial_\mu A^\mu = 0$), Maxwell's equations reduce to:

$$\partial^2 A_\mu = -j_\mu \tag{15.33}$$

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

The formal solution is:

$$A_\mu(x) = \int d^4y\, G(x-y) j_\mu(y) \tag{15.34}$$

where $G(x-y)$ is the Green's function satisfying $\partial^2 G(x) = -\delta^{(4)}(x)$.

In momentum space, this is trivial:

$$q^2 \tilde{G}(q) = 1 \implies \tilde{G}(q) = \frac{1}{q^2} \tag{15.35}$$

This is exactly the photon propagator (up to the tensor structure and factors of i from our conventions). The propagator *is* the Green's function—it describes how electromagnetic disturbances propagate.

Retarded Propagation

What about causality? The Feynman propagator $1/q^2$ isn't obviously causal—it seems to treat past and future symmetrically.

In position space, the Feynman propagator is:

$$D_F(x) = \int \frac{d^4q}{(2\pi)^4} \frac{e^{-iq\cdot x}}{q^2 + i\epsilon} \tag{15.36}$$

The $i\epsilon$ prescription determines how we handle the poles at $q^0 = \pm|\mathbf{q}|$. Evaluating by contour integration:

- For $x^0 > 0$: close the contour in the lower half-plane, picking up the pole at $q^0 = +|\mathbf{q}| - i\epsilon$
- For $x^0 < 0$: close in the upper half-plane, picking up $q^0 = -|\mathbf{q}| + i\epsilon$

The result involves both positive and negative frequency components. For scattering processes, this is what we want: particles can propagate both forward and backward in time (the latter interpreted as antiparticles going forward).

For classical electromagnetism, we want the *retarded* Green's function, which only propagates forward in time:

$$G_{\text{ret}}(x) = \frac{\theta(x^0)}{4\pi|\mathbf{x}|}\delta\left(x^0 - |\mathbf{x}|\right) \tag{15.37}$$

This says: a disturbance at the origin at $t = 0$ reaches point $\mathbf{x}$ at time $t = |\mathbf{x}|/c$ (with $c = 1$)—light-cone propagation. The $\theta(x^0)$ enforces causality.

The Feynman propagator and retarded propagator differ by terms that contribute to vacuum fluctuations but not to classical field configurations. In the classical limit (many photons), these differences average out.

15.10 Why $1/r^2$ Forces?

The Coulomb force falls off as $1/r^2$. In QED, this comes from the photon being massless. This connection between particle mass and force range is deep and general.

The Yukawa Potential

Suppose the photon had a mass M. The propagator would become:

$$\frac{-ig_{\mu\nu}}{q^2 - M^2 + i\epsilon} \tag{15.38}$$

In the non-relativistic limit, $q^0 \approx 0$, so $q^2 \approx -\mathbf{q}^2$:

$$\tilde{V}(\mathbf{q}) \propto \frac{1}{\mathbf{q}^2 + M^2} \tag{15.39}$$

The Fourier transform gives the potential in position space:

$$V(r) = \int \frac{d^3q}{(2\pi)^3} \frac{e^{-i\mathbf{q}\cdot\mathbf{r}}}{\mathbf{q}^2 + M^2} \tag{15.40}$$

Let's evaluate this. In spherical coordinates:

$$V(r) = \frac{1}{(2\pi)^3} \int_0^\infty q^2\, dq \int_{-1}^{1} d(\cos\theta) \int_0^{2\pi} d\phi\, \frac{e^{-iqr\cos\theta}}{q^2 + M^2} \tag{15.41}$$

The angular integrals give:

$$V(r) = \frac{1}{2\pi^2 r}\int_0^\infty \frac{q\sin(qr)}{q^2+M^2}\,dq \tag{15.42}$$

This integral can be done by contour methods. Write $\sin(qr) = \mathrm{Im}(e^{iqr})$ and extend to the complex plane. The integrand has poles at $q = \pm iM$. Closing the contour in the upper half-plane (since $r > 0$ makes e^{iqr} decay there), we pick up the pole at $q = iM$:

$$\int_0^\infty \frac{q\sin(qr)}{q^2+M^2}\,dq = \frac{\pi}{2}e^{-Mr} \tag{15.43}$$

Therefore:

$$\boxed{V(r) = \frac{e^{-Mr}}{4\pi r}} \tag{15.44}$$

This is the *Yukawa potential*, discovered in 1935 when Yukawa proposed that the nuclear force is mediated by a massive particle (now called the pion).

Interpretation

The Yukawa potential has two factors:

- $1/r$: the familiar Coulomb-like behavior from three spatial dimensions
- e^{-Mr}: exponential screening from the particle mass

At distances $r \ll 1/M$, the exponential is close to 1, and we recover the $1/r$ potential. At $r \gg 1/M$, the exponential kills the interaction.

