

A slightly different take on renormalization

April 2023

Note: This article was originally written for publication in my university student journal, but since my professor's review has been pending for like 2 months now, I'll publish it here too.

Abstract

Renormalization is an important concept in quantum field theory, which is often described as subtracting infinities from each other to assign finite values to initially divergent Feynman diagrams. While this is certainly a formally correct way to describe the working principle of renormalization, it is of little use in understanding *why* we need renormalization, or the physical significance of counterterms, renormalized parameters or various cut-off methods.

In this article, I present an alternative interpretation of renormalization that puts a focus on conceptual clarity. The reader should have basic knowledge of standard perturbative quantum field theory, i.e. second quantization, the interaction picture and the perturbative expansion of the interaction Hamiltonian, Wick's theorem and Feynman rules in position and momentum space and a general idea on how to derive them.

1 The problem

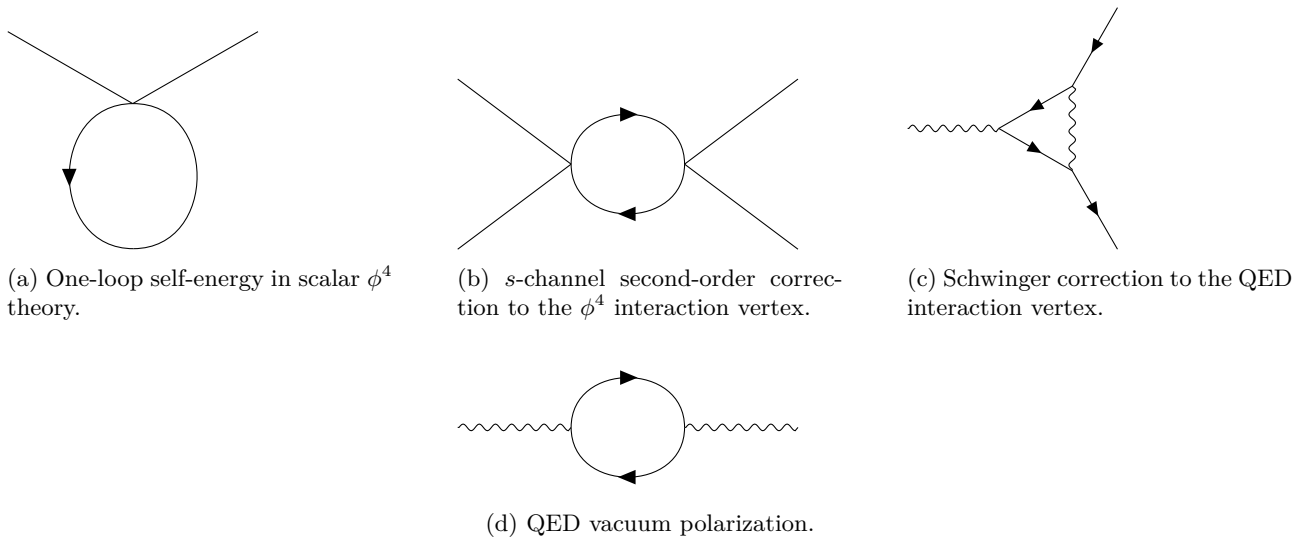


Figure 1: Various divergent Feynman diagrams.

In most interacting field theories, we will invariably encounter Feynman diagrams that are divergent because of an integral over a loop. Figure 1 shows some examples. The integral arises because the rule for momentum conservation doesn't fully fix all the internal momenta, so there are "leftover" integrals that don't get absorbed by the momentum conservation factors¹

$$\tau^4 \delta^{(4)} \left(\sum_i p_i^\mu \right).$$

A simple example is the first-order, one-loop self-energy of the phi particle in scalar ϕ^4 theory (Figure

¹In this article, we use $\tau = 2\pi$ for cleaner notation.

1a):

$$\begin{aligned}
1 + iM_{fi} &= \langle f | \hat{S} | i \rangle \\
&= \langle f | T [1 + (-i)\hat{H}_{\text{int}} + (-i)^2\hat{H}_{\text{int}}^2 + \dots] | i \rangle \\
&= 1 + (-i\lambda) \langle 0 | \hat{a}(p) \overbrace{\phi\phi\phi\phi}^{\overbrace{\quad\quad\quad\quad}^{\overbrace{\quad\quad\quad\quad}^{\overbrace{\quad\quad\quad\quad}}}} \hat{a}^\dagger(p) | 0 \rangle + \mathcal{O}(\lambda^2) \\
&= 1 - i\lambda \int \frac{d^4q}{\tau^4} \frac{i}{q^2 - m^2 + i\epsilon} + \mathcal{O}(\lambda^2)
\end{aligned}$$

We count four powers of momentum in the numerator (from d^4q), and two powers of momentum in the denominator (from q^2), resulting in a $q^4/q^2 = q^2$ divergence.

Let's take a step back and reconsider what we just did: We asked for the amplitude of the process of one phi going from momentum p to momentum p . In the interaction picture, the obvious possibility here is "nothing happens", represented by the 1 term in the expansion. The first-order process is the phi self-interaction depicted in Figure 1a. However, if we find that this term is infinite, we get a result that looks like $\langle f | \hat{S} | i \rangle = 1 - i\infty$, which certainly violates unitarity and hence doesn't make much sense.

