Quantum Field Theory Extension

MSc Physics and Astronomy, Theoretical Physics track

Dr Juan Rojo

VU Amsterdam and Nikhef Theory Group

http://www.juanrojo.com/

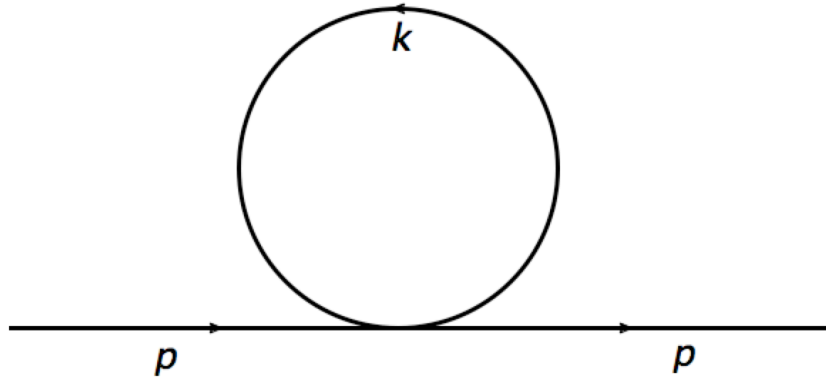j.rojo@vu.nl

Lecture notes Academic Year 2019-2020

Current version: **January 31, 2020**

Table of Contents

Figure 1.1. One of the loop diagrams that arise in calculations using $\lambda\phi^4$ theory beyond the Born approximation, with k^μ being the four-momentum running in the loop.

1   Renormalization in scalar QFTs

Beyond the Born approximation (the so-called *leading order* in the perturbative expansion), the calculation of scattering amplitudes in generic QFT involves the presence of *loops* of virtual particles. These particles will be in general different from those present in the initial and final states of the scattering. Indeed, basic considerations in quantum theory indicate that any particle with the right quantum numbers will circulate in virtual loops, even if the energies involved are much smaller than the actual mass of the particle.

A crucial aspect of these *virtual diagrams* is that in many cases they are *divergent*, that is, it is not possible to associate a finite result to them, which is obviously problematic. In this first part of the course, we will present how infinities arise in calculations in Quantum Field Theories beyond the Born level, and illustrate how they can be dealt with by a suitable redefinition of the parameters of the Lagrangian by means of the *renormalization process*.

Here we will focus on the $\lambda\phi^4$ scalar field theory, defined by the following Lagrangian:

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\,, \tag{1.1}$$

where λ is a dimensionless coupling constant and m is the mass of the (real) scalar field. The choice of $\lambda\phi^4$ theory to illustrate the renormalization paradigm is motivated by its simplicity, which yet still captures the most relevant physical aspects of the problem. In the second part of the course we will take a look at the implications of renormalisation for QED, the quantum theory of the electromagnetic interactions.

How do infinite results appear in QFT calculations beyond the Born approximation? Let us consider for example the diagram in Fig. 1.1, which arises in loop calculations of $\lambda\phi^4$ theory, in particular for the first non-trivial correction to the free-field scalar particle propagator. Using the Feynman rules of the theory, one can find that the matrix element associated to this specific amplitude $\mathcal{M}$ is given by

$$i\mathcal{M} = i\lambda \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 + m^2 - i\epsilon}\,, \tag{1.2}$$

where k^μ is the four-momentum running in the loop. It is easy to convince yourselves that this integral is

divergent. To show this, in the region where momenta are much larger than the mass, $k^\mu \gg m$, we have that

$$\mathcal{M} \sim \lambda \int d^4 k \frac{1}{k^2} \sim \int k \, dk \to \infty \,, \tag{1.3}$$

where we have used that in Euclidean space (which is more adequate to compute loop integrals) then $d^4 k = k^3 dk d\Omega_4$. So we find that the one-loop amplitude show in Fig. (1.1), is not only divergent, but also *quadratically* divergent, irrespective of the specific value of the coupling constant λ (that is, even for a vanishingly small λ). This kind of divergent integral appears in almost every attempt to compute matrix elements beyond leading order in perturbation theory. The problem was so severe that by the late 1930s the founding fathers of quantum mechanics such as Dirac, Bohr and Oppenheimer were about to give up QFT as a fundamental theory of Nature because of these divergent integrals.

How we can deal with these divergences and make physical sense out of the QFT predictions beyond the Born approximation? The crucial point here is that it is possible to reorganize the calculation in way that all divergences are absorbed into a redefinition of the (unobservable) parameters of the theory Lagrangian, in a way that all physical observables are finite. This program is generically known as *renormalization*, and will be the main topic of this first part of the course. But first of all, let us gain a bit more of physical insight about the origin of these infinities by studying the famous *Casimir effect*.

Before doing that, let me mention that within the modern understanding of QFTs, then infinities never actually arise. Indeed, QFTs are regarded as *effective theories*, which are only valid in a restricted region of energies, for example $E \lesssim \Lambda$. In this case the above loop integral reads

$$\mathcal{M} \sim \lambda \int^\Lambda d^4 k \frac{1}{k^2} \propto \int^\Lambda k \, dk \propto \Lambda^2 \,. \tag{1.4}$$

On the other hand, even with this understanding such result is quite problematic: it implies that the separation of scales in which the effective field theory is constructed is not very good, since results are extremely sensitive of the specific cutoff (namely on the physics of the high-energy scale $E \simeq \Lambda$). Well-defined effective field theories depend on the high-energy cutoff in a much milder way, at most logarithmically.

1.1 The Casimir effect

Let us illustrate how infinities can appear in QFT calculations with a simple example, namely the zero-point energy of a free (non-interacting) scalar field ϕ. The Hamiltonian for this scalar field in terms of the creation and annihilation operators $a_k^\dagger$ and a_k can be written as

$$H = \frac{1}{2} \int \frac{d^3 k}{(2\pi)^3} \left(a_k^\dagger a_k + \omega \left(\vec{k} \right) V \right) \,, \tag{1.5}$$

where V is the volume where the scalar field is confined, and by virtue of the dispersion relation of the free-field theory we have

$$\omega \left(\vec{k} \right) = \sqrt{\vec{k}^2 + m^2} \,, \tag{1.6}$$

with m being the particle mass. We can use Eq. (1.5) to compute the *vacuum energy*, defined as the expectation value of the Hamiltonian in the ground state. If we carry out this calculation we find:

$$E_0 \equiv \langle 0|H|0\rangle = V \int \frac{d^3k}{(2\pi)^3} \frac{1}{2}\omega(k) \sim V \int \frac{d^3k}{(2\pi)^3} \frac{k}{2} \sim V \int k^3 dk \to \infty \,, \tag{1.7}$$

so even a simple non-interacting scalar QFT leads to an infinite zero-point energy E_0. Moreover, this is a very severe divergence since it scales as the fourth power of the momentum k.

However, the fact that we have a divergent result is *per se* not a problem, since zero-point energies are not physical observables, and only differences in energy can be accessed experimentally. So one might think that this infinite vacuum energy is some kind of theoretical oddity with no implications for real physical systems. Still, this seems to be a rather remarkable effect, so maybe this infinite vacuum energy can have some observable effects after all?

To verify that indeed this vacuum energy leads to observable effects, let us consider the following setup: a box of size L, where inside we place a moving plate at position l. We impose Dirichlet boundary conditions for the scalar field both at the sides of the box and at the position of the plate $x = l$: the scalar field should vanish both for $x = 0$ and $x = L$ as well as for $x = l$, namely

$$\phi(x = 0) = \phi(x = L) = \phi(x = l) = 0\,. \tag{1.8}$$

In other words, we assume that we have a perfectly reflective plate as well as perfectly reflective walls of the box. The setup is illustrated in Fig. 1.2.

For this system, the total energy of the vacuum state in this specific configuration of quantum fields will then be the sum of the contributions from two regions in the box, namely

$$E_0(l) = E_0(x \le l) + E_0(x \ge (L - l))\,. \tag{1.9}$$

Now, as one moves the plate inside the box, that is, as one changes the value l, from general considerations there should appear a force exerted on the plate given by the variation of the energy with respect to l,

$$F_c = -\frac{dE_0(l)}{dl}\,, \tag{1.10}$$

pushing the system to a minimum of the energy. This force is the so-called *Casimir force*, predicted by the Dutch physicist Hendrik Casimir in 1948 and observed experimentally 10 years later. The Casimir force is therefore a real physical effect, and a direct consequence of the (infinite) zero-point energy of a quantum field theory. Note that the fact that the vacuum energy is infinite is not necessary a problem: any offset in E_0 will be canceled out when we take the derivative: only variations of E_0 with respect to l are physically meaningful, while the actual value of E_0 is not.

Let us now compute this Casimir force, where for the sake of simplicity we will restrict ourselves to a one dimensional box. In this case, in the expression for the vacuum energy Eq. (1.7) one replaces the integral over k by a finite sum over the field modes ω_n that satisfy the Dirichlet boundary conditions, namely

$$E_0(a) = \langle 0|H|0\rangle = \frac{1}{2}\sum_{n=1}^{\infty} \omega_n = \frac{1}{2}\sum_{n=1}^{\infty} \frac{\pi n}{a}\,, \tag{1.11}$$
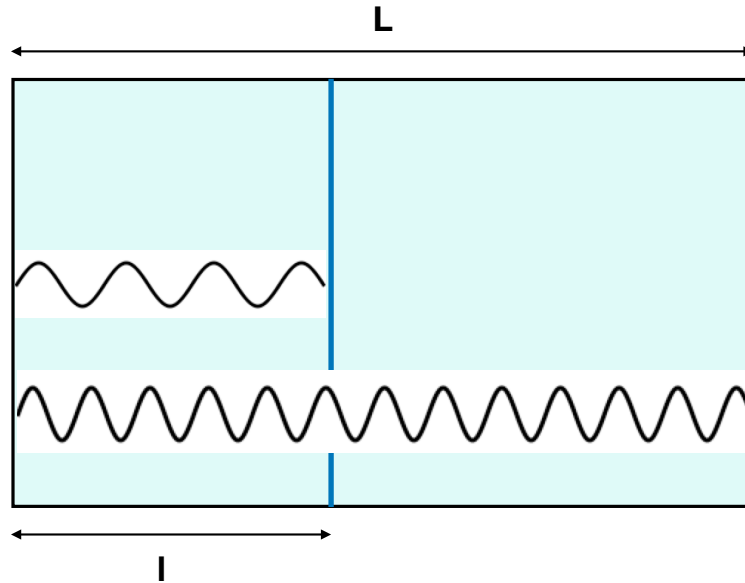
Figure 1.2. Setup for the Casimir effect. A box of length L contains a perfectly reflective plate at a distance l from one side of the box. We impose Dirichlet boundary conditions both at the sides of the box, $\phi(x = 0) = \phi(x = L) = 0$ and at the plate $\phi(x = l)$.

where a is the length of the region of the box that contributes to E_0, and we have used that $\omega_n = \pi n/a$. Therefore, the total vacuum energy taking into account the two regions of the box from Fig. 1.2 is

$$E_0 = E_0(l) + E_0(L - l) = \left(\frac{1}{l} + \frac{1}{L - l} \right) \frac{\pi}{2} \sum_{n=1}^{\infty} n \,, \tag{1.12}$$

which is infinite, as in the continuum case. Now, as we have discussed above, while in principle there is nothing wrong with E_0 being infinite (since it is not an observable quantity), the Casimir force Eq. (1.10), being experimentally observable, must be finite. Computing this force, we find

$$F_C = -\frac{dE_0}{dl} = \left(\frac{1}{l^2} - \frac{1}{(L - l)^2} \right) \frac{\pi}{2} \sum_{n=1}^{\infty} n \,, \tag{1.13}$$

which is still infinite, unless $l = L/2$, where for which the Casimir force vanishes, $F_c = 0$. So it seems that just by putting a plate inside an empty box and nothing else we can obtain an infinite repulsive force, which of course makes no physical sense.

How can we understand this result? To begin with, one notes that the Dirichlet boundary conditions for the scalar field at $x = 0, l$, and L that force the quantization of its modes arise from the interaction between the field and the plates, which are themselves not infinitely thin and continuous but rather compose by finite-size atoms. So the problem here is that the free theory is too simplistic to give realistic predictions: a full theory should account for the interactions of the field ϕ with the atoms in the plate. In particular, field modes with very small wavelengths (that is, very high energies) will go straight through the plates and will not interact with the atoms that compose it. Therefore, assuming that *all* field modes will be quantized due to the plate is an inconsistent assumption.

In order to take this physical requirement into account, namely the the high energy modes should not

contribute to the Casimir force, what we can do is to regulate the sum in Eq. (1.12) by suppressing the contribution of the high frequency modes as follows:

$$E(a, \Lambda) = \frac{1}{2} \sum_{n=1}^{\infty} \omega_n(a) e^{-\omega_n(a)/\pi\Lambda}, \quad \omega_n(a) = \frac{\pi n}{a}, \tag{1.14}$$

where Λ is for the time being an arbitrary parameter with the same units as the frequencies ω_n, whose purpose will be revealed in short. The strategy then is to compute the zero-point energy $E_0(l)$ and the Casimir force $F_c(l)$ as a function of Λ and then at the end of the calculation take the limit of $\Lambda \to \infty$, that is, take the limit where the high-energy modes are not suppressed.