The characteristic length scale is:

$$\lambda = \frac{1}{M} = \frac{\hbar}{Mc} \tag{15.45}$$

This is the *Compton wavelength* of the exchanged particle. It sets the range of the force.

Examples

Electromagnetism: $M_\gamma = 0 \implies \lambda = \infty$. The Coulomb force has infinite range.

Weak force: $M_W \approx 80$ GeV $\implies \lambda_W = \hbar/(M_W c) \approx 2\times 10^{-18}$ m. This is about 1/1000 the size of a proton. The weak force is truly short-ranged.

Strong force (naive): Yukawa predicted $M_\pi \approx 100$ MeV to get $\lambda \approx 2$ fm (the range of nuclear forces). The pion was discovered in

1947 with $M_\pi \approx 140$ MeV, confirming his prediction. (The full story of the strong force is more complex—QCD and confinement—but Yukawa's picture captures the right physics at nuclear scales.)

Gravity: $M_{\text{graviton}} = 0$ (we think) $\implies$ infinite range, just like electromagnetism.

The inverse relationship between mass and range is a fundamental feature of quantum field theory. It's why we need particle accelerators to probe short distances: to create the heavy particles that mediate short-range forces, we need high energies.

What If the Photon Had a Small Mass?

Experimentally, the photon mass is constrained to be extremely small: $M_\gamma < 10^{-18}$ eV. This corresponds to a Compton wavelength larger than the observable universe!

If the photon had a tiny but nonzero mass, the Coulomb potential would become:

$$V(r) = \frac{\alpha}{r} e^{-M_\gamma r} \tag{15.46}$$

At distances $r \ll 1/M_\gamma$, this is indistinguishable from Coulomb. But at very large distances (galactic or cosmological scales), deviations would appear. The tight experimental bounds come from searches for such deviations in precision electromagnetism and cosmological observations.

15.11 Classical Fields from Coherent States

We've seen that forces arise from particle exchange, and that the propagator is the classical Green's function. But there's still a puzzle: how do we get a classical electromagnetic *field* from discrete photons?

The answer involves *coherent states*—quantum states that behave as classically as quantum mechanics allows.

The Quantum Electromagnetic Field

In QED, the photon field operator is:

$$A_\mu(x) = \sum_{\mathbf{k},\lambda} \frac{1}{\sqrt{2\omega_k V}} \left[\epsilon_\mu^{(\lambda)}(\mathbf{k}) a_{\mathbf{k},\lambda} e^{-ikx} + \text{h.c.}\right] \tag{15.47}$$

where $a_{\mathbf{k},\lambda}$ annihilates a photon of momentum $\mathbf{k}$ and polarization λ, and $\epsilon_\mu^{(\lambda)}$ is the polarization vector.

The operators satisfy the commutation relations:

$$[a_{\mathbf{k},\lambda}, a^\dagger_{\mathbf{k}',\lambda'}] = \delta_{\mathbf{k}\mathbf{k}'}\delta_{\lambda\lambda'} \tag{15.48}$$

In the vacuum state $|0\rangle$, the field has zero expectation value:

$$\langle 0|A_\mu(x)|0\rangle = 0 \tag{15.49}$$

But this doesn't mean the field is zero! The vacuum has fluctuations:

$$\langle 0|A_\mu(x)A_\nu(y)|0\rangle \neq 0 \tag{15.50}$$

These are the vacuum fluctuations that contribute to effects like the Lamb shift.

What Are Coherent States?

A coherent state is an eigenstate of the annihilation operator:

$$a|\alpha\rangle = \alpha|\alpha\rangle \tag{15.51}$$

where α is a complex number.

Coherent states can be written explicitly as:

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle = e^{\alpha a^\dagger - \alpha^* a}|0\rangle \tag{15.52}$$

where $|n\rangle$ is the state with exactly n photons.

Key properties of coherent states:

1. **Superposition of number states:** A coherent state is a superposition of states with 0, 1, 2, ... photons. The probabilities follow a Poisson distribution with mean $|\alpha|^2$.

2. **Nonzero field expectation:** Unlike number states, coherent states have a nonzero expectation value for the field:

$$\langle\alpha|A_\mu(x)|\alpha\rangle = A_\mu^{\text{classical}}(x) \tag{15.53}$$

   This *is* the classical field.

3. **Minimum uncertainty:** Coherent states saturate the Heisenberg uncertainty relation for the field quadratures. They're as "classical" as quantum mechanics allows.