Apparently, our ϕ^4 theory is inherently inconsistent. However, the other (non-infinite) terms of quantum field theories can make quite useful physical predictions - consider for instance quantum electrodynamics, which predicts hydrogen orbitals and transitions quite well if the divergent terms are just ignored. In the early years of QFT, physicists thought that they could get around this problem by simply dismissing the divergent terms as nonphysical. We'll shortly see that this is not the case at all.

2 Phonons: An analogous problem

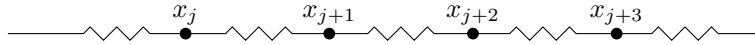


Figure 2: A chain of coupled harmonic oscillators.

To understand the meaning of these infinities, we'll consider a system of atoms on a crystal lattice. We know that on such scales, atoms behave as quantum harmonic oscillators, and we know that quantum harmonic oscillators are both well-defined and exactly solvable. To tackle the problem of renormalization, we'll see if we can construct a "real-life" version of scalar ϕ^4 theory using a crystal lattice and take a look at whether it exhibits the ill-defined divergent counterterms. Note that the following is not intended to be a rigorous treatment of phonons, but simply an illustration of some important concepts, so we will omit some details irrelevant to the illustration of the matter.

When considering a crystal lattice in the framework of quantum mechanics, we find that the classical approximation of spherical atoms bouncing against each other doesn't hold anymore - we need to consider the individual atoms as quantum harmonic oscillators coupled to each other. A typical interaction potential between two such harmonic oscillators on a crystal lattice is given by a potential composed of a $1/r^{12}$ repulsive term and an ionic $1/r$ attractive term:

$$V(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right) \right]$$

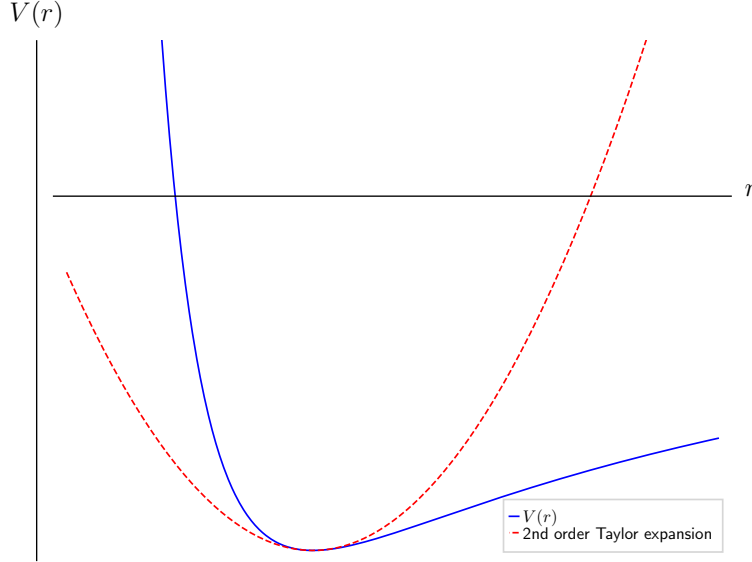


Figure 3: The lattice potential and its Taylor expansion at the equilibrium point.

We expand $V(r)$ around its equilibrium point r_e to second order, as seen in Figure 3. This is a good approximation for weak lattice vibrations and the base of the so-called Debye model of specific heat.

The Hamiltonian of a one-dimensional crystal lattice with N atoms, potential $V(r)$ and atom mass m is:

$$\begin{aligned}\hat{H} &= \sum_{j=1}^N \frac{\hat{p}_j^2}{2m} + \sum_{\langle jl \rangle} \hat{V}(|r_j - r_l|) \\ &= \sum_{j=1}^N \frac{\hat{p}_j^2}{2m} + \frac{1}{2} \hat{V}(|r_{j+1} - r_j|)\end{aligned}$$

where $\langle jl \rangle$ denotes the sum over nearest neighbors². We now define x_j as the position of the j th atom relative to its equilibrium position, i.e. $x_j = r_j - r_e$. Inserting the second-order Taylor expansion $V(x) = Kx^2$ then gives the Hamiltonian for N coupled harmonic oscillators:

$$\hat{H} = \sum_{j=1}^N \frac{\hat{p}_j^2}{2m} + \frac{1}{2} K(x_{j+1} - x_j)^2 \quad (1)$$

To solve this system, we decouple the system by taking the discrete Fourier transform of x_j :

$$\begin{aligned}\tilde{x}_k &= \sum_i x_i e^{\tau i \frac{kj}{N}} \\ x_j &= \frac{1}{N} \sum_k \tilde{x}_k e^{-\tau i \frac{kj}{N}}\end{aligned}$$

While $x_j, j = 1, \dots, N$ represent the displacement of the atom with equilibrium position jr_e , the new coordinates $\tilde{x}_k, k = -(N-1)/2, \dots, N/2$ represent the displacement of the vibrational mode with wavenumber $\tau k/Na$. Rewriting (1) in terms of the new coordinates gives³:

$$\hat{H} = \sum_k \left[\frac{1}{2m} \hat{p}_k \hat{p}_{-k} + \frac{1}{2} m \omega(k)^2 \hat{p}_k \hat{p}_{-k} \right],$$

or more simply in terms of creation and annihilation operators $\hat{a}(p), \hat{a}^\dagger(p)$ of the new coordinates:

$$\hat{H} = \sum_k \omega(k) \left(\hat{a}^\dagger(k) \hat{a}(k) + \frac{1}{2} \right), \quad (2)$$

which is the Hamiltonian for N uncoupled oscillators, where $\omega(k)$ denotes the *phonon dispersion relation*

$$\omega(k) = \sqrt{\frac{4K}{m}} \left| \sin\left(\frac{ka}{2}\right) \right|.$$

²For simplicity, we assume periodic boundary conditions, i.e. the last atom of the chain being linked back to the first atom so that the system becomes a “ring” of atoms.