Adopting this strategy, what one gets is

$$E(a, \Lambda) = \frac{\pi}{2a} \sum_{n=1}^{\infty} n e^{-\epsilon n}, \qquad \epsilon = \frac{1}{a\Lambda}, \tag{1.15}$$

which can be expressed in terms of an infinite sum with a finite result,

$$E(a, \Lambda) = -\frac{\pi}{2a} \frac{\partial}{\partial \epsilon} \left(\sum_{n=1}^{\infty} e^{-\epsilon n} \right) = -\frac{\pi}{2a} \frac{\partial}{\partial \epsilon} \left(\frac{1}{e^\epsilon - 1} \right) = \frac{\pi}{2a} \left(\frac{1}{\epsilon^2} - \frac{1}{12} + \mathcal{O}\left(\epsilon^2\right) \right), \tag{1.16}$$

where terms suppressed when $\epsilon \to 0$ can be neglected, since this corresponds to the physical limit where $\Lambda \to \infty$. Combining now the contributions from the two parts of the box, one finds

$$E_0(l, \Lambda) = \frac{\pi}{2l} \left(l^2 \Lambda^2 - \frac{1}{12} + \mathcal{O}\left(\frac{1}{\Lambda^2}\right) \right), \tag{1.17}$$

$$E_0(L - l, \Lambda) = \frac{\pi}{2(L-l)} \left((L-l)^2 \Lambda^2 - \frac{1}{12} + \mathcal{O}\left(\frac{1}{\Lambda^2}\right) \right), \tag{1.18}$$

so that their sum is

$$E_0 = E_0(l) + E_0(L - l) = \frac{\pi}{2} L \Lambda^2 - \frac{\pi}{24l} + \mathcal{O}\left(\frac{1}{\Lambda^2}\right), \tag{1.19}$$

where in the last step we have assumed that the plate is not too close to the center of the box, so that the approximation $L \gg l$ holds.

Therefore we now find that the zero-point energy is still infinite in the $\Lambda \to \infty$ limit, but the Casimir force $F_c(l)$ is instead *finite*, since

$$F_c = -\frac{dE_0}{dl} = -\frac{\pi}{24l^2} + \mathcal{O}\left(\frac{1}{\Lambda^2}\right), \tag{1.20}$$

a result which is found to be in excellent agreement with the experiment. So what we have learned with this calculation? That observable quantities are independent on the specific details of how we treat the modes with the highest energy, in order words, on the details of the interaction between the scalar field ϕ and the plate. This means that once we properly regulate the UV physics, then the IR behaviour of the theory is well-defined, as one could have expected on general grounds.

The exponential suppression of the high energy modes of scalar field implemented by Eq. (1.14) is not the only possible way to regulate the ultraviolet divergences. One can check that other regulation strategies

lead to the same result for the Casimir force, including for instance:

- A hard cut-off:

$$E(a) = \frac{1}{2} \sum_n \omega_n \Theta(\pi\Lambda - \omega_n) \,, \tag{1.21}$$

- A Gaussian cut-off:

$$E(a) = \frac{1}{2} \sum_n \omega_n \exp\left(-\frac{\omega_n^2}{\pi^2 \Lambda^2}\right) \,, \tag{1.22}$$

- A ξ-function suppression:

$$E(a) = \frac{1}{2} \sum_n \omega_n \left(\frac{\omega_n}{\Lambda}\right)^{-\xi} \,. \tag{1.23}$$

While each of this regulators gives a different result for the (unobservable) vacuum energy, the physically measurable Casimir force is always the same. More generally, it is possible to show that any regulation strategy that has the form

$$E(a) = \sum_n \omega_n f\left(\frac{n}{a\Lambda}\right) \,, \tag{1.24}$$

and that satisfies that $f(0) = 1$ and

$$\lim_{x \to \infty} \left(x \partial^{(i)} f(x)\right) = 0 \tag{1.25}$$

will lead to the same result for the IR physics. In particular, note that the condition $f(0) = 1$ ensures that the regulator of the high energy divergences does not affect the infrared behaviour of the theory.

In summary, we have learned from this exercise that the Casimir force is independent of the regulator and that it can be understood as a purely infrared (as opposed to ultraviolet) effect: the IR physics is effectively *decoupled* from the behaviour of the theory in the ultraviolet. In this respect, the infinite result that we obtained in Eq. (1.13) was a consequence of applying the theory beyond its regime of validity, in particular by neglecting the interactions between the scalar field and the plate.

1.2   One-loop calculations in $\lambda\phi^4$ theory

As we will show now, the main features of the renormalization procedure of a generic QFT can be already discussed in the relatively simple $\lambda\phi^4$ theory, so this is what we will use in the following. This theory contains a massless scalar field ϕ with a quartic self-interaction, and is described by the Lagrangian Eq. (1.1). For the $2 \to 2$ scattering amplitude, it is easy to verify that there exists a single tree level diagram, which gives

$$i\mathcal{M}_0 = -i\lambda \,, \tag{1.26}$$

by using the Feynman rules of the theory. The goal of this section is to compute $\mathcal{M}_1$, the next term in the perturbative expansion of the amplitude $\mathcal{M}$, which is the term of the order $\mathcal{O}(\lambda^2)$:

$$\begin{aligned} \mathcal{M} &= \mathcal{M}_0 + \mathcal{M}_1 + \mathcal{M}_2 + \ldots \\ &= \mathcal{O}(\lambda) + \mathcal{O}(\lambda^2) + \mathcal{O}(\lambda^3) + \ldots \end{aligned} \tag{1.27}$$

When doing the calculation, we will then encounter a number of infinities, which will be dealt with by means of a regularization procedure in the same spirit as that of the Casimir effect. We will then show how by
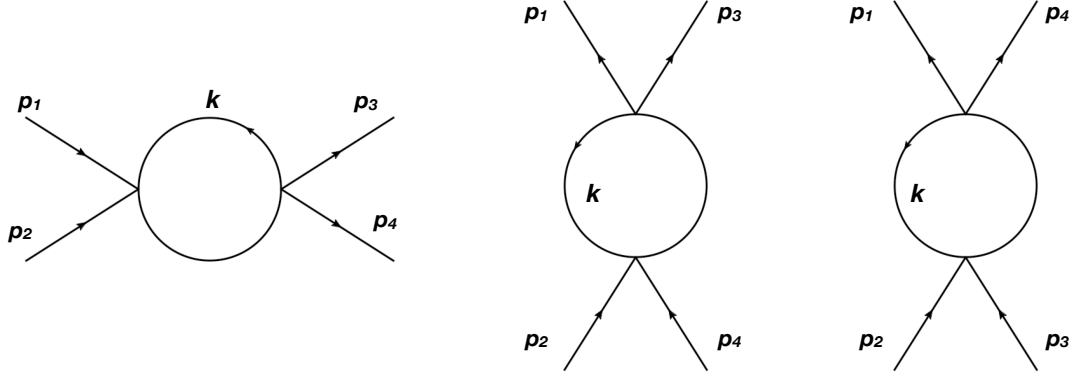
Figure 1.3. The Feynman diagrams relevant the calculation of the $\mathcal{O}(\lambda^2)$ contribution to the $2 \to 2$ scattering amplitude in $\lambda\phi^4$ theory. From left to right the topologies correspond to s-channel, t-channel, and u-channel respectively.

means of a suitable renormalisation procedure physical predictions of the theory become strictly independent of the regulator, as one would expect for a sensibly behaved theory.

There are three diagrams that contribute to the one-loop amplitude $\mathcal{M}_1$, corresponding to the s-, t- and u-channels diagrams respectively. These Feynman diagrams are represented in Fig. 1.3. Let us focus for the time being in the contribution from the s-channel diagram, the leftmost one in Fig. 1.3, which already highlights the most representative features of the problem. If we define $p \equiv (p_1 + p_2) = (p_3 + p_4)$ (from momentum conservation), then the application of the Feynman rules of the $\lambda\phi^4$ theory leads to the following amplitude:

$$\mathcal{M}_{1,s} = \frac{(i\lambda)^2}{2} \int \frac{d^4k}{(2\pi)^4} \frac{1}{i} \frac{1}{k^2 - i\epsilon} \frac{1}{i} \frac{1}{(p-k)^2 - i\epsilon} \,, \tag{1.28}$$

where $1/2$ is a symmetry factor and we integrate over the four-momentum k^μ going through the loop. For the time being we are setting the scalar field to be massless, $m = 0$, later we will study the physical consequences of a finite mass.

To evaluate loop integrals that arise in quantum field theory, there are a number of mathematical identities that can be used, and go over the name of *Feynman integral formulae*. The one that we need here is the following:

$$\frac{1}{AB} = \int_0^1 dxdy\delta(x+y-1)\frac{1}{(xA+yB)^2} = \int_0^1 dx\frac{1}{(xA+(1-x)B)^2} \,, \tag{1.29}$$

which is based on the following idea:

- We start with a product of the form $(AB)^{-1}$, where A and B are typically propagators corresponding to particles appearing in loop diagrams.

- Then we express this product in terms of integrals over auxiliary variables x and y.

- We carry out integration over y to end up with an integral over the auxiliary variable x, known as the *Feynman parameter*.

- The advantage is that in the resulting expression the inverse propagators A and B appear in an additive rather than multiplicative form, which facilitates the calculation of the loop matrix elements.

- Moreover, an analogous expression can be found for any number of propagators: The Feynman integral manipulations allows one to express a product of propagators in terms of their sum, which is much simpler to deal with.

In deriving Eq. (1.29), we have used the fact that one can write

$$\frac{1}{AB} = \int_0^1 \frac{dx}{(A + (B - A)x)^2} \,.$$

$$(1.30)$$

Indeed, one can derive the corresponding Feynman formula Eq. (1.29) for the combination of an arbitrary number of propagators to find:

$$\frac{1}{A_1 \ldots A_n} = \int dF_n \left(\sum_{k=1}^n x_k A_k \right)^{-n} ,$$

$$(1.31)$$

with the integration measure over the Feynman parameters x_i is given by

$$\int dF_n = (n-1)! \int_0^1 \left(\prod_{k=1}^n dx_k \right) \delta \left(x_1 + \ldots + x_n - 1 \right) .$$

$$(1.32)$$

You can check that Eq. (1.31) reduces to Eq. (1.29) in the case that $n = 2$.

Let us exploit the mathematical relationship of Eq. (1.29) to compute our loop integral Eq. (1.28). We observe that if we assign $A = (p - k)^2 - i\epsilon$ and $B = k^2 - i\epsilon$, then the amplitude can be rewritten using Eq. (1.29) and some additional algebra as follows:

$$i\mathcal{M}_{1,s} = \frac{\lambda^2}{2} \int \frac{d^4k}{(2\pi)^4} \int_0^1 \frac{dx}{\left((p^2 - 2p \cdot k)(1 - x) + k^2 - i\epsilon \right)^2} \,.$$

$$(1.33)$$

Here at this point it becomes useful to point out a couple of observations:

- Recall here that $p = p_1 + p_2$ corresponds to the total energy available for the particle scattering in the center-of=mass frame, $s = p^2 \neq 0$.

- Note also that the scalar particle in the loop is not on-shell, so we cannot impose that $k^2 = 0$ even if the particle is massless.

Next, in order to compute the loop matrix element Eq. (1.33), we need to perform the following change of variables within the integral over the virtual particle four-momentum k:

$$\widetilde{k} \equiv k - p(1 - x) \,, \quad d\widetilde{k} = dk \,.$$

$$(1.34)$$

which means that

$$k = \widetilde{k} + p(1 - x) \,, \qquad \widetilde{k}^2 = k^2 - 2k \cdot p(1 - x) + p^2(1 - x)^2 \,.$$

$$(1.35)$$

Note that this change of variables does not affect the integration ranges, since it is just a constant offset (for a fixed x) and the integral in Eq. (1.33) is unbounded. So the integration limits of the loop integral, that determine the ultraviolet behaviour of the scattering amplitude, are not modified by this constant shift. One can verify that this change of variables implies

$$p^2(1 - x) + k^2 - 2p \cdot k(1 - x) = p^2(1 - x) + \widetilde{k}^2 - p^2(1 - x)^2 = \widetilde{k}^2 + p^2 x(1 - x) \,,$$

$$(1.36)$$

and therefore (going back to denoting k the loop integration variable for simplicity, since it is a dummy variable anyway) we see that Eq. (1.33) can be expressed as

$$i\mathcal{M}_{1,s} = \frac{\lambda^2}{2} \int \frac{d^4k}{(2\pi)^4} \int_0^1 dx \frac{1}{\left(k^2 + p^2 x(1-x) - i\epsilon\right)^2} \, , \tag{1.37}$$

which as we will see now is rather simpler to handle with than the original loop integral. In particular, what we have achieved with these manipulations is that

- Now the loop momentum k only appears in the integrand as squared, k^2, and not contracted with any other four-vector.

  This is easier to integrate since we can go to a coordinate system where we only need to do a single "radial" integration and the integral over the "angular" variables is trivial.

The next step which is necessary in order to compute the integral is to interchange the integration order between x and k, and this way one finds the following result:

$$i\mathcal{M}_{1,s} = \frac{\lambda^2}{2} \int_0^1 dx \int \frac{d^4k}{(2\pi)^4} \frac{1}{\left(k^2 + p^2 x(1-x) - i\epsilon\right)^2} \, , \tag{1.38}$$

To make further progress, we now need to evaluate the integral over the loop four-momentum k^μ. To do this, one notices that this integral has the following form

$$I_{1,s}(x) \equiv \int \frac{d^4k}{(2\pi)^4} \frac{1}{\left(k^2 + \Delta - i\epsilon\right)^2} \, , \tag{1.39}$$

with $\Delta \equiv p^2 x(1-x)$ being a constant (k-independent) term from the point of view of the integration over k.

Just from dimensional analysis reasons, it is easy to convince yourselves that an integral of the form Eq. (1.39) has a logarithmic divergence in the ultraviolet regime, since in this case

$$I_{1,s} \simeq \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^4} \simeq \int \frac{dk}{k} \, , \tag{1.40}$$

since note that, in the $k \to \infty$ limit that determines the UV behaviour of the integral we can discard the finite $\Delta \propto p^2$ term. Therefore, before we can try to make sense of our loop integral, we need to regularise this ultraviolet divergence. This is analogous to what we did in the case of the Casimir effect, where we regularised the small wavelength / large energy divergence by damping the corresponding modes when computing the vacuum energy.