Classical Limit

For a coherent state with amplitude α, the mean photon number is $\langle n\rangle = |\alpha|^2$ and the variance is also $|\alpha|^2$. The fractional fluctuation is:

$$\frac{\Delta n}{\langle n\rangle} = \frac{1}{\sqrt{|\alpha|^2}} = \frac{1}{\sqrt{\langle n\rangle}} \tag{15.54}$$

When $\langle n\rangle \gg 1$, the fractional fluctuations become negligible. A laser beam with 10^{15} photons has fractional fluctuations of 10^{-7}—completely negligible for practical purposes.

Similarly, the quantum fluctuations in the field itself become small compared to the classical amplitude:

$$\frac{\Delta A}{|\langle A \rangle|} \sim \frac{1}{\sqrt{\langle n \rangle}} \tag{15.55}$$

This is the classical limit: when many photons are present, their quantum nature averages out, and we see a smooth classical field.

Why Lasers Produce Classical Light

Lasers produce coherent states (approximately). This is why laser light behaves so classically: it has a well-defined phase and amplitude, with minimal quantum fluctuations.

In contrast, thermal light (from a hot filament) is in a thermal state—an incoherent mixture of number states. Thermal light has larger fluctuations and no well-defined phase. But when averaged over many modes and long times, it too approaches classical behavior.

The point is this: classical electromagnetism isn't wrong—it's the limit of QED when many photons are present. The discrete nature of photons only matters when photon numbers are small, as in single-photon detectors, quantum optics experiments, or atomic transitions.

15.12 *Summary*

QED reproduces classical electromagnetism in the appropriate limits:

- The Coulomb potential $V = \alpha/r$ comes from tree-level photon exchange in the non-relativistic limit.
- The $1/r$ dependence arises from the massless photon propagator $1/q^2$.
- Magnetic forces come from the velocity-dependent (γ^i) parts of the vertex.
- Classical electromagnetic waves are coherent states of many photons.
- The range of a force equals the inverse mass of the exchanged particle.

This is deeply satisfying. The elaborate machinery of quantum field theory—propagators, vertices, Feynman diagrams—isn't just an abstract formalism. It describes real physics. Photon exchange is how charged particles interact, and in the classical limit, this exchange manifests as the familiar electromagnetic force.

The "virtual particles" that we integrate over in loop diagrams aren't figments of mathematical imagination. They're the quantum mechanical origin of forces. When you feel the push of two magnets repelling, you're experiencing the collective effect of countless virtual photon exchanges.

QED is not a replacement for Maxwell's equations—it's their quantum completion. At macroscopic scales, QED reduces to classical electromagnetism. At atomic scales, quantum effects become important. And at very short distances (high energies), the quantum corrections we've computed throughout these lectures become dominant.

The success of this reduction—that the most fundamental theory we know reproduces the most successful classical theory we know—is powerful evidence that we're on the right track.

Bibliography

[1] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, Boulder, CO, 1995.

[2] Steven Weinberg. *The Quantum Theory of Fields*, volume 1. Cambridge University Press, 1995.

[3] Mark Srednicki. *Quantum Field Theory*. Cambridge University Press, 2007.

[4] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.

[5] A. Zee. *Quantum Field Theory in a Nutshell*. Princeton University Press, 2nd edition, 2010.

[6] Julian Schwinger. On quantum-electrodynamics and the magnetic moment of the electron. *Physical Review*, 73:416–417, 1948.

[7] Gerard 't Hooft and Martinus Veltman. Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44:189–213, 1972.

[8] Kenneth G. Wilson. The renormalization group and the ϵ expansion. *Physics Reports*, 12:75–199, 1974.

[9] Kenneth G. Wilson. *The Renormalization Group and Critical Phenomena*. 1982. Nobel Lecture.

[10] Freeman J. Dyson. The radiation theories of tomonaga, schwinger, and feynman. *Physical Review*, 75:486–502, 1949.

[11] Richard P. Feynman. Space-time approach to quantum electrodynamics. *Physical Review*, 76:769–789, 1949.

[12] Sin-Itiro Tomonaga. On a relativistically invariant formulation of the quantum theory of wave fields. *Progress of Theoretical Physics*, 1:27–42, 1946.

[13] Willis E. Lamb and Robert C. Retherford. Fine structure of the hydrogen atom by a microwave method. *Physical Review*, 72: 241–243, 1947.

[14] T. Aoyama et al. The anomalous magnetic moment of the muon in the standard model. *Physics Reports*, 887:1–166, 2020.

[15] Claude Itzykson and Jean-Bernard Zuber. *Quantum Field Theory*. McGraw-Hill, 1980.

[16] Pierre Ramond. *Field Theory: A Modern Primer*. Benjamin/Cummings, 1981.

[17] H. David Politzer. Reliable perturbative results for strong interactions? *Physical Review Letters*, 30:1346–1349, 1973.

[18] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Physical Review Letters*, 30:1343–1346, 1973.