³The details of the derivation are irrelevant for the purposes of this article; refer to [2] for details (and a clear and intuitive introduction to QFT in general).

2.1 The particle picture

To summarize, we've turned a system of N coupled harmonic oscillators representing individual atoms into a system of N uncoupled oscillators representing the possible wavelength modes of the chain. In the “particle picture”, we can say that an excitation of the wavelength mode k to the n th level corresponds to n phonon particles of wavelength k being present in the system. Identifying wavelength modes with particles in this way is the key element of second quantization. Again, [2] provides an excellent in-depth discussion of this notion. In the case of many-body physics, the particle-like excitations arising from large systems of other particles are known as *quasiparticles*.

Now, let's have a look at the details of the phononic dispersion relation.

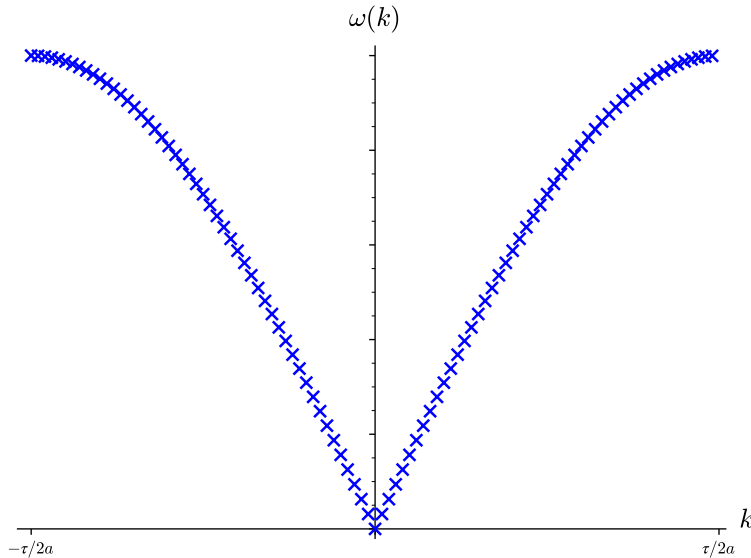


Figure 4: The phononic dispersion relation for $N = 100$.

In Figure 4, we can see that at low energies $\omega(k) \ll \sqrt{4K/m}$, the dispersion relation is approximately equal to the linear (massless) dispersion relation $\omega(k) = c_s|k|$. At high energies however, it exhibits non-linear behaviour.

There is a point crucial to our treatment of renormalization to emphasize that is often overlooked in introductory solid state physics: The dispersion relation function is not “periodic” after $\Delta k = \tau/a$ - it *ends* there, because there are no possible phonon modes beyond $-\tau/2a < k < \tau/2a$. This concept is known as the “first Brillouin zone” in solid state physics, but it is misleading in our formal treatment. Instead of a periodic function with multiple, identical Brillouin zones, the reader should view $\omega(k)$ as a function with a bounded domain. In the particle picture, this means that phonons with a wavenumber $k > \tau/2a$ or an energy $\omega > \sqrt{4K/m}$ simply do not exist.

2.2 Interaction between phonons

Figure 5: Interaction between phonons

To make the system a bit more interesting, we introduce a “toy model interaction” between phonons:

$$\mathcal{H}_{\text{int}} = \frac{\lambda}{4!} x^4 \quad (3)$$

Note that the actual interaction terms resulting from a higher-order expansion of the lattice potential would not look like this, but instead additionally carry terms containing the momenta of the interacting particles. This, however, would unnecessarily complicate matters here, so we'll leave those terms aside.

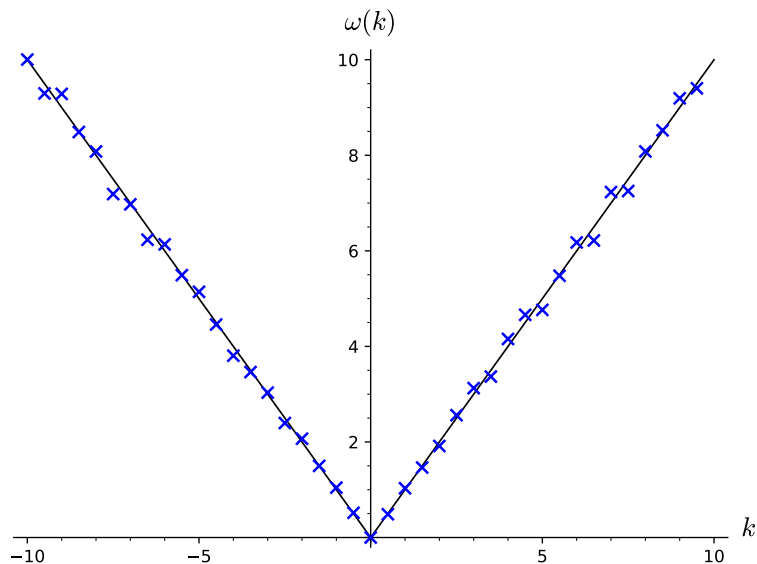


Figure 6: Phonon dispersion relation, as measured by a quasi-physicist

2.3 A tale of quasi-physicists

Phonons are by far not the only quasiparticles found in crystals. Some common examples for other quasiparticles are:

- magnons: collective spin excitations,
- polarons: interactions between electrons and ions,
- excitons: holes and electrons in a bound state
- etc.