From the physical point of view, some considerations are relevant at this point:

- The fact that a regularisation procedure is necessary indicates that our theory Eq. (1.1) is an *effective theory* only valid below a certain energy cutoff, and we cannot expect it to be applicable to arbitrarily large energies (or that is the same, arbitrarily small distances).

- It is only once we manage to show that physical observables are independent of the regulator (as well as the specific regularisation strategy adopted) that we know that physical predictions will be reliable: this is what is known as renormalisation procedure.

- If we regulator does not disappear, this means that predictions for physical processes that take place at an energy $E \ll \Lambda$ depend on features of the theory at $E \simeq \Lambda$, in other words, that we are extremely sensitive to the (potentially unknown) physics of the ultraviolet, and therefore the results obtained cannot be trusted.

In analogy with the discussion of the Casimir effect, there exist different methods that can be used to regularise the divergent loop integrals that appear in quantum field theories. Each of these methods has their advantages and disadvantages, but in any case for renormalizable theories the results must the independent on the choice of regularisation strategy.

A popular regularization strategy is known as the *dimensional regularization* method, introduced by the Dutch physicists Gerhard 't Hooft and Martinus Veltman in 1972. This method is based on the realization that if instead of working in $d = 4$ space-time dimensions we assume a generic d-dimensional space-time, then the integral

$$I_{1,s}^{(d)} \equiv \int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 + \Delta - i\epsilon)^2} \,, \tag{1.41}$$

is actually convergent for $d < 4$, rather than logarithmically divergent as happens for the $d = 4$ case. So the idea of the dimensional regularization method is to evaluate $I_{1,s}^{(d)}$ for those values of d for which it converges, and then analytically continue the result to the physical limit $d \to 4$.

Alternatively, this integral could also be regularized using the Λ cutoff strategy adopted for the study of the Casimir effect. In this regularisation method, the integration over UV momenta is cut off for $|k| \geq \Lambda$, and afterwards one needs to demonstrate the physical predictions are unchanged if one takes the $\Lambda \to \infty$ limit. Indeed a wide range of alternative regularization methods exist, all leading to equivalent predictions for physical observables.

In order to be able to compute Eq. (1.41) by means of dimensional regularisation, the first step is to perform a *Wick rotation* and go to an *Euclidean signature*. By this we mean that instead of using the Minkowski flat-space time metric

$$g_{\mu\nu} = (-1, +1, +1, +1) \,, \tag{1.42}$$

we transform coordinates in order to end up with the Euclidean metric

$$g_{\mu\nu} = (+1, +1, +1, +1) \,, \tag{1.43}$$

To achieve this, the Wick rotation is defined by the transformation $k^0 = i k_E^d$ (in other words, rotating to imaginary time) and $k^i = k_E^i$, such that $k^2 = k_E^2$ where

$$k_E^2 = (k_E^1)^2 + \ldots + (k_E^d)^2 \,. \tag{1.44}$$

In addition, it can be show that Wick-rotated momenta satisfy the following property:

$$\int d^d k f \left(k^2 - i\epsilon \right) = i \int d^d k_E f \left(k_E^2 \right) \,, \tag{1.45}$$

provided $f(k_E^2) \to 0$ faster than k_E^{-d} as $k_E \to \infty$. The latter property is true for the divergent loop integral that we are considering here in $d < 4$ dimensions, so that can use this property to replace $k^2 - i\epsilon$ by k_E in Eq. (1.41). Note that here the change of variables that we performed in Eq. (1.34) was crucial since else we

would have factors of the type $k \cdot p$ in the integrand and carrying out the Wick rotation would have been more difficult (since one needs to carry out angular integrals).

After applying the Wick rotation the integral Eq. (1.41) can be expressed as follows using spherical coordinates,

$$I_{1,s}^{(d)} = i \int \frac{d^d k_E}{(2\pi)^d} \frac{1}{(k_E^2 + \Delta)^2} = \frac{i}{(2\pi)^d} \int d\Omega_d \int_0^\infty dk_E k_E^{d-1} \frac{1}{(k_E^2 + \Delta)^2} \,, \tag{1.46}$$

where Ω_d is the differential solid angle of the d-dimensional sphere, which is given by

$$\Omega_d \equiv \int d\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)} \,, \tag{1.47}$$

and $\Gamma(x)$ is Euler's Gamma function, ubiquitous in QFT loop calculations.

Euler's Gamma function

The Gamma function $\Gamma(x)$ is a generalization of the factorial function to real and complex numbers. For n a positive integer one has

$$\Gamma(n) = (n-1)! \,, \tag{1.48}$$

while for all complex numbers except non-positive integers one has

$$\Gamma(x) = \int_0^\infty z^{x-1} e^{-z} dz \,. \tag{1.49}$$

The Gamma functions has poles for $n = 0, -1, -2, -3, \ldots$. Some useful relations are

$$\Gamma\left(n + \tfrac{1}{2}\right) = \frac{(2n!)}{n! 2^{(2n)}} \sqrt{\pi} \,, \tag{1.50}$$

$$\Gamma(-n + x) = \frac{(-1)^n}{n!} \left[\frac{1}{x} - \gamma_E + \sum_{k=1}^n \frac{1}{k} + \mathcal{O}(x) \right] \,, \tag{1.51}$$

with $\gamma_E = 0.5772$ the Euler-Mascheroni constant, and

$$\Gamma\left(\frac{\epsilon}{2}\right) = \frac{2}{\epsilon} - \gamma_E + \mathcal{O}(\epsilon) \,. \tag{1.52}$$

which is useful in the limit $\epsilon \to 0$.

The computation of the radial integral gives the following result:

$$\int_0^\infty dk_E k_E^{d-1} \frac{1}{(k_E^2 + \Delta)^2} = \Delta^{d/2-2} \Gamma(d/2) \Gamma\left(\frac{4-d}{2}\right) \frac{1}{2\Gamma(2)} \,. \tag{1.53}$$

You can convince yourselves that this integral is only convergent for $d < 4$, and is divergent otherwise. The integral Eq. (1.53) is a particular case of the general result

$$\int \frac{d^d k_E}{(2\pi)^d} \frac{(k_E^2)^a}{(k_E^2 + \Delta)^b} = \frac{\Gamma(b - a - d/2)\Gamma(a + d/2)}{(4\pi)^{d/2} \Gamma(b) \Gamma(d/2)} \Delta^{-(b-a-d/2)} \,, \tag{1.54}$$

in the case $a = 0$ and $b = 2$. This result is very useful for general loop calculations in QFT.

Combining the results from angular and radial integrals, we find that the integral that we wanted to compute gives

$$\int \frac{d^d k_E}{(2\pi)^d} \frac{1}{(k_E^2 + \Delta)^2} = \frac{1}{(4\pi)^{d/2}} \Gamma\left(\frac{4-d}{2}\right) \Delta^{d/2-2} \,, \tag{1.55}$$

where we have used that $2\Gamma(2) = 2$. The next step, in preparation for taking the limit $d \to 4$, we define $d \equiv 4 - \epsilon$ and expand around the physical $\epsilon \to 0$ limit, where terms that go to zero in this limit can be neglected, and we find

$$
\begin{aligned}
\int_0^\infty k_E^{d-1} dk_E \frac{1}{(k_E^2 + \Delta)^2} &= \frac{1}{(4\pi)^2} \left(\frac{2}{\epsilon} - \gamma_E + \mathcal{O}(\epsilon)\right)\left(1 - \frac{\epsilon}{2} \ln \Delta + \mathcal{O}(\epsilon^2)\right) \\
&= \frac{i}{(4\pi)^2} \left(\frac{2}{\epsilon} - \gamma_E - \ln \Delta + \mathcal{O}(\epsilon)\right) \,,
\end{aligned}
\tag{1.56}
$$

where we have Wick-rotated back to Minkowski metric.

Wick rotation

This mathematical operation is motivated by the fact that the Minkowski metric in natural units

$$ds^2 = -(dt)^2 + dx^2 + dy^2 + dz^2 \,, \tag{1.57}$$

is equivalent to the Euclidean metric in four spatial dimensions,

$$ds^2 = d\tau^2 + dx^2 + dy^2 + dz^2 \,, \tag{1.58}$$

provided that the time t is allowed to take negative values, $t = i\tau$ with τ a real variable. Wick rotation allows in certain cases to rewrite integration over Minkowski momenta in terms of integration over four-dimensional Euclidean momenta. Specifically, if k is a Minkowskian four-momentum and k_E is the corresponding Wick-rotated Euclidean version, it can be shown that in general:

$$\int d^d k f\left(k^2 - i\epsilon\right) = i \int d^d k_E f\left(k_E^2\right) \,, \tag{1.59}$$

provided $f(k_E^2) \to 0$ faster than k_E^{-d} as $k_E \to \infty$.

Now that we have computed Eq. (1.41), the next step is to perform the integral over the Feynman parameter x, Eq. (1.39). Using Eq. (1.56), we find the following result for the one-loop s-channel matrix element,

$$i\mathcal{M}_{1,s} = \frac{\lambda^2}{2} \left(\frac{-i}{16\pi^2}\right) \int_0^1 dx \left[\ln(p^2 x(1-x)) + \gamma_E - \frac{2}{\epsilon}\right] = -\frac{i\lambda^2}{32\pi^2} \ln \frac{p^2}{\Lambda^2} \,, \tag{1.60}$$

where

$$\ln \Lambda = \frac{1}{\epsilon} + \text{constant} \,, \tag{1.61}$$

and where we have used the definite integral computed below. In the above expression, $\mathcal{O}(\epsilon)$ terms have been neglected. Recall that p^2 is the square of the center-of-mass energy of the scattering particles. This is

---

the final result for the s-channel contribution to the one-loop scattering amplitude in $\lambda\phi^4$ theory, see Fig. 1.3.

Loop integrals

For the calculation of loop integrals in QFT a number of definite integrals are useful, such as

$$\int_0^1 dx \left[\ln(p^2 x(1-x)) + B\right] = \ln(p^2) + B - 2\,. \tag{1.62}$$

Another integral that appears frequently in loop calculations in QFT is

$$\int_0^\infty dk_E k_E^{d-1} \frac{1}{(k_E^2 + \Delta)^2} = \Delta^{d/2-2}\Gamma(d/2)\Gamma\left(\frac{4-d}{2}\right)\frac{1}{2\Gamma(2)}\,, \tag{1.63}$$

with k_E being a d-dimensional Euclidean momentum. This integral is a particular case of the more general expression:

$$\int \frac{d^d k_E}{(2\pi)^d} \frac{\left(k_E^2\right)^a}{\left(k_E^2 + \Delta\right)^b} = \frac{\Gamma(b-a-d/2)\Gamma(a+d/2)}{(4\pi)^{d/2}\Gamma(b)\Gamma(d/2)}\Delta^{-(b-a-d/2)}\,. \tag{1.64}$$

Note that the two integrals above can be divergent in the physical limit $d = 4$ and thus need to be regulated. Another integral that might arise in some calculations is of the form

$$\int_0^1 dx x(1-x)\left[\ln\left(\frac{-s}{m^2}\right) + \ln\left(\frac{x(1-x)}{1-x(1-x)}\right)\right] = \ln(-s/m^2) + 4 - \pi\sqrt{3}\,, \tag{1.65}$$

which appears for example when computing the self-energy of a massive scalar field.

At this point we can add together the tree-level and the one-loop contribution to construct the partial $\mathcal{O}(\lambda^2)$ (next-to-leading order, or NLO for short) matrix element for $2 \to 2$ scattering in $\lambda\phi^4$ theory. Note that this is not the full result, since the t- and u-channels are still missing, but as we will show it captures the essential physical features. The final result reads

$$\mathcal{M}_{\rm NLO}(s) = \mathcal{M}_0 + \mathcal{M}_{1,s} = -\lambda - \frac{\lambda^2}{32\pi^2}\ln\left(-\frac{s}{\Lambda^2}\right)\,, \tag{1.66}$$

with $s = -p^2$ being the center-of-mass energy of the $2 \to 2$ scattering. At a first glance, this result does not seem to make much sense, since it is infinite in the $\Lambda \to \infty$ ($\epsilon \to 0$) physical limit. Even the Born approximation becomes meaningless, since the next term in the perturbative expansion appears to be infinitely larger. How can we understand this behaviour?

The origin of the infinity in the NLO scattering amplitude Eq. (1.66) is not too different from that of the Casimir effect discussed in Sect. 1.1, namely extrapolating the theory beyond its range of validity, in this case the deep ultraviolet regime. We will next show how, despite these infinities, it is possible to make sense of the theory and provide finite well-defined predictions for the physical observables of the theory. The key observation to achieve this is that the coupling constant λ in Eq. (1.66), that is, the Lagrangian parameter, is *not* a physical observable.

1.3   The renormalized coupling constant

While the (partial) NLO amplitude $\mathcal{M}_{\mathrm{NLO}}(s)$ Eq. (1.66) seems to be still divergent, it is easy to check that the *difference* between two center-of-mass energies s_1 and s_2 is actually finite, since

$$\mathcal{M}_{\mathrm{NLO}}(s_1) - \mathcal{M}_{\mathrm{NLO}}(s_2) = \frac{\lambda^2}{32\pi^2} \ln \frac{s_2}{s_1} \,. \tag{1.67}$$

In other words, given either a *calculation* or a *measurement* of the amplitude $\mathcal{M}(s_1)$, we can use Eq. (1.67) to predict the behaviour of the scattering amplitude for any other value of s_2 without ever encountering infinities. So the question is: should we expect that $\mathcal{M}(s)$ is finite? It seems so, since after all this amplitude is the building block of the physical cross-sections. But maybe the situation is subtler than that?