All those quasiparticles can interact with each other in many different ways. Now, let’s imagine a crystal in which the interaction between different quasiparticles is so strong that they readily form all sorts of bound states, just like real elementary particles. The quasiparticles on our hypothetical crystal might bond together to quasi-nucleons, quasi-atoms, quasi-molecules, and so on, until we have a full “quasi-chemistry” happening on the crystal lattice.

When the right conditions are met, i.e. the crystal is large enough and there is a steadily available energy source (for instance, a laser in our laboratory exciting a specific spot of the crystal), it is conceivable that some form of “quasi-life” might evolve on our crystal. Of course, such lifeforms would be orders of magnitude larger in size than life as we know it because the typical bound state between quasiparticles will necessarily be larger than the distance between two atoms on the lattice, and hence much larger than the typical bound state between elementary particles (e.g. a proton).

If at some point intelligent life evolves on our crystal, “quasi-physicists” will start to investigate the nature of their universe. Gradually, they’ll discover their constituent quasi-molecules, quasi-atoms, and so on, until they reach the level of single quasiparticles, which would of course be labelled “elementary particles” by them. In their studies of these pseudo-elementary particles, the quasi-physicists would face challenges similar to the ones real-world physicists face - some quasiparticles might only appear in bound states or at very high energies, necessitating the construction of quasiparticle accelerators by our quasiparticle physicists. Above all, the quasiparticles will be very tiny compared to the quasi-physicists, just like real elementary particles are unimaginably tiny compared to human bodies.

2.3.1 Experiments by quasi-physicists

Now let’s say a quasi-physicist makes a first attempt to measure the dispersion relation of the phonon by accelerating phonons to a known momentum and then measuring their energy. The result of this experiment will look something like Figure (6). This is because the quasi-physicists and the quasi-measurement devices will be very large in comparison to the lattice length a , and hence the typical momentum will be a lot smaller than $\tau/2a$, so that we land in the nearly-linear regime of Figure (4).

The quasi-physicists however know nothing of atom lattice lengths or maximum phonon momentums. They can only observe the low-energy regime, and seeing that the linear fit perfectly describes the phonon dispersion relation up to experimental error, they might come up with the following Lagrangian to describe the phonon field they observe:

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

which is essentially the Lagrangian for massless ($m = 0$) scalar bosons.

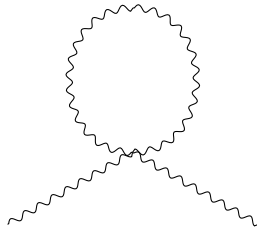


Figure 7: One-loop correction to the phonon propagator

We find that this low-energy phonon theory diverges just as well as our previous scalar boson theory: 7:

$$iM_{fi} = -i\lambda \int \frac{d^4k}{\tau^4} \frac{i}{k^2 - m^2 + i\epsilon} \quad (5)$$

Now, the quasi-physicists have an analogous renormalization problem. Similar to us, they are probably scratching their heads over how it is possible that a theory that predicts the behaviour of phonon propagation and interaction so nicely now suddenly yields infinite terms on a closer look. Should they just ignore the divergent terms, or come up with some entirely new theory?

2.3.2 The answer to phonon renormalization

The answer to the problem of our quasi-physicists lies in the physical meaning of the integral in (5). The process we are describing here actually is not just one, but many different ones - the virtual phonon propagating from and back to the interaction vertex can have any arbitrary four-momentum k , so in order to calculate the total amplitude of the one-loop correction, we need to integrate over all possible four-momenta k the virtual phonon could have.

At this point, the *real* physicists in the lab with the crystal might observe the quasi-physicists discussing this apparent inconsistency of their low-energy phonon theory. Contrary to them however, we know exactly what is wrong with the theory (4) - it glosses over the high-energy behaviour of phonons! At very high momenta, the phonons begin to exhibit a sinusoidal dispersion relation instead of a linear one, and more crucially, simply *don't exist* beyond some $p \geq \Lambda$. We can smile at the futile efforts of the phonon physicists to make sense of their theory and immediately see that the integral in (5) should not go up to infinite four-momenta, but just to $|p| < \tau/2a$. This obviously makes the integral non-divergent, and the renormalization problem of the phonon is solved. If we were to establish contact with the quasi-physicists, we could tell them about this high-momentum cutoff, which would allow them to write an “ $p^2 < \Lambda^2$, $\Lambda =$ (the value we gave them)” next to all of their integrals. If their technology is roughly analogous to ours, this will be an unattainably high energy value for them and hence hard to verify, but it makes the low-energy phonon theory self-consistent, which in turn allows the quasi-physicists to calculate cross-sections and fit the data from their quasiparticle accelerators to their theory.

2.3.3 Back to the real world

We have seen that while the phonon theory is perfectly divergence-free, its low-energy limit can appear to contain divergent terms if we fail to consider the high-energy behaviour of the theory. Going back to normal ϕ^4 theory, we can start to see why the integral for the one-loop self-energy of the phi particle

$$i\Pi(p) = -i\lambda \int \frac{d^4q}{\tau^4} \frac{i}{q^2 - m^2 + i\epsilon} \quad (6)$$

diverges. We have integrated over all possible momenta, from zero to infinity, implicitly assuming that our theory is valid up to arbitrary energy scales. Doing this results in divergent terms, rendering our theory useless. Therefore, in a classical *reductio ad absurdum*, we can conclude that in order for our ϕ^4 theory to have any chance to describe physical reality at all, we need to assume that the theory stops being valid at some point and some other, higher-level theory takes over⁴. In a way, the theory *predicts its own failure* at high energy regimes. For instance, it is a common view among physicists that spacetime might be “coarse-grained” at very small scales - similar to how a lattice of atoms can be approximated as a continuous fields at large scales/low momenta/low energies, but becomes coarse-grained at small scales/high momenta/high energies.

Unfortunately, we don't have any experimental evidence of such a “maximum momentum scale” or signs of physicists from a higher plane of reality establishing contact with us so far, so we're forced to guess. A group of particle physicists at CERN might assume this cutoff to lie somewhere on the order of the Planck scale and proceed to fit the cross-sections at LHC to the theory, obtaining values for the masses and coupling constants. Then again, some theoretical physicists from Heidelberg might speculate that the

⁴There are several other ways to coerce a theory into behaving and yielding useful predictions, most notably *dimensional regularization*, which assumes that the theory has a dimension slightly lower than 4 to make the integrals finite. The working principle outlined here is essentially the same.

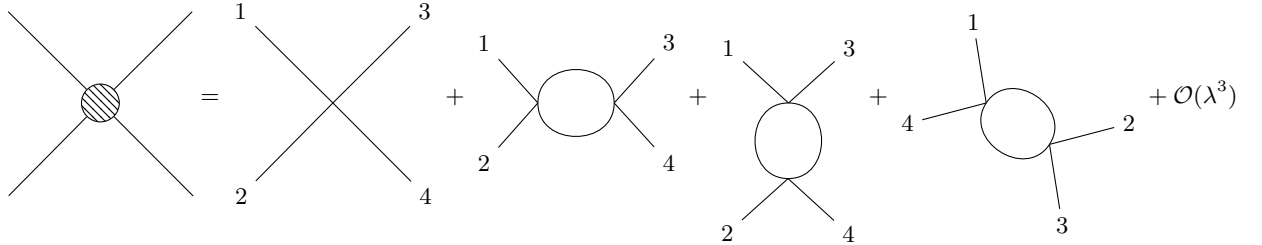
cutoff is somewhere on the order of the **Oh-My-God particle** and obtain different values for the masses and coupling constants. This would obviously cause friction in the physics community, so we need a way to formulate a theory with a high-momentum cutoff while leaving the value of this cutoff open and still obtaining predictions from the theory. This is the basic concept of renormalization.

3 The systematics of renormalization

Now that we have established the theoretical need for renormalization, let's try to introduce a high-momentum cutoff to our theory while maintaining deliberate ambiguity about its value.

3.1 Vertex renormalization

We start by considering the effects of the formerly divergent vertex function terms on the coupling strength of the ϕ^4 vertex. Up to second order, the bare vertex and the s -, t - and u -channel one-loop diagrams contribute:



$$-i\lambda_P(s, t, u) = -i\lambda + i\Gamma_s(s) + i\Gamma_t(t) + i\Gamma_u(u) + \mathcal{O}(\lambda^3)$$

where s, t, u are the Mandelstam variables, λ_P is the physical (as in: measured in a particle accelerator by cross-sections) coupling, and

$$i\Gamma_s = (-i\lambda)^2 \int \frac{d^4q}{\tau^4} \frac{i}{q^2 - m^2} \frac{i}{(p_1 + p_2 - q)^2 - m^2}$$

$$i\Gamma_t = (-i\lambda)^2 \int \frac{d^4q}{\tau^4} \frac{i}{q^2 - m^2} \frac{i}{(p_1 - p_3 - q)^2 - m^2}$$

$$i\Gamma_u = (-i\lambda)^2 \int \frac{d^4q}{\tau^4} \frac{i}{q^2 - m^2} \frac{i}{(p_1 - p_4 - q)^2 - m^2}$$

are the vertex function contributions. These terms are divergent, so we postulate a high-momentum cutoff to make them finite:

$$i\Gamma_{s,t,u} \stackrel{q^2 < \Lambda^2}{\rightleftharpoons} i\Gamma_{s,t,u}(\Lambda)$$

For additional clarity, we write $i\Gamma_s + i\Gamma_t + i\Gamma_u = (-i\lambda)^2 A(s, t, u)$.

In the end, we obtain a prediction for the full vertex function (and hence for the matrix element M_{fi} and the cross-section $\propto |M_{fi}|^2$) that is dependent on the value of the bare coupling λ and the cutoff:

$$-i\lambda_P(s, t, u, \lambda, \Lambda) = -i\lambda + (-i\lambda)^2 A(s, t, u, \Lambda) + \mathcal{O}(\lambda^3) \quad (7)$$

This notations should raise an eyebrow, as the physical coupling is a real quantity we measure in experiments - if we scatter two phis with some Mandelstam variables s, t, u , the result of the experiment will certainly only depend on the actual physical parameters of the experiment! So (7) should actually look like

$$-i\lambda_P(s, t, u) = -i\lambda + (-i\lambda)^2 A(s, t, u, \Lambda) + \mathcal{O}(\lambda^3) \quad (8)$$

The exact values of s, t, u at which we perform the scattering experiment are called the *renormalization scale*.