The solution of this conundrum lies in the interpretation of the coupling constant λ, that determines the strength of the interactions in the theory. By construction, it is a measure of how strong is the self-interaction of the scalar field in the theory. At the Born level (leading order), the value of λ is determined directly from a measurement of the $\sigma(\phi\phi \to \phi\phi)$ scattering cross-section, recall Eq. (1.26). Now, at NLO the relation between the scattering amplitude and the coupling λ is more involved, since we need to account also for the $\mathcal{O}(\lambda^2)$ terms, see Eq. (1.66).

Let us do the following thing now. It will appear at first a bit arbitrary but its rationale will be soon revealed. We *define* a *renormalized coupling constant* λ_R as $-\mathcal{M}(s_0)$, the value of the $2 \to 2$ scattering amplitude $\sigma(\phi\phi \to \phi\phi)$ for a given fixed center of mass energy s_0. Note that by construction $\lambda_R = \lambda$ at the Born level, but in general the two coupling constants will be different once we account for loop (higher order) effects. From this definition we have that at the one-loop level, see Eq. (1.66), the renormalized coupling is given by

$$\lambda_R \equiv -\mathcal{M}(s_0) = \lambda + \frac{\lambda^2}{32\pi^2} \ln \left(-\frac{s_0}{\Lambda^2} \right) + \mathcal{O}(\lambda^3) \,, \tag{1.68}$$

or in other words, the difference between λ and λ_R is given by

$$\lambda_R - \lambda = \frac{\lambda^2}{32\pi^2} \ln \left(-\frac{s_0}{\Lambda^2} \right) + \mathcal{O}(\lambda^3) \,. \tag{1.69}$$

This equation relates λ, the parameter that appears in the Lagrangian of the theory, with λ_R, the measured value of the $\phi\phi \to \phi\phi$ scattering amplitude at a reference center of mass energy s_0. This identity can also be reversed to give

$$\lambda = \lambda_R - \frac{\lambda_R^2}{32\pi^2} \ln \left(-\frac{s_0}{\Lambda^2} \right) + \mathcal{O}(\lambda_R^3) \,, \tag{1.70}$$

where note that we have used that $\lambda = \lambda_R + \mathcal{O}\left(\lambda_R^2 \right)$, Eq. (1.69).

So far, we do not seem to have made much progress: after-all, the relation Eq. (1.68) still is ill-defined, since we need to take $\Lambda \to \infty$ in the physical limit. However, the situation changes once we write the scattering matrix element in terms of λ_R. Indeed, of express the NLO matrix element $\mathcal{M}_{\mathrm{NLO}}(s)$, Eq. (1.66), in terms of the renormalized coupling λ_R. What we find in this case is that

$$\mathcal{M}_{\mathrm{NLO}}(s) = -\lambda_R - \frac{\lambda_R^2}{32\pi^2} \ln \left(\frac{s}{s_0} \right) + \mathcal{O}(\lambda_R^3) \,, \tag{1.71}$$

which is not only perfectly finite in the physical $\Lambda \to \infty$ limit, but in addition gives a experimentally testable

prediction of how the scattering amplitude s depends with the center-of-mass energy of the collision, given a measurement of the same quantity at a reference center-of-mass-energy s_0.

So where have all the infinities gone? The key point here is that the parameter λ in the theory Lagrangian, Eq. (1.1), is *not an observable quantity* once we account for genuinely quantum (loop) effects. If we express our predictions in terms only of measurable quantities, such as λ_R, the value of the scattering amplitude at s_0, then the results turn out to be well-behaved. Eq. (1.71) illustrates the renormalization paradigm of quantum field theories, where UV infinities can be reabsorbed into a redefinition of the *bare* parameters of the Lagrangian. If you feel that this is some kind of dirty trick you are not alone: some of the founding fathers of quantum theory shared the same feeling when presented with the renormalization program. In the following, we study in some more details how the process works in more generality.

1.4 Absorbing infinities: counterterms

The starting point of this whole discussion was the Lagrangian of $\lambda\phi^4$ theory, Eq. (1.1). We found that if we compute amplitudes in terms of the coupling constant λ that appears in the Lagrangian, beyond the Born level we get unphysical (infinite) results, Eq. (1.66). On the other hand, we found that if the same amplitude if expressed in terms of the renormalized coupling constant λ_R, which is defined in terms of a physical cross-section (the value of the $\phi\phi \to \phi\phi$ scattering amplitude at $s = s_0$) then the amplitude is perfectly well defined, Eq. (1.71).

This observation suggest that in general, in order to tame the infinities that arise in QFT loop calculations, the following approach seems advisable:

(a) write down the Lagrangian of the theory in terms of the physical couplings and other observable parameters, rather than the bare ones; and then

(b) remove the UV divergences by adding additional terms to the Lagrangian, known as the *counter-terms*.

Coupling constant renormalization. In order to illustrate this strategy, let us consider the following redefinitions of the scalar field and the coupling constant in the Lagrangian of the $\lambda\phi^4$ theory:

$$\phi \to Z_\phi^{1/2}\phi\,, \tag{1.72}$$

$$\lambda \equiv \lambda_R Z_4\,, \tag{1.73}$$

so that in terms in terms of the renormalized field and coupling the Lagrangian of the theory now reads

$$\mathcal{L} = -\frac{1}{2}Z_\phi\partial_\mu\phi\partial^\mu\phi - \frac{\lambda_R}{4!}Z_\lambda\phi^4\,, \tag{1.74}$$

with $Z_\lambda \equiv Z_\phi^2 Z_4$. In Eq. (1.74) the terms Z_ϕ and Z_λ are general functions of the renormalized coupling λ_R, the energy scale s, and the cutoff of the theory Λ (or what is the same, of the regulator ϵ). That is, in general we have that

$$Z = Z(\lambda_R, s, \Lambda)\,. \tag{1.75}$$

These terms are called technically *counterterms*, and their role is to make sure that the theory gives finite predictions for the amplitude and other observables. To ensure a smooth matching with the tree-level (Born)

results, we should have that

$$Z_\phi = 1 + \mathcal{O}(\lambda_R)\,, \tag{1.76}$$

$$Z_\lambda = 1 + \xi_\lambda \lambda_R + \mathcal{O}(\lambda_R^2)\,, \tag{1.77}$$

where the coefficients of these λ_R expansions can be fixed by explicit loop calculations (as well as by more general methods).

It is easy to check that if one repeats the one-loop calculation presented in Sect. 1.2, but using the Lagrangian Eq. (1.74), one gets the following result,

$$\mathcal{M}_{\mathrm{NLO}} = -\lambda_R Z_\lambda - \frac{\lambda_R^2 Z_\lambda^2}{32\pi^2} \ln\left(-\frac{s}{\Lambda^2}\right)\,, \tag{1.78}$$

by comparing with Eq. (1.66). Note that by using Eq. (1.77) we can express the above equation as

$$\mathcal{M} = -\lambda_R \left(1 + \xi_\lambda \lambda_R\right) - \frac{\lambda_R^2 (1 + \xi_\lambda \lambda_R)^2}{32\pi^2} \ln\left(-\frac{s}{\Lambda^2}\right)\,. \tag{1.79}$$

Next, since we have *defined* the renormalized coupling as $\mathcal{M}(s_0) = -\lambda_R$, the above equation can be expressed as

$$\mathcal{M}(s_0) = -\lambda_R \left(1 + \xi_\lambda \lambda_R\right) - \frac{\lambda_R^2}{32\pi^2} \ln\left(-\frac{s_0}{\Lambda^2}\right) + \mathcal{O}\left(\lambda_R^3\right) = -\lambda_R\,, \tag{1.80}$$

which requiring the cancellation of the coefficient of the $\mathcal{O}\left(\lambda_R^2\right)$ term gives

$$Z_\lambda = 1 - \frac{\lambda_R}{32\pi^2} \ln\left(-\frac{s_0}{\Lambda^2}\right) + \mathcal{O}\left(\lambda_R^2\right)\,. \tag{1.81}$$

Therefore, we have determined that that renormalizing the bare fields and coupling constants of the $\lambda\phi^4$ theory as indicated in Eq. (1.74), one automatically obtains finite results at the one-loop order provided that the coupling constant counterterm is taken to be Eq. (1.81).

Note that the counterterm is also infinite in the physical limit $\Lambda \to \infty$, but this has no impact on observable quantities since these turn out to be independent of Λ. We also note here that that to cure the ultraviolet divergence that appears in $\phi\phi \to \phi\phi$ scattering we do not need the scalar field renormalization counter-term Z_ϕ, at least at the one-loop level, but this will be eventually necessary for other processes.

While here we have only sketched the surface of a rather complex procedure, it is enough to understand that renormalization is a program that systematically makes observables finite which gives testable predictions about the IR physics despite the UV divergences. The core idea is that in renormalizable QFT physical observables are well-defined once a finite number of counterterms are introduced, whose values in turn are fixed by a finite number of measurements. Also note that loops can produce behaviour which are completely different from anything that can be generated at tree level. In particular, non-analytic behaviour such as $\ln(s/s_0)$ is characteristic of loops, while tree-level amplitudes are always rational polynomials.

Mass renormalization and the renormalized propagator. Another important aspect of the renormalization program is the renormalization of the free-particle propagator. Similarly to the case of the scattering amplitudes, this will lead to the realization that the bare mass m of the Lagrangian is not a physical observable, and we will introduce a better defined *renormalized mass.*

Our starting point to study the renormalized propagator is the Lagrangian of the $\lambda\phi^4$ theory before
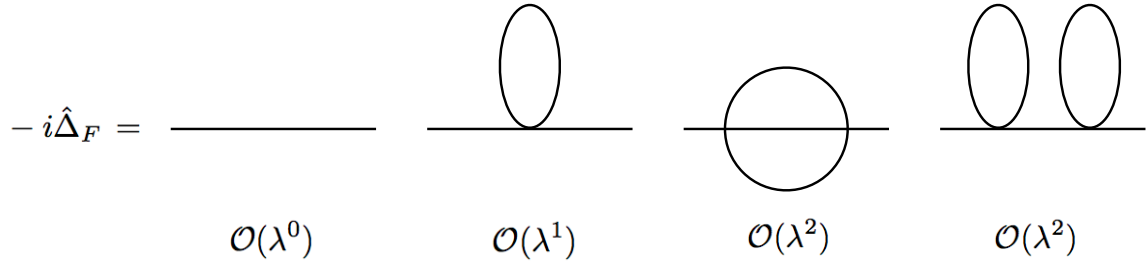
Figure 1.4. Some of the the Feynman diagrams that contribute to the calculation of the propagator $\mathbf{\Delta}_F$ in the interacting $\lambda\phi^4$ theory up to $\mathcal{O}\left(\lambda^2\right)$.

adding the counterterms, this time accounting for the mass of the scalar field:

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 \,, \tag{1.82}$$

and now try to compute the two-point function $\phi \to \phi$, in other words, the scalar field propagator. Recall the definition of the Feynman propagator:

$$\Delta_F(x_1 - x_2) = i\left\langle T\left(\phi(x_1)\phi(x_2)\right)\right\rangle_{(0)} \,, \tag{1.83}$$

where the sub-index (0) indicates that the expectation value should be computed using the free-theory states, rather than those of the interacting theory, and where T indicates time ordering. Once we account for the self-interactions of the scalar field, the two point function will be modified, leading to an *interacting propagator* $\mathbf{\Delta}_F$ given by

$$\mathbf{\Delta}_F(x_1 - x_2) \equiv i\left\langle T\left(\phi(x_1)\phi(x_2)\right)\right\rangle \,. \tag{1.84}$$

Our goal here is to evaluate the interacting propagator, but before that it is useful to understand its structure in perturbation theory. The propagator in the interacting theory will be determined by computing those diagrams in Fig. 1.4, where we show some of the diagrams that contribute to the result up to $\mathcal{O}\left(\lambda^2\right)$. As illustrated by these diagrams, the self-coupling of the scalar field ϕ affects how this same field propagates in space-time.

An alternative way to represent the perturbative expansion of the propagator in Fig. 1.4 is achieved by reorganizing it in terms of *one-particle-irreducible* (1PI) diagrams, which are defined by those diagrams that cannot be reduced (by cutting) in terms of diagrams with two or more propagators.[1] As shown in Fig. 1.5, the propagator is the sum of the free theory terms, plus all 1PI diagrams, plus all diagrams that contain 2 1PI contributions, and so on. Note that the 1PI diagrams cannot be associated a unique power of the coupling λ, since in general they receive contributions to all orders.

Furthermore, from the Feynman rules of the theory, one observes that the dependence on the momenta k of the various pieces shown in Fig. 1.5 will be the following:

- free theory: $\frac{1}{i}\Delta_F(k^2)$,

- 1PI diagrams: $\frac{1}{i}\Delta_F(k^2)i\Sigma(k^2)\frac{1}{i}\Delta_F(k^2)$,

- diagrams with 2 1PI contributions: $\frac{1}{i}\Delta_F(k^2)i\Sigma(k^2)\frac{1}{i}\Delta_F(k^2)i\Sigma(k^2)\frac{1}{i}\Delta_F(k^2)$,

---

[1] In other words, a diagram is 1PI if it is still connected after any one line is cut.
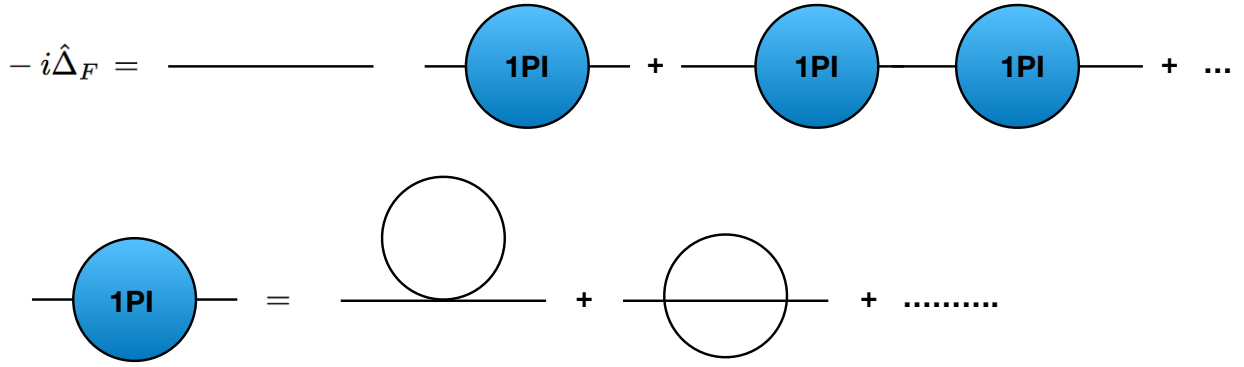
Figure 1.5. The perturbative expansion of the interacting propagator $\mathbf{\Delta}_F$ in Fig. 1.4 can also be organized in terms of diagrams with different number of 1PI diagrams.

and so on, where $i\Sigma(k^2)$ is known as the *self-energy* contribution and contains the effects of all 1PI diagrams to all powers of λ without the propagators from the two external lines.

The reason to organize the perturbative expansion of the propagator of $\lambda\phi^4$ theory as illustrated in Fig. 1.5 is that the infinite series

$$\frac{1}{i}\mathbf{\Delta}_F(k^2) = \frac{1}{i}\Delta_F(k^2) + \frac{1}{i}\Delta_F(k^2)i\Sigma(k^2)\frac{1}{i}\Delta_F(k^2) + \frac{1}{i}\Delta_F(k^2)i\Sigma(k^2)\frac{1}{i}\Delta_F(k^2)i\Sigma(k^2)\frac{1}{i}\Delta_F(k^2) + \dots, \quad (1.85)$$

has the form of a geometric series and therefore can be summed to all orders, assuming of course that the perturbative expansion is convergent. By doing this, we end up with the following expression of the interacting propagator expressed in terms of the self-energy $\Sigma(k^2)$ of the scalar field:

$$\mathbf{\Delta}_F(k^2) = \frac{1}{(\Delta_F(k^2)/i)^{-1} - i\Sigma(k^2)} = \frac{1}{k^2 + m^2 - i\epsilon - \Sigma(k^2)}, \quad (1.86)$$

where we have used the expression for the free-theory propagator

$$\Delta_F(k^2) = \frac{1}{k^2 + m^2 - i\epsilon}, \quad (1.87)$$

which has a pole at $k^2 = -m^2$, with m being the bare mass. From Eq. (1.86), we observe that the net effect of the loop corrections (to all orders) in the propagator of $\lambda\phi^4$ theory is a redefinition of the effective mass of the propagating particles, by an amount which is determined by the self-energy $\Sigma(k^2)$ which in turn is computed from the 1PI diagrams.[2]

At this point we see how the classical and QFT definitions of a particle mass start to diverge. At the classical level, the *physical mass m_p* of a particle is simply $m_p = m$, the parameter of the Lagrangian. However from Eq. (1.86) we see that a more robust definition of m_p is the solution $k^2 = -m_p^2$ of

$$k^2 + m^2 - \Sigma(k^2) = 0, \quad (1.88)$$

namely the pole of the all-orders propagator. In other words, m_p becomes dependent on the perturbative order at which we are working (just as in the case of the renormalized coupling). In addition, this implies

---

[2]Since the self-energy is in general complex, loop corrections will modify not only the physical mass but also the *width* of the propagating particles.

that at the quantum level the physical mass $m_p(k^2)$ in general depends on the value of k^2, rather than being fixed as in the classical theory, just as the renormalized coupling depends on s. To be a bit more precise, $k^2 + m^2 - \mathrm{Re}(\Sigma(k^2)) = 0$ determines the physical mass of the particle m_p, while $\mathrm{Im}(\Sigma(k^2))$ determines the its width Γ_p, which fixes the decay rate of the particle. For stable particles, $\Gamma_p = 0$.

In order to determine the quantum corrections to the mass in $\lambda \phi^4$ theory (analogously as that we have done for the quantum corrections to λ from the scattering amplitude), we need to evaluate the one-loop correction to the self-energy $\Sigma(k^2)$. From the Feynman diagram in Fig. 1.5, we see that the one-loop correction to the self-energy is given by

$$i\Sigma(k^2) = -\frac{i\lambda}{2} \int \frac{d^4p}{(2\pi)^4} \frac{1}{i} \frac{1}{p^2 + m^2 - i\epsilon} + \mathcal{O}(\lambda^2) , \tag{1.89}$$

where we have used the fact that the coupling constant λ does not depend on the energy k at this perturbative order, or in other words, that $\lambda - \lambda_R = \mathcal{O}\left(\lambda^2\right)$. This integral is quadratically divergent: to see this, note that in the UV limit one has that

$$i\Sigma(k^2) \simeq -\frac{i\lambda}{2} \int \frac{p^3 dp}{p^2} \simeq \lambda \int p \, dp \to \infty . \tag{1.90}$$

However, we are now more confident that we can make sense of this apparently unphysical result by using a similar renormalization strategy as in the case of the coupling constant.

Therefore let us try to compute the integral Eq. (1.89) using dimensional regularization. Computing the integral in d Euclidean dimensions after a Wick rotation, we find that

$$\int \frac{d^d p_E}{(2\pi)^d} \frac{1}{p_E^2 + m^2} = \Gamma\left(1 - \frac{d}{2}\right) (m^2)^{d/2-1} \frac{1}{(4\pi)^{d/2}} \tag{1.91}$$

$$= \frac{1}{(4\pi)^{d/2}} \left(-\frac{2}{\epsilon} - \gamma_E + \mathcal{O}(\epsilon)\right) \left(m^2 - \epsilon \frac{m^2}{2} \ln m^2 + O(\epsilon^2)\right) , \tag{1.92}$$

where in the second step as usual we have used that $d = 4 - \epsilon$, together with some of the properties of the Gamma function. Defining the UV cutoff as $\ln \Lambda \equiv 2/\epsilon + \gamma_E$, we have that the one-loop expression for the scalar field self-energy in $\lambda \phi^4$ theory is

$$\Sigma(k^2) = \lambda \frac{m^2}{32\pi^2} \ln\left(\frac{m^2}{\Lambda^2}\right) + \mathcal{O}\left(\lambda^2\right) . \tag{1.93}$$

The fact that the self-energy is divergent in the physical $\Lambda \to \infty$ limit is now something that does not worry us, since we know that the bare mass m is not an observable quantity. The relevant physical requirement is that the physical mass (the pole of the all-orders propagator) is finite, and therefore we need to impose that

$$k^2 + m^2 - \mathrm{Re}\left(\Sigma(k^2)\right) = 0 , \tag{1.94}$$

when $k^2 = -m_p^2$, is finite as well. In this case we find that the Lagrangian mass m can be expressed in terms of the physical mass m_p and the renormalized coupling λ_R as follows

$$m^2 = m_p^2 \left(1 + \frac{\lambda_R}{32\pi^2} \ln\left(\frac{m_p}{\Lambda^2}\right) + \mathcal{O}(\lambda_R^2)\right) , \tag{1.95}$$

where we have used the mass-shell condition $k^2 = -m^2$. Note that as in the case of the renormalized coupling, one has $m - m_p = \mathcal{O}(\lambda_R)$. In general, when we compute the self-energy $\Sigma(k^2)$ in perturbation theory, its structure will have the form

$$\Sigma(k^2) = Ak^2 + Bm^2 \,, \tag{1.96}$$

where A and B are functions of the parameters of the theory as well as of the relevant scales such as λ_R, Λ, m_p^2 and k.

The renormalized Lagrangian. The formal way to regulate the divergences that arise in the two point function is the same as in the case of the coupling constant, namely one needs to add suitable counterterms to the Lagrangian of the theory

$$\mathcal{L} = -\frac{1}{2}Z_\phi \partial_\mu \phi \partial^\mu \phi - \frac{1}{2}Z_m m_p^2 \phi^2 - \frac{\lambda_R}{4!}Z_\lambda \phi^4 \,, \tag{1.97}$$

which is the same as Eq. (1.74) with the additional mass counterterm Z_m. This Lagrangian can also be written as

$$
\begin{aligned}
\mathcal{L} &= \mathcal{L}_{\text{bare}} + \mathcal{L}_{\text{ct}} \,, \\
\mathcal{L}_{\text{bare}} &= -\frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}m_p^2 \phi^2 - \frac{\lambda_R}{4!}\phi^4 \,, \\
\mathcal{L}_{\text{ct}} &= -\frac{1}{2}(Z_\phi - 1)\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}(Z_m - 1)m_p^2 \phi^2 - \frac{\lambda_R}{4!}(Z_\lambda - 1)\phi^4 \,,
\end{aligned}
\tag{1.98}
$$

explicitly separating the bare terms from the counterterms, which vanish at the Born level where $Z = 1$.

From the renormalized Lagrangian Eq. (1.98) one can derive a new set of Feynman rules from which finite observables will be obtained by construction. These Feynman rules are collected in Fig. (1.6), where now the scalar field propagator is expressed in terms of the physical mass

$$\frac{-i}{k^2 + m_p^2 - i\epsilon} \,, \tag{1.99}$$

and there are three interaction terms: two four-point interactions, with couplings $-\lambda_R$ and $-\lambda_R (Z_\lambda - 1)$ respectively, and a two-point interaction vertex with an associated Feynman rule

$$-i(Z_\phi - 1)k^2 - i(Z_m - 1)m_p^2 \,. \tag{1.100}$$

The main advantage of this formalism is that once the expressions for the counterterms Z_i is known at any given order in perturbation theory, any physical quantity computed from Eq. (1.98) will automatically lead to well-defined, finite results.

To summarize, arbitrary physical predictions in $\lambda \phi^4$ will be finite and well-defined once we determine the three counterterms Z_ϕ, Z_m and Z_λ at a given order in perturbation theory. These counterterms must be fixed by requiring the following conditions:

- Scattering amplitudes should be finite, *e. g.* $\mathcal{M}(s_0) = -\lambda_R$

- The interacting propagator $\mathbf{\Delta}_F$ has a pole at the physical mass m_p with residue one, which in turn

$$\frac{-i}{k^2 + m_p^2 - i\epsilon}$$

$$- i \left(Z_\phi - 1 \right) k^2 - i \left(Z_m - 1 \right) m_p^2$$

$$- \lambda_R$$

$$- \lambda_R \left(Z_\lambda - 1 \right)$$

Figure 1.6. The modified Feynman rules of $\lambda \phi^4$ theory once the bare Lagrangian is supplemented by the counterterms required to regulate the UV divergences, see Eq. (1.98). The second term should be understood as an interaction vertex term.

implies

$$\Sigma(k^2 = -m_p^2) \;\; = \;\; 0 \,, \tag{1.101}$$

$$\frac{\partial}{\partial k^2} \Sigma(k^2 = -m_p^2) \;\; = \;\; 0 \,.$$

Imposing these conditions, one finds that the $\mathcal{O}(\lambda_R)$ expressions for the counterterms are

$$
\begin{aligned}
Z_\phi &= 1 + \mathcal{O}(\lambda_R^2) \,, \\
Z_m &= 1 + \frac{\lambda_R}{32\pi^2} \ln \left(\frac{m_p^2}{\Lambda^2} \right) + \mathcal{O}(\lambda_R^2) \,, \\
Z_\lambda &= 1 - 3 \frac{\lambda_R}{32\pi^2} \ln \left(\frac{-s}{\Lambda^2} \right) + \mathcal{O}(\lambda_R^2) \,,
\end{aligned}
\tag{1.102}
$$

where the factor 3 in the last equation arises from other one-loop diagrams beyond the s-channel one, as will be shown below.

1.5   The renormalization group equations

In the previous lectures we have seen how we can absorb the divergences that appear in loop calculations in quantum field theory by means of a redefinition of the physical parameters of the theory. This renormalisation process made possible to obtain well-defined, finite results for scattering amplitudes and propagators beyond the Born level in $\lambda \phi^4$ theory. Now we explore some of the more general consequences of this renormalisation procedure, in particular studying the information that we can obtain for the scale dependence of physical parameters and amplitudes that is a direct consequence of this procedure.

Indeed, perhaps one of the most important consequences of the renormalization procedure is that quan-

tities such as the mass or the coupling constants are not fixed parameters anymore, but rather become functions of the energy of the process we are considering. For instance, in the case of $\lambda\phi^4$ theory, we have seen that for the scattering amplitude $\mathcal{M}_{\rm NLO}(s)$, once it is measured at $s = s_0$, then its dependence for any other value of s is determined by the result of our one-loop calculation. As we have seen, this also fixes the scale dependence of the coupling constants and the masses of the theory. As will be discussed in the tutorial session, the full renormalisation of the theory at the one loop level gives for the relation

$$\lambda = \lambda_R - 3\frac{\lambda_R^2}{32\pi^2}\ln\left(-\frac{\mu^2}{\Lambda^2}\right),\tag{1.103}$$

which specifies how $\lambda_R(s)$ depends on the center of mass energy s.

Here we formalize a bit more these ideas by presenting the *renormalization group equations*, a central construction in QFTs which allows to relate physical quantities at different scales. Let us start with the renormalized NLO scattering amplitude in $\lambda\phi^4$ theory, Eq. (1.71),

$$\mathcal{M}_{\rm NLO}(s) = -\lambda_R - \frac{\lambda_R^2}{32\pi^2}\ln\frac{s}{s_0} + \mathcal{O}\left(\lambda_R^3\right),\tag{1.104}$$

where the renormalized coupling had been defined by the condition $\mathcal{M}(s_0) = -\lambda_R$, namely the behaviour of λ_R is determined one one specifies the value of the scattering amplitude at the reference scale s_0. However, this reference scale s_0 is completely arbitrary: it was nowhere to be seen neither in the original nor in the renormalized Lagrangian of the theory. From this observation we can derive the following powerful consequence on the four-point scattering amplitude:

> Since the specific choice of s_0 is ultimately arbitrary, *physical observables must be independent of this value of s_0*. Therefore, from general considerations one must require that the one-loop scattering amplitude is independent of s_0,
> $$\frac{d\mathcal{M}_{\rm NLO}(s)}{ds_0} = 0,\tag{1.105}$$
> a condition which leads to the so-called renormalization group equations (RGE).