To actually make the theory fit the experiment, we need to adjust λ based on the cut-off Λ we guessed. Imposing the condition that the theory with the parameter $\lambda(\Lambda)$ matches our observations, we write:

$$-i\lambda_P(s, t, u) \stackrel{!}{=} -i\lambda(\Lambda) + (-i\lambda(\Lambda))^2 A(s, t, u, \Lambda) + \mathcal{O}(\lambda(\Lambda)^3)$$

We exploit that we only work up to second order in λ :

$$\begin{aligned} & (-i\lambda_P(s, t, u))^2 = (-i\lambda(\Lambda))^2 + \mathcal{O}(\lambda(\Lambda)^3) \\ \Leftrightarrow & \quad (-i\lambda_P(s, t, u))^2 A(s, t, u, \Lambda) = (-i\lambda(\Lambda))^2 A(s, t, u, \Lambda) \\ \Leftrightarrow & \quad -i\lambda(\Lambda) = -i\lambda_P(s, t, u) + (-i\lambda_P(s, t, u))^2 A(s, t, u, \Lambda) \\ \Leftrightarrow & \quad \lambda(\Lambda) = \lambda_P(s, t, u) + -i\lambda_P(s, t, u)^2 A(s, t, u, \Lambda) \end{aligned}$$

We see that we can guess any arbitrary cutoff Λ and adjust $\lambda \rightarrow \lambda(\Lambda)$ such that the predictions of the theory remain the same regardless of our guess - up to the relevant second-order diagrams. If it turns out that the final predictions of the theory are indeed independent of our guess for Λ *to all orders* in λ , the theory is said to be *renormalizable*. In this case, the particle physics working group at CERN could guess some Λ_1 , determine the corresponding $\lambda(\Lambda_1)$, and would arrive at the exact same physical predictions as the group from Heidelberg who guessed Λ_2 , determined $\lambda(\Lambda_2)$ and used that coupling constant in their further calculations.

Now, we can easily give an expression for the physical coupling at any arbitrary Mandelstam variables $-i\lambda_P(s', t', u')$:

$$-i\lambda_P(s', t', u') = -i\lambda(\Lambda) + (-i\lambda(\Lambda))^2 A_{s,t,u}(s', t', u', \Lambda) + \mathcal{O}(\lambda(\Lambda)^3)$$

Inserting $\lambda(\Lambda) = \lambda_P(s, t, u) + (-i\lambda_P(s, t, u))^2 A(s, t, u, \Lambda)$:

$$\begin{aligned} -i\lambda_P(s', t', u') &= -i[\lambda_P(s, t, u) + (-i\lambda_P(s, t, u))^2 A(s, t, u, \Lambda)] \\ &\quad - i[\lambda_P(s, t, u) + (-i\lambda_P(s, t, u))^2 A(s, t, u, \Lambda)]^2 A_{s,t,u}(s', t', u', \Lambda) + \mathcal{O}(\lambda(\Lambda)^3) \\ &= -i\lambda_P(s, t, u) - i\lambda_P(s, t, u)^2 [A(s', t', u', \Lambda) - A(s, t, u, \Lambda)] + \mathcal{O}(\lambda(\Lambda)^3) \end{aligned}$$

In our case, the value of the integral $A(s, t, u, \Lambda)$ is:

$$A(s, t, u, \Lambda) = C \left[\log \frac{\Lambda^2}{s} + \log \frac{\Lambda^2}{t} + \log \frac{\Lambda^2}{u} \right]$$

where C is some constant that doesn't concern us. Our prediction for λ_P at s', t', u' based on our initial measurement of λ_P at s, t, u hence is:

$$-i\lambda_P(s', t', u') = -i\lambda_P(s, t, u) - i\lambda_P(s, t, u)^2 C \left[\log \frac{s}{s'} + \log \frac{t}{t'} + \log \frac{u}{u'} \right] + \mathcal{O}(\lambda(\Lambda)^3)$$

which is even better, because we now have an expression for the prediction of our theory that is *manifestly* independent of the guessed value of the high-energy cutoff.

3.2 Mass renormalization

Revisiting the phi one-loop self-energy (6), we again guess a high-energy cutoff, hence making the integral finite:

$$i\Pi(p, \lambda, \Lambda) = -i\lambda \int_{q^2 < \Lambda^2} \frac{d^4 q}{\tau^4} \frac{i}{q^2 - m^2 + i\varepsilon} \quad (9)$$

where p is the momentum of the particle, q the internal momentum of the loop, λ the ϕ^4 coupling constant and Λ the high-energy cutoff. We note that the expression is independent of the external momentum p , however we keep p as a parameter of Π in order to not lose generality.

As we have already determined that λ can be determined as a function of Λ based on scattering experiments, we write $i\Pi(p, \lambda, \Lambda) = i\Pi(p, \lambda(\Lambda), \Lambda) = i\Pi(p, \Lambda)$.