The reason for calling equations of the form of Eq. (1.105) as renormalisation group equations will become clear very soon.

Let us study what are the consequences of this invariance under the choice of s_0 imposed by Eq. (1.105). For this, we note that $\mathcal{M}_{\rm NLO}(s)$ in Eq. (1.104) has both explicit dependence on s_0 as well as an implicit dependence via the renormalized coupling, since recall that from Eq. (1.68) we had found that λ_R depends on the bare coupling constant, the regulator Λ, and the reference scale s_0 as follows:

$$\lambda_R = \lambda + \frac{\lambda^2}{32\pi^2}\ln\left(-\frac{s_0}{\Lambda^2}\right) + \mathcal{O}(\lambda^3),\tag{1.106}$$

and therefore we need to take into account this implicit dependence on s_0 when we evaluate the total derivative in Eq. (1.105).

Taking into account both the explicit and the implicit dependences of Eq. (1.104) on the reference scale

s_0, we find that the invariance requirement Eq. (1.105) implies that

$$\frac{d}{ds_0} \mathcal{M}_{\mathrm{NLO}}(s) = \frac{d}{ds_0} \left(\lambda_R + \frac{\lambda_R^2}{32\pi^2} \ln \frac{s}{s_0} \right) = 0 \,, \tag{1.107}$$

and using the chain rule to account the implicit dependence of λ_R in s_0 we find that

$$\frac{d\lambda_R}{ds_0} \cdot \left(1 + \frac{\lambda_R}{16\pi^2} \ln \frac{s}{s_0} \right) - \frac{\lambda_R^2}{32\pi^2} \frac{1}{s_0} = 0 \,. \tag{1.108}$$

At this point we can do some rearranging and we obtain that the invariance of $\mathcal{M}_{\mathrm{NLO}}$ with respect to the arbitrary choice for s_0 leads to the following equation for the dependence of the renormalized coupling constant λ_R on s_0:

$$\frac{d\lambda_R}{ds_0} = \frac{\lambda_R^2}{32\pi^2} \frac{1}{s_0} \left(1 + \frac{\lambda_R}{16\pi^2} \ln \frac{s}{s_0} \right)^{-1} = \frac{\lambda_R^2}{32\pi^2} \frac{1}{s_0} + \mathcal{O}\left(\lambda_R^3 \right) \,, \tag{1.109}$$

where as usual we can neglect terms which are beyond the perturbative order that we are computing. Eq. (1.109) determines how λ_R must depend on s_0 in order that physical quantities such as the $\mathcal{M}_{\mathrm{NLO}}$ amplitude do not depend on this arbitrary choice. In other words, Eq. (1.109) fixes the *scale dependence* of the renormalized coupling constant. What we can learn from this exercise? The main take-away message is that

> The physical requirement that physical observables such as scattering amplitudes must be independent of the reference scale s_0 fixes the scale dependence of the renormalized coupling constant $\lambda_R(s)$.

Indeed, if we inspect the differential equation that we have obtained we see that

$$\frac{d\lambda_R}{ds_0} = \frac{\lambda_R^2}{32\pi^2} \frac{1}{s_0} + \mathcal{O}\left(\lambda_R^3 \right) \,, \tag{1.110}$$

it does not depend on any other scale but s_0 which in this context it is just a dummy scale. Therefore, solving this equation determines the full dependence of $\lambda_R(s)$ for any value of the scale s, subject to the appropriate boundary conditions.

Let us now solve this equation and study what we can learn from the scale dependence of $\lambda_R(s)$. The differential equation that we need to solve is:

$$\frac{d\lambda_R(s)}{ds} = \frac{\lambda_R^2(s)}{32\pi^2 s} \,, \tag{1.111}$$

subject to the boundary condition that at the reference scale s_0, one has that $\lambda_R(s = s_0)$ coincides with the value of $-\mathcal{M}(s = s_0)$ (since this is how the renormalized coupling constant is defined). The solution of this differential equation gives

$$\frac{1}{\lambda_R(s_0)} - \frac{1}{\lambda_R(s)} = \frac{1}{32\pi^2} \ln \frac{s}{s_0} \,, \tag{1.112}$$

which can be rearranged to give

$$\lambda_R(s) = \frac{\lambda_R(s_0)}{1 - \frac{\lambda_R(s_0)}{32\pi^2} \ln \frac{s}{s_0}} \,, \tag{1.113}$$

which is the sought-for expression of the scale dependence of $\lambda_R(s)$ as a function of $\lambda_R(s_0)$.

> The renormalized coupling $\lambda_R(s)$ is also known as the *running coupling*, given that its value is not
> fixed but rather it *runs* with the energy of the process.

Recall that $\lambda_R(s)$ coincides with the value of a physically observable quantity, since it was identified with a measurement of the $2 \to 2$ scattering amplitude. Since this dependence with the scale of $\lambda_R(s)$ (and thus on the associated two-point scattering amplitude) is determined entirely by the outcome of the renormalization procedure, Eq. (1.111) is called a *renormalization group equation* (RGE).

> The importance of RGEs such as Eq. (1.113) in quantum field theory arises because they allow us to
> relate the *physics of different scales*. In the case we are studying now, knowledge of the physics at
> the reference scale s_0 determines the physics at any other value of the energy s.

It is easy to show that the solution of the RGE equation for $\lambda_R(s)$, Eq. (1.113), can also be expanded in Taylor series to give the following result:

$$\lambda_R(s) = \lambda_R(s_0) \sum_{n=0}^{\infty} \left(\frac{\lambda_R(s_0)}{32\pi} \ln \frac{s}{s_0} \right)^n = \lambda_R(s_0) + \frac{\lambda_R^2(s_0)}{32\pi} \ln \frac{s}{s_0} + \dots , \tag{1.114}$$

which is useful to understand what the RGE is actually doing: one is resumming *to all orders* in the perturbative expansion of $\lambda_R(s)$ a specific class of terms, specifically those terms that are logarithmically enhanced in the energy s. To understand why this is quite powerful, not that with a one-loop calculation we have determined the complete tower of *leading logarithmic terms* to all orders in perturbation theory. On the other hand, the full expansion of $\lambda_R(s)$ will also contain terms of the form $\lambda_R^k (\ln s/s_0)^{k-2}$, and to determine these we would need to perform a two-loop calculation, and so on.

This is a generic feature of renormalisation group equations: they allow us to construct a more stable and robust perturbative expansion for the quantities that we are computing since they resum certain classes of enhanced higher-order terms to all orders in the expansion in powers of λ_R.

The derivative of the renormalized coupling with respect to the energy scale,

$$\frac{d\lambda_R(s)}{ds} \equiv \beta(\lambda_R) , \tag{1.115}$$

which determines the running of λ_R with the energy s, is called the *beta function* of the theory (for historical reasons). The beta function $\beta(\lambda_R)$ can be computed in perturbation theory, with each term providing a full tower of logarithmic terms to all orders. In general, quantum field theories will have beta functions $\beta_{\lambda_i}(\{\lambda_j\})$ associated to each of its coupling constants λ_i. The knowledge of these beta functions specifies the dependence with the scale of both coupling constants and physical observables of the theory. The first non-trivial term of the beta function in Eq. (1.115) requires a one-loop calculation, the second a two-loop calculation and so on.

Implications of the RGE for $\lambda_R(s)$. In the specific case of $\lambda\phi^4$ theory, the solution of the RGE associated to the scale dependence of its coupling constant, Eq. (1.113), indicates that $\lambda_R(s)$ increases with the energy,

in other words, perturbation theory will work better at low scales than at high scales. In this respect there are two interesting limiting behaviors of Eq. (1.113):

(a) For very small energies as compared to the reference value, $s \ll s_0$, then $\lambda_R(s) \to 0$ and the theory becomes effectively free (non-interacting). So in this limit all scattering cross-sections vanish.

> We say that in the limit $s \ll s_0$ the $\lambda\phi^4$ quantum field theory becomes *trivial*: all couplings vanish and therefore we end up with a theory of freely propagating fields.

(b) For very high energies as compared to the reference value, $s \gg s_0$, the coupling constant increases so perturbation theory will converge more slowly. If we keep increasing the energy, one finds that at some finite value s given by

$$s = s_0 \times \exp\left(\frac{32\pi^2}{\lambda_R(s_0)}\right), \tag{1.116}$$

the renormalized coupling becomes *infinite*. This is a sign that we are reaching the non-perturbative regime, where the perturbative description of the theory is not valid anymore.

> The value of the energy for which the renormalized coupling becomes infinite, Eq. (1.116), known as the *Landau pole* of the theory, and indicates the breakdown of the perturbative expansion, where non-perturbative dynamics set it.
>
> Note that in general QFTs can have their Landau poles either in the ultraviolet or in the infrared. For example, QED has its Landau pole in the UV (in the same was as $\lambda\phi^4$) while QCD instead has its Landau pole in the infrared.

Note however how perturbation theory ceases to be useful well before reaching the Landau pole. For a perturbative expansion to be convergent, the expansion parameter needs to satisfy $\lambda_R < 1$. Therefore, perturbative calculations in $\lambda\phi^4$ theory stop being trustable when

$$\lambda_R(s) \simeq 1 \quad \to \quad s \simeq s_0 \exp\left(-32\pi^2(1 - \lambda_R(s_0)^{-1})\right). \tag{1.117}$$

Recall that $\lambda_R(s_0)$ is a parameter fixed from the experiment and therefore not calculable from first principles. For here we see that the ratio s/s_0 between the maximum energy scale where one can use perturbation theory and the reference scale where λ_R is defined depends strongly on the value of $\lambda_R(s_0)$. If we start from a point where $\lambda_R(s_0) \ll 1$, when s can be much larger than s_0 and the theory still admit a perturbative description.

In Fig. 1.7 we show the one-loop running of the renormalized coupling constant $\lambda_R(s)$ in $\lambda\phi^4$ theory, Eq. (1.113), as a function of the scale ratio s/s_0. The reference value of the coupling constant has been taken to be $\lambda_R(s_0) = 1/2$. We observe how for $s \gg s_0$ the coupling constant increases, reaching eventually the Landau pole where $\lambda_R(s) \to \infty$, and for $s \ll s_0$ the coupling constant becomes actually smaller. This is one of the most striking features of quantum field theories: the strength of a given interaction is not just a fixed number, but inherently a function of the typical energy scale at which we are applying the theory. As we have discussed, this is a generic consequence of the renormalization procedure of the theory.

Coupling constant running
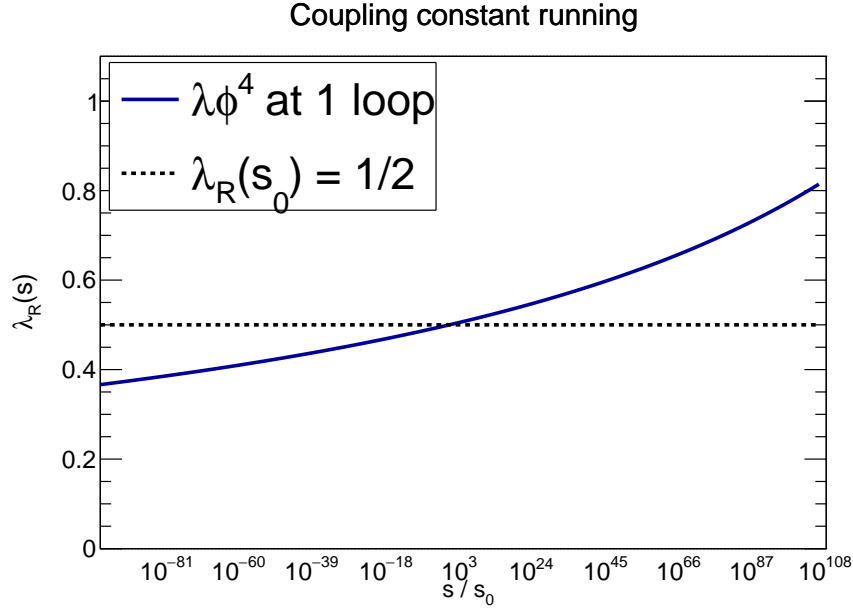
Figure 1.7. The one-loop running of the renormalized coupling constant $\lambda_R(s)$ in $\lambda\phi^4$ theory, Eq. (1.113), as a function of the scale ratio s/s_0. The reference value of the coupling constant has been taken to be $\lambda_R(s_0) = 1/2$.

From Fig. 1.7 we also see that the coupling in $\lambda\phi^4$ theory runs with the energy but in an extremely slow way: to increase the value of $\lambda_R(s)$ from its reference value of 0.5 up to 0.8 one needs to increase the energy by more than 100 orders of magnitude. This very slow running of the coupling constant of $\lambda\phi^4$ is also visible if we compute the value of the Landau pole of the theory, Eq. (1.116): for $\lambda_R(s_0) = 0.5$, we find that $s/s_0 \simeq 10^{274}$ - so we can be confident that this theory can always be treated with perturbation theory. Note that this is not a generic behaviour of all QFTs: some run much faster, and others, such as QCD, exhibit the inverse behaviour where the coupling becomes stronger as we decrease the energy scale s.

This slow running of the coupling in $\lambda\phi^4$ theory can be traced back to the small numerical factor, $1/32\pi^2$, that multiplies the logarithm in Eq. (1.113). RGE solutions with a larger value of this numerical coefficient would then run much faster than in the case of $\lambda\phi^4$ theory.

1.6 Formal derivation of the renormalization group equations

In the previous discussion, Sect. 1.5, we have found that imposing that physical observables, such as the $2 \to 2$ scattering amplitudes, should be independent of arbitrary choices such as the reference value of the energy s_0, one can derive renormalization group equations (RGEs) that *determine the energy dependence* of the renormalized coupling constants.

In this section we develop a more *general framework to construct RGEs*, based on the dimensional regularization method that has been used before to regularize loop scattering amplitudes in $\lambda\phi^4$ theory (see the discussion in Sect. 1.2). The key observation here is that the logarithms resummed by the Renormalisation Group can be associated with the ultraviolet divergences of the theory.

> Within dimensional regularisation, the renormalization of ultraviolet divergences introduces an *arbitrary scale* μ into physical predictions, which should cancel in the all-order calculation. The RGE equations will then arise from requiring the independent of physical observables on the arbitrary scale μ, which in general is referred to as the *renormalization scale*, at any given order in the perturbative expansion.

This is a more general case than the specific situation that we studied before, where the RGE for λ_R can constructed by demanding the independence of physical cross-sections with respect to the value of the arbitrary scale s_0.

As in the rest of this part of the course, we work with the $\lambda\phi^4$ theory defined by the Lagrangian Eq. (1.82). As we have demonstrated, the *bare parameters* in the Lagrangian are formally UV divergent in $d = 4$, recall Eq. (1.70) for the coupling constant and the fact that λ_R is a physical observable (and thus finite). We have shown that these infinities can be tamed using dimensional regularization, that is, by working in $d = 4 - \epsilon$ space-time dimensions. Recall that the mass dimensions of the fields and parameters that appear in the Lagrangian are in d space-time dimensions

$$[\phi] = \frac{d-2}{2}, \qquad [m] = 1, \qquad [\lambda] = 4 - d, \tag{1.118}$$

given that $[\mathcal{L}] = d$ and $[\partial_\mu] = 1$, since in natural units the action $S \sim \int d^4x \mathcal{L}$ is dimensionless. In particular, the coupling constant λ is not dimensionless anymore in $d \neq 4$ dimensions, but rather has mass dimensions of $4 - d$. This fact will be crucial in our general approach to construct renormalisation group equations.

In the renormalized theory, once we have introduced the counterterms as explained in Sect. 1.4, the Lagrangian is expressed in terms of physical parameters (related directly to measurable quantities) and of renormalized fields, and all dependence on the regulators has disappeared.

Formal derivation of RGEs

In order to formally derive general RGEs, we require two conditions:

- One needs to be able to expand any physical quantities in the renormalized theory in terms of a *dimensionless parameter*, to have a well-defined perturbative expansion.

  Note that one cannot perform a perturbative expansion in terms of a dimensionful parameter since then the size of the ratio between one term and the next one depends on the choice of units (is also a dimensionful number).

- One needs to ensure that the fields of the theory maintain their canonical dimension, irrespective of the specific value of that is being used d.

In order to ensure that these two requirements are satisfied, it is convenient to rescale the fields, masses, and coupling constants of the theory as follows:

$$\phi \to \sqrt{Z_\phi}\phi, \qquad m_p = \frac{1}{\sqrt{Z_1}}m, \qquad \lambda_R = \frac{1}{Z_4}\mu^{d-4}\lambda. \tag{1.119}$$

In Eq. (1.119), the scale μ has been introduced in order to keep the rescaled coupling λ_R dimensionless in any number of dimensions, and where the counterterms themselves, Z_ϕ, Z_1, and Z_4 are dimensionless. We have indicated as m_p the renormalized mass of the theory, which does not require to introduce μ since it already has the correct mass dimensions.

If we now express the Lagrangian of $\lambda\phi^4$ theory, Eq. (1.82),

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\,, \tag{1.120}$$

in terms of the renormalized fields, masses, and coupling constants as indicated in Eq. (1.119), we find that in d space-time dimensions the Lagrangian of our theory reads

$$\mathcal{L} = -\frac{1}{2}Z_\phi\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_p^2 Z_m\phi^2 - \frac{\lambda_R}{4!}Z_\lambda\mu^{4-d}\phi^4\,, \tag{1.121}$$

where we have defined the following counterterm combinations

$$Z_m \equiv Z_1 Z_\phi\,, \qquad Z_\lambda = Z_4 Z_\phi^2\,. \tag{1.122}$$

An important feature of the rescaled Lagrangian Eq. (1.121) is that both the physical coupling λ_R and the counterterms Z_i are *dimensionless* irrespective of the specific value of d. So the requirements of the method that we are trying to implement are indeed satisfied.

> If the renormalisation procedure has been successful, all physical observables computed from the Lagrangian Eq. (1.121) will be finite and independent of the regulator used. On the other hand, they will depend on the unphysical value of μ, the renormalisation scale of the theory.

Now, the crucial observation here is that scale μ in Eq. (1.121) is completely *arbitrary*: indeed, the original Lagrangian of the theory Eq. (1.82) did not exhibited any dependence on μ. Therefore, the *bare* parameters of the Lagrangian, such as the bare coupling constant λ and the mass m, *cannot depend on the choice* of the renormalization scale μ. This requirement leads to the following differential equation for the bare coupling constant of the theory:

$$\mu\frac{d}{d\mu}\lambda = 0\,, \tag{1.123}$$

which combined with the fact that the bare and physical couplings are related by

$$\lambda = Z_4\mu^{4-d}\lambda_R\,, \tag{1.124}$$

as indicated in Eq. (1.119), we have that

$$0 = \mu\frac{d}{d\mu}\lambda = \mu\frac{d}{d\mu}\left(Z_4\mu^{4-d}\lambda_R\right) = \mu^\epsilon Z_4\lambda_R\left[\epsilon + \frac{\mu}{\lambda_R}\frac{d\lambda_R}{d\mu} + \frac{\mu}{Z_4}\frac{dZ_4}{d\mu}\right]\,, \tag{1.125}$$

where we have used the chain rule of the derivative and accounted for the fact that both Z_4 and λ_R need to also depend on the renormalization scale μ. This is necessary since else it would not be possible to satisfy Eq. (1.123): the dependence on μ of λ_R follows directly from its definition, while the dependence of the

counterterm Z_4 is implicit via its dependence on the running coupling λ_R. In the last step we have used that $\epsilon = 4 - d$, in order to facilitate the connection with the physical limit $\epsilon \to 0$.

Therefore, we end up with a differential equation, Eq. (1.125), that determines how the renormalized coupling of the theory *runs with the energy scale*, and can be solved order by order in perturbation theory. We have derived this equation only from requiring that the bare parameters of the Lagrangian do not depend on the value of the arbitrary renormalisation scale μ. Since the counterterm Z_4 is known for the renormalized theory, then the RGE associated to λ_R and that determines its scale dependence will be given by

$$\frac{1}{\lambda_R} \frac{d\lambda_R}{d \ln \mu} = -\epsilon + \frac{\mu}{Z_4} \frac{dZ_4}{d\mu} \,. \tag{1.126}$$

For instance, at leading order in the perturbative expansion (Born approximation) we have that

$$Z_4 = Z_\lambda Z_\phi^{-2} = 1 + \mathcal{O}(\lambda_R) \quad \rightarrow \quad \left(\frac{dZ_4}{d\mu} \right)_0 = 0 \,, \tag{1.127}$$

and then from Eq. (1.125) we find that

$$\mu \left(\frac{d\lambda_R}{d\mu} \right)_0 = -\epsilon \lambda_R \,. \tag{1.128}$$

which determines the running of the coupling with the scale μ at leading order (Born level) in $d = 4 - \epsilon$ space-time dimensions. Consistently, we find that if $\epsilon = 0$ then λ_R is scale invariant at tree level, as we know from the explicit computation of scattering diagrams. It is remarkable to note that in $\lambda\phi^4$ theory the coupling constant runs even at leading order for $d \neq 4$. In other words, the *beta function* of the theory is different from zero already at the Born level for $d \neq 4$.

At the one-loop level (NLO), we have already derived the explicit $\mathcal{O}(\lambda_R)$ expressions for the three counterterms of the theory, from explicit computations of two- and four-point functions using dimensional regularization. These three counterterms are associated to the renormalisation of the scalar field, its mass, and its coupling constant respectively. The results were the following:

$$\begin{aligned}
Z_\phi(\lambda_R, \epsilon) &= 1 + \mathcal{O}\left(\lambda_R^3\right) \,, \\
Z_\lambda(\lambda_R, \epsilon) &= 1 + \frac{3\lambda_R}{16\pi^2} \left(\frac{1}{\epsilon} + \text{constant} \right) + \mathcal{O}\left(\lambda_R^2\right) \,, \\
Z_4(\lambda_R, \epsilon) &= 1 + \frac{3\lambda_R}{16\pi^2} \frac{1}{\epsilon} + \mathcal{O}\left(\lambda_R^2\right) \,,
\end{aligned} \tag{1.129}$$

Therefore, at NLO the renormalization group equation for the counterterm Z_4 reads, taking into account that its only dependence on the renormalisation scale μ is via its implicit dependence on the running coupling λ_R,

$$\mu \left(\frac{dZ_4}{d\mu} \right)_{\text{NLO}} = \frac{3}{16\pi^2} \left(\mu \frac{d\lambda_R}{d\mu} \right) \frac{1}{\epsilon} + \mathcal{O}\left(\lambda_R^2\right) = -\frac{3}{16\pi^2} \lambda_R + \mathcal{O}\left(\lambda_R^2\right) \,, \tag{1.130}$$

where we have used Eq. (1.128) that gives the scale dependence of μ_R at the Born level. Next, we can substitute this result into Eq. (1.125) and doing some algebra leads to the RGE equation for the renormalized

coupling constant. We first that at the one-loop order, this RGE is given by

$$\mu \frac{d\lambda_R}{d\mu} = \frac{3}{16\pi^2} \lambda_R^2 \equiv \beta(\lambda_R) \,, \tag{1.131}$$

where now we have been able to take the $\epsilon \to 0$ limit since we had the next term in the perturbative expansion. Therefore, we find that we can recover the previous result in terms of the beta function of the $\lambda\phi^4$ theory at the first non-trivial order, recall Eq. (1.112).

> Note that this derivation has not required any explicit calculation, provided the *perturbative expansion of the counterterms* is known at any fixed perturbative order: the knowledge of the renormalized Lagrangian is all we need in order to determine the renormalisation group equations for the couplings and masses of the theory.

This result highlights that this new method is rather more general and powerful than the one used in the previous section: no additional calculations are required provided that the structure of the renormalized Lagrangian is known.

Given the structure of the counterterm $Z_4 = Z_\lambda Z_\phi^{-2}$, in principle the beta function of the theory receives contributions from many different diagrams, in particular from all those that are required to evaluate Z_λ and Z_ϕ. For this reason, building the RGEs from the bare Lagrangian is a more systematic way than looking at explicit observables. On the other hand, the calculation of the beta function in terms of physical observables can be more intuitive, so in this respect the two methods are *complementary* to each other.

Using a similar strategy, we can derive RGEs for other parameters of the Lagrangian and determine their dependence with the unphysical renormalisation scale (which is a necessary byproduct of the renormalisation procedure). Specifically, we can determine the RGEs corresponding to the physical mass m_p and to the renormalized scalar field ϕ. From Eq. (1.119) we have that

$$
\begin{aligned}
m &= \sqrt{Z_1} m_p = \sqrt{Z_m Z_\phi^{-1}} m_p \,, \\
\phi_0 &= \sqrt{Z_\phi} \phi \,,
\end{aligned}
\tag{1.132}
$$

where we denote by ϕ_0 is the bare field and ϕ the renormalized one. To derive the RGE equations, we impose the physical requirement that the bare mass m and scalar field ϕ_0 cannot depend on the arbitrary renormalization scale μ, since it was absent in the original Lagrangian of the theory. That is, we need to impose the same conditions as was done in Eq. (1.123) for the case of the bare coupling but now for the bare fields and masses. Therefore we have that:

$$\mu \frac{d}{d\mu} m = 0 \,, \qquad \mu \frac{d}{d\mu} \phi_0 = 0 \,, \tag{1.133}$$

which lead to two differential equations that relate the Z_i counterterms to the renormalized mass m_p and field ϕ. At this point, it becomes convenient to define the following quantity:

$$\gamma_m \equiv \frac{\mu}{m_p} \frac{dm_p}{d\mu} \,, \tag{1.134}$$

which is called the *anomalous dimension* of the mass, and

$$\gamma_\phi \equiv \frac{1}{2} \frac{\mu}{Z_\phi} \frac{dZ_\phi}{d\mu} \, , \tag{1.135}$$

which is known as the *anomalous dimension* of the scalar field. The reason of these names will become apparent in short. In the case of the physical mass m_p, its anomalous dimension γ_m determines its running with the energy, analogously to the beta function in Eq. (1.131) which fixes the running of the renormalized coupling λ_R with the scale.

Let us here work out explicitly the renormalization group equation for the physical mass m_p. Taking into account that the mass counterterm Z_m is given, see Eq. (1.102), by

$$Z_m = 1 + \frac{\lambda_R}{32\pi^2} \ln \left(\frac{m_p^2}{\Lambda^2} \right) + \mathcal{O}\left(\lambda_R^2 \right) \, , \tag{1.136}$$

which combined with the result for the other counterterms, Eq. (1.129), leads to the following relation between the physical and renormalized masses:

$$m = \sqrt{Z_m Z_\phi^{-1}} \, m_p \simeq \left(1 + \frac{\lambda_R}{64\pi^2} \ln \left(\frac{m_p^2}{\Lambda^2} \right) + \mathcal{O}\left(\lambda_R^2 \right) \right) \times m_p \, , \tag{1.137}$$

and therefore by imposing the invariance condition Eq. (1.133) using Eq. (1.137) we can compute the anomalous dimension γ_m of the physical mass, which determines its running with the energy. As mentioned above, this is the analogous relation of Eq. (1.131) for the renormalized coupling. By doing this we will have

$$\mu \frac{dm}{d\mu} = 0 = \mu \frac{d}{d\mu} \left[\left(1 + \frac{\lambda_R}{64\pi^2} \ln \left(\frac{m_p^2}{\Lambda^2} \right) + \mathcal{O}\left(\lambda_R^2 \right) \right) \times m_p \right] \, , \tag{1.138}$$

which upon some rearrangement leads to

$$\left(\mu \frac{d\lambda_R}{d\mu} \right) \ln \left(\frac{m_p^2}{\Lambda^2} \right) + \lambda_R \left(\frac{\mu}{m_p} \frac{dm_p}{d\mu} \right) \left[2 + \ln \left(\frac{m_p^2}{\Lambda^2} \right) \right] = 0 \, , \tag{1.139}$$

where we have used that for the renormalized coupling λ_R and physical mass m_p will depend on the renormalisation scale μ.