The full propagator is now given by the geometric series

$$\begin{aligned} \frac{i}{p^2 - m_P(p, \Lambda)^2} &= + \frac{i}{p^2 - m^2} \\ &\quad + \frac{i}{p^2 - m^2} i\Pi(p, \Lambda) \frac{i}{p^2 - m^2} \\ &\quad + \frac{i}{p^2 - m^2} i\Pi(p, \Lambda) \frac{i}{p^2 - m^2} i\Pi(p, \Lambda) \frac{i}{p^2 - m^2} \\ &\quad + \frac{i}{p^2 - m^2} i\Pi(p, \Lambda) \frac{i}{p^2 - m^2} i\Pi(p, \Lambda) \frac{i}{p^2 - m^2} i\Pi(p, \Lambda) \frac{i}{p^2 - m^2} \\ &\quad + \dots \end{aligned}$$

where $m_P(p, \Lambda)$ is the position of the pole of the full propagator, i.e. the physical mass of the particle we measure in particle accelerators. This can be solved via the geometric series formula:

$$\frac{i}{p^2 - m_P(p, \Lambda)^2} = \frac{i}{p^2 - m^2 + \Pi(p, \Lambda)}$$

obtaining

$$m_P(p, \Lambda)^2 = m^2 - \Pi(p, \Lambda).$$

We see that once we postulate a high-energy cutoff Λ , our theory makes a prediction for the physical mass of a particle. Just as the physical coupling constant in the last section however, this mass should only perhaps depend on the momentum p of the phi:

$$m_P(p)^2 = m^2 - \Pi(p, \Lambda).$$

Based on an experiment measuring the physical mass of the particle at some momentum (*renormalization scale*) p , we can again determine the bare mass parameter $m(\Lambda)$ as a function of our guess for the high-energy cutoff:

$$m(\Lambda)^2 = m_P(p)^2 - \Pi(p, \Lambda).$$

Again, the point of m depending on Λ is admitting that we do not know anything about the high-momentum cutoff - but *if* it lies at Λ , we can say for sure that the bare mass of the phi must be $m(\Lambda)$. Our renormalization scheme makes sure that the physical predictions stay the same.

3.3 Quasiparticle renormalization

The last important parameter of our theory to renormalize is the so-called *quasiparticle weight* Z . To understand what it does and why we need it, first consider a free field theory like the free scalar boson theory. The eigenstates of the Hamiltonian will be excitations of the individual wavelength modes, i.e. particles flying around. The lowest eigenvalue is that of the vacuum $|0\rangle$, and we can produce more eigenstates by acting on it with the creation operators of our theory. For instance, a state with one particle with momentum p would be

$$|p_{\text{free}}\rangle = a^\dagger(p) |0\rangle \quad (10)$$

which is an eigenstate of H_0 with energy ω_p . However, if we add in an interaction Hamiltonian, the resulting Hamiltonian will surely have different eigenstates. For instance, some eigenstates might be particles bound together by the new interaction, particles in hyperbolic trajectories to each other, and so on. One of them will probably be our dressed particle state $|p\rangle$ we are looking for - in QED, for instance, it might be an electron surrounded by a cloud of photons, positrons and other electrons. Expressing this as a formula, we can say that our dressed particle eigenstate $|p\rangle$ of the full Hamiltonian is a sum of the original eigenstates and some unrelated states $|\lambda\rangle$, like for instance a photon flying by the electron:

$$|p\rangle = \alpha |p_{\text{free}}\rangle + \sum_{\lambda} \beta_{\lambda} |\lambda\rangle$$

Turning this around, we can also write the original free state as a sum of the dressed particle state $|p\rangle$ we are looking for, and some other interacting eigenstates $|\gamma\rangle$:

$$|p_{\text{free}}\rangle = a^\dagger(p) |0\rangle = A |p\rangle + \sum_{\gamma} B_{\gamma} |\gamma\rangle$$

This equation tells us that if we act on the interacting vacuum with our original creation operator, it will not just create the single-particle state we are looking for - it will also create some other, unrelated states with amplitudes B_{γ} we are not interested in. The actual single-particle state is merely created with amplitude A . Rewriting (10) with the more familiar $Z = A^2$, we get

$$\sqrt{Z} |p\rangle = a^\dagger(p) |0\rangle + \dots, \quad (11)$$

ignoring the other states for now. An obvious interpretation of this relation is that $a^\dagger(p)$ creates the dressed particle state in question with probability Z , and any other state, like a bound or multiparticle state, with probability $1 - Z$ in total. Therefore, Z is also called the *quasiparticle weight*. The reason why we define it as $Z = A^2$ (and not, for instance, as $Z = A$) is that it makes the full propagator come out nicely:

$$\langle \phi \phi \rangle = \frac{iZ}{p^2 - m^2} \quad (12)$$

In analogy with the vertex and mass renormalization, we can absorb this factor into a new “physical field” operator:

$$\phi_P(x) = \frac{1}{\sqrt{Z}} \phi(x) \quad (13)$$

(12) then becomes:

$$\langle \phi_P \phi_P \rangle = \frac{i}{p^2 - m^2} \quad (14)$$

This procedure is called *field strength renormalization*. Again, note that (14) is what we actually measure in particle accelerators - it is physically impossible to create bare particle states, because we live in a world full of interactions. Now, without doing explicit calculations, we can reasonably assume that there are some loop terms involved in the quasiparticle propagator, so actually, we need to write

$$Z \longrightarrow Z(\Lambda)$$

and, based on the bare field operators $\phi(x)$, the theory would start to make a prediction for $\phi_P(x, \Lambda)$ and the full propagator $D_{\text{full}}(y - x, \Lambda)$. However, we already know where this is supposed to go - we can measure the full propagator and the physical field strength in experiments! So we pin D_{full} /the physical field strength ϕ_P to our experimental observations and make the bare field ϕ cutoff-dependent. Again, this is like saying that we have no idea what Λ is - but *if* it is some Λ_1 , the bare field has to be $\sqrt{Z(\Lambda_1)}$ times the field we measure:

$$\phi(x, \Lambda_1) = \sqrt{Z(\Lambda_1)} \phi_P(x) \quad (15)$$

3.4 Counterterms

So far, we have structured the renormalization procedure from the point of view of an experimentalist - we measure some physical parameters ϕ_P, m_P, λ_P , and then try to explain it with a theory involving bare parameters ϕ, m, λ . We notice that the predictions of the theory for the physical parameters are dependent of the cutoff Λ we guess, so we make the bare parameters Λ -dependent for the theory to match the observations:

$$\begin{aligned}\phi(x) &\longrightarrow \phi(x, \Lambda) \\ m &\longrightarrow m(\Lambda) \\ \lambda &\longrightarrow \lambda(\Lambda)\end{aligned}$$

This scheme emphasises conceptual clarity, but can be cumbersome to work with in practice. There is an equivalent scheme involving *counterterms*, which essentially does exactly the same, but just with slightly different maths. We take inspiration from (15) and absorb the Λ -dependence in multiplicative factors to the physical parameters:

$$\begin{aligned}\phi(\Lambda) &= Z_\phi^{1/2}(\Lambda)\phi_P \\ m(\Lambda) &= Z_m^{1/2}(\Lambda)m_P \\ \lambda(\Lambda) &= Z_\lambda(\Lambda)\lambda_P\end{aligned}$$

where $Z_\phi = Z$ from the previous section. We can rewrite the bare ϕ^4 Lagrangian in terms of these renormalized parameters:

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\partial_\mu\phi(\Lambda))^2 - \frac{1}{2}m(\Lambda)^2\phi^2(\Lambda) - \frac{\lambda(\Lambda)}{4!}\phi(\Lambda)^4 \\ &= \frac{1}{2}Z_\phi(\Lambda)(\partial_\mu\phi_P)^2 - \frac{1}{2}Z_m^2(\Lambda)Z_\phi(\Lambda)m_P^2\phi_P^2 - Z_\lambda(\Lambda)Z_\phi(\Lambda)^2\frac{\lambda_P}{4!}\phi_P^4\end{aligned}$$

Rearranging this yields:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi_P)^2 - \frac{1}{2}m_P^2\phi_P^2 - \frac{\lambda_P}{4!}\phi_P^4 \quad (16)$$

$$- \frac{1}{2}(1 - Z_\phi(\Lambda))(\partial_\mu\phi_P)^2 + \frac{1}{2}(1 - Z_\phi(\Lambda)Z_m(\Lambda))m_P^2\phi_P^2 + (1 - Z_\lambda(\Lambda)Z_\phi(\Lambda)^2)\frac{\lambda_P}{4!}\phi_P^4 \quad (17)$$

We notice that the first part of the Lagrangian looks like a perfectly normal ϕ^4 field theory - but this time with the *physical* parameters straightaway! The inherent Λ -dependence of the theory can now be absorbed into the interaction part of the Hamiltonian, i.e.

$$\begin{aligned}\mathcal{H}_{\text{int}} &= \frac{\lambda_P}{4!}\phi_P^4 \\ &+ \frac{1}{2}(1 - Z_\phi(\Lambda))(\partial_\mu\phi_P)^2 - \frac{1}{2}(1 - Z_\phi(\Lambda)Z_m(\Lambda))m_P^2\phi_P^2 - (1 - Z_\lambda(\Lambda)Z_\phi(\Lambda)^2)\frac{\lambda_P}{4!}\phi_P^4,\end{aligned}$$

so we can just treat it as a set of additional Feynman rules:

$$\begin{aligned}\text{---} \times \text{---} &= -i\left(\frac{1}{2}(1 - Z_\phi(\Lambda))p^2 + \frac{1}{2}(1 - Z_\phi(\Lambda)Z_m(\Lambda))m_P^2\right) \\ \begin{array}{c} | \\ \circ \\ / \quad \backslash \end{array} &= -i(1 - Z_\lambda(\Lambda)Z_\phi(\Lambda)^2)\frac{\lambda_P}{4!}\end{aligned}$$

These counterterms need to be inserted whenever an infinity arising from naively integrating over infinite momenta needs to be cancelled. They are often more convenient in practice, because we can just pick the first part of the re-scaled Lagrangian (16) when doing “normal” computations, and concern ourselves with the counterterms (17) when we want to calculate loop corrections.

4 Conclusion

Renormalization is often seen as some sort of black magic or wacky trick to cover up inconsistent theories. In this article, we’ve seen that this is the case - our theory is in fact inconsistent, but rightfully so. Renormalization is just a very clever way of simultaneously admitting that we have no idea where our theories start to fail, but still obtaining useful predictions from them.

In this article, we’ve only superficially covered the practical mathematical tools of renormalization. For a more thorough introduction to the mathematical underpinnings of what we’ve discussed, the reader is referred to e.g. [1].

References

- [1] Bertrand Delamotte. “A hint of renormalization”. In: *American Journal of Physics* 72.2 (2004), pp. 170–184. DOI: [10.1119/1.1624112](https://doi.org/10.1119/1.1624112). URL: <https://doi.org/10.1119%2F1.1624112>.
- [2] Tom Lancaster and Stephen J. Blundell. “Quantum Field Theory for the Gifted Amateur”. In: Oxford University Press, Apr. 2014, pp. 25–27. ISBN: 9780199699322. DOI: [10.1093/acprof:oso/9780199699322.001.0001](https://doi.org/10.1093/acprof:oso/9780199699322.001.0001). URL: <https://doi.org/10.1093/acprof:oso/9780199699322.001.0001>.