This equation relates the beta function of the renormalized coupling with the anomalous mass dimension γ_m determining the scale dependence of the physical mass, Eq. (1.134):

$$\beta(\lambda_R) \ln \left(\frac{m_p^2}{\Lambda^2} \right) + \lambda_R \gamma_m(\lambda_R) \left[2 + \ln \left(\frac{m_p^2}{\Lambda^2} \right) \right] = 0 \, , \tag{1.140}$$

Putting everything together, we find that in $\lambda \phi^4$ theory the mass anomalous dimension is given by

$$\gamma_m(\lambda_R) = \frac{\mu}{m_p} \frac{dm_p}{d\mu} = -\frac{\frac{3}{16\pi^2} \lambda_R \ln \left(\frac{m_p^2}{\Lambda^2} \right)}{2 + \ln \left(\frac{m_p^2}{\Lambda^2} \right)} \rightarrow \frac{3}{16\pi^2} \lambda_R \, , \tag{1.141}$$

where we have inserted the value of $\beta(\lambda_R)$ at one-loop, and where in the last step we have taken the physical limit $\epsilon \to 0$, which is equivalent to $\Lambda \to \infty$. Eq. (1.141) is the sought-for result of the mass anomalous dimension in $\lambda \phi^4$ theory, which is indeed finite and free of any dependence on the regulator, as expected for a renormalizable theory.

To summarize, in this section we have shown how the renormalization group equations are powerful tools, basic on basic invariance requirements, which allows one to resum contributions to all orders in perturbation theory and to infer the running of the masses and the couplings with the scale for any given theory. As such they are ubiquitous in many applications of QFT from particle physics to condensed matter systems. We have also presented a method that allows constructing general RGEs without the need of explicit calculations, based only on the structure of the renormalized Lagrangian of the theory.

2 Effective Field Theories and the all-order behaviour of QFTs

Up to know, we have presented explicit calculations of renormalised QFT at the one loop level. We will discuss now what we can say from general principles about the all-order behaviour of QFT amplitudes. In doing so, we will also introduce the ideas underlying Effective Field Theories (EFTs) and their relation with UV-complete renormalizable EFTs. These ideas are an essential component of the modern understanding of QFTs.

2.1 The EFT philosophy

The basic idea underlying the concept of effective theories is that in a wide range (but not all) of situations, the physics that determines the behaviour of a system a given length or energy scale is decoupled from the physics that determines the behaviour of the same system at very different scales.

Some examples of this separation of scales include:

- We don't need to know that Quantum Chromodynamics is the correct theory of the strong nuclear force (or even to know the existence of the strong nuclear force itself) to compute with high precision the electronic transitions in atoms.

  This property is a consequence of the fact that electronic transitions in atoms take place at length scales of $r_{\rm atom} \simeq 10^{-10}$ m, while the effects of the strong force arise only at much smaller scales $r_{\rm nucl} \simeq 10^{-15}$ m.

- We can compute with good accuracy the properties of the β radioactive decay of the neutron, $n \to p + e^+ + \nu_e$ without knowing about the existence of the W, Z bosons, or even about the quark substructure of hadrons for that matter.

- We can predict the motion of planets and solar eclipses with extremely high precision using only classical mechanics, which is only an approximation to the full theory, Special Relativity, valid when $v \ll c$.

In summary, the principle of separation of scales can be described as follows:

> **Principle of separation of scales**
>
> The physics that take place at a given distance or energy scale should be describable without the need of reference to the physics that take place at widely separated distance or length scales.

An Effective Field Theory (EFT) is one of the possible realizations of this concept, valid in Quantum Field Theories (QFT) that involve widely separated mass or energy scales. For example, consider a QFT composed by two scalar fields, a *light* field ϕ with mass m_ϕ and a *heavy* field ψ with mass m_ψ, so that $m_\phi \ll m_\psi$. Depending on the energy E involved in the process, for example in a scattering amplitude, we will have two different regimes as illustrated by the drawing below:

- For $E \gtrsim m_\psi$, both fields are dynamical degrees of freedom, and in particular appear in the initial and final states of scattering amplitudes. We denote this as the *full theory* or the *UV-complete theory*.
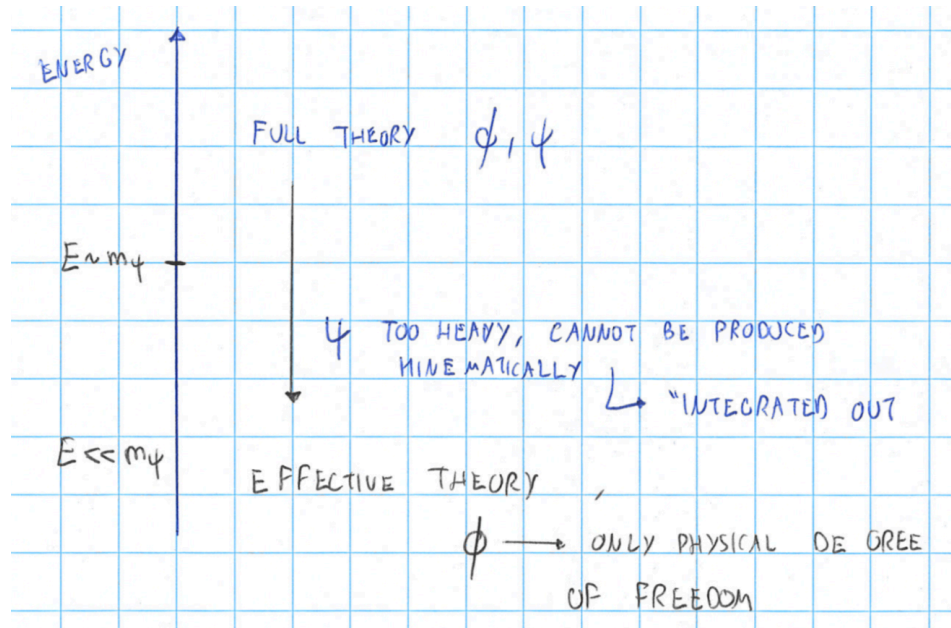
Figure 2.1. Separation of scales in a QFT composed by two scalar fields with a wide hierarchy of masses among them.

- For $E \ll m_\psi$ the heavy field is *not dynamical*: for example it cannot be kinematically produced in the final state of a scattering process. We say that this field can then be *integrated out* of the theory, to reflect that at low energies the only dynamical degree of freedom is the light field ϕ. We call the theory where ψ has been integrated out as the *effective theory*.

Note that the full and the effective theory are not different theories: they are alternative formulations of the same physical dynamics. One important advantage of the EFT framework is that by concentrating on the physical degrees of freedom relevant at low energy one can in general significantly simplify calculations. Furthermore, in general one does not need to know the full theory to construct its effective version: EFTs can be constructed based on general bottom-up considerations without the requirement to refer to a full theory. If the full theory is known, then the parameters such as coupling constants in the EFT can be determined by means of a procedure denotes as *matching*.

2.2   The Fermi EFT of the weak interactions

Let us illustrate the main ideas of the EFT framework by means of Fermi Theory, perhaps the paradigmatic example of an effective theory. When originally formulated, it was a phenomenological model to describe radioactive processes such as the beta decay of the neutron,

$$n \to p + e^- + \bar{\nu}_e \,. \tag{2.1}$$

At the Lagrangian level, this process takes place via a four-fermion interaction of the following form

$$\mathcal{L} = \frac{G_F}{\sqrt{2}} \left(\bar{\psi}_e \gamma^\mu (1 + \gamma_5) \psi_\nu \right) \left(\bar{\psi}_n \gamma^\mu (1 + \gamma_5) \psi_p \right) \,. \tag{2.2}$$

Note that proton and neutrons appear here as elementary fermions, since their quark substructure was unknown by the time the theory was formulated. The coupling constant of the theory G_F, known as Fermi constant, is not dimensionless but rather dimensionful, $[G_F] = -2$, whose value fixed from the experiment is $G_F = 1.166 \times 10^{-5}$ GeV^{-2}.

This theory seemed to describe pretty well all available data from neutron beta decays. However, from the conceptual point of view there remained a number of important questions, such as the mechanism that determined the value of the Fermi constant. In addition to this conceptual conundrum, there was also another concern from the practical point of view, related to the behaviour of scattering amplitudes at high energies. Basic crossing considerations in QFT indicate that if the decay $n \to p + e^- + \bar{\nu}_e$ was allowed by the theory, then the scattering process $\nu_e + n \to e^- + p$ should also be allowed, as depicted below.
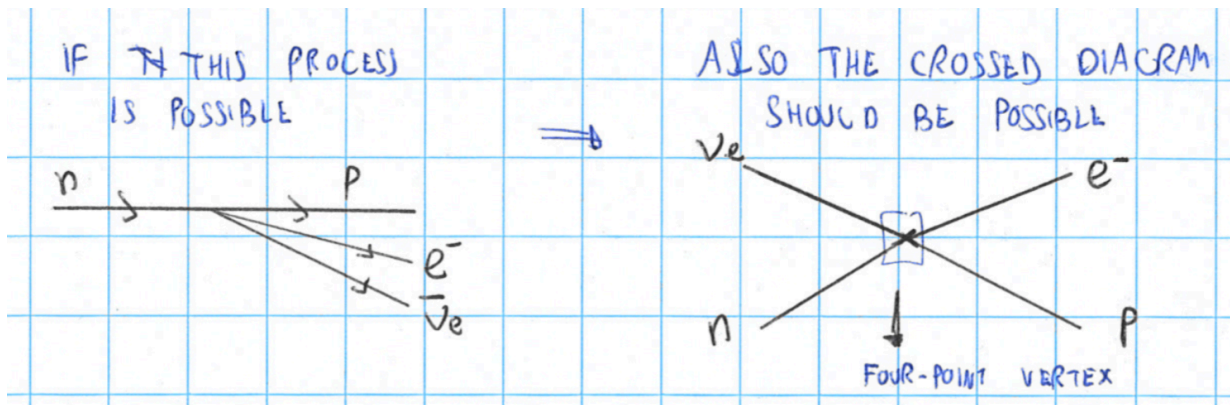
Figure 2.2.   The existence of neutron's beta decay implied that Fermi theory should also be able to describe neutrino-neutron scattering.

Based on dimensional analysis, it is easy to see that the cross-section for $\nu_e + n \to e^- + p$ scattering should scale as $\sigma \propto G_F^2$, given that the amplitude is proportional to $\mathcal{A} \propto G_F$. Now remembering that a cross-section is nothing but an area, its mass-dimensions are $[\sigma] = -2$ and thus purely based on dimensional reasons we can derive that at large energies the scattering cross-section should scale as

$$\sigma \left(\nu_e + n \to e^- + p \right) \propto G_F^2 E_{\text{cm}}^2 \,, \tag{2.3}$$

where E_{cm} is the center of mass energy, and we assume that E_{cm} is large enough that all masses involved can be neglected (and thus E_{cm} is the only magnitude in the problem with the right units to obtain dimensionally consistent expressions).

Eq. (2.3) indicates that the scattering cross-section grows quadratically with the energy. But this is a serious problem: if we go to high enough energies the theory will violate unitarity, meaning that the probability for this scattering process will become larger than one. Therefore this result indicates that there is something weird with Fermi theory: it seems unlikely that it is a fundamental theory valid at all energies. On the other hand, it describes very well the physics around the proton and neutron mass scales, so one would not want to throw it away entirely.

Can we estimate the regime of validity of Fermi theory? Assume that we can compute one loop corrections in the theory. From dimensional considerations, we have that the cross-section once we include one-loop

corrections should be given by

$$\sigma \left(\nu_e + n \rightarrow e^- + p \right) = c_1 G_F^2 E_{\text{cm}}^2 + c_2 G_F^3 E_{\text{cm}}^4 + \dots \,, \tag{2.4}$$

where c_1, c_2 are dimensionless numbers, and where the second term arises from the interference between the Born and one-loop scattering amplitudes. For this expansion to have any sense, the one-loop term should be smaller than the Born contribution, and this means that $G_F E_{\text{cm}}^2 \lesssim 1$, which implies in turn that $E_{\text{cm}} \lesssim G_F^{-1/2} \simeq 300$ GeV: the theory cannot really be used around and above that scale.

So this rather succinct inspection of Fermi theory has lead us to two main questions:

- What mechanism determines the value of G_F?

- Is it possible to make sense of the high energy behaviour of Fermi theory for $E_{\text{cm}} \gtrsim G_F^{-1/2} \simeq 300$ GeV, given that perturbative expansion seems to break down there?

Of course, we know very well the answers to these two questions: Fermi theory is not a fundamental but rather an effective theory, whose regime of validity is limited by some energy cutoff Λ. And we also know that the corresponding fundamental theory is nothing but the Standard Model, and that $\Lambda \simeq M_W$ is the mass of the W boson that in Fermi theory has been integrated out, as represented schematically below. In the full theory, the Standard Model, the scattering process $\nu_e + n \rightarrow e^- + p$ proceeds via the exchange of a W^+ gauge boson. However at low energies the W boson propagator collapses to a single point, and therefore one ends up with the four-fermion interaction characteristic of Fermi theory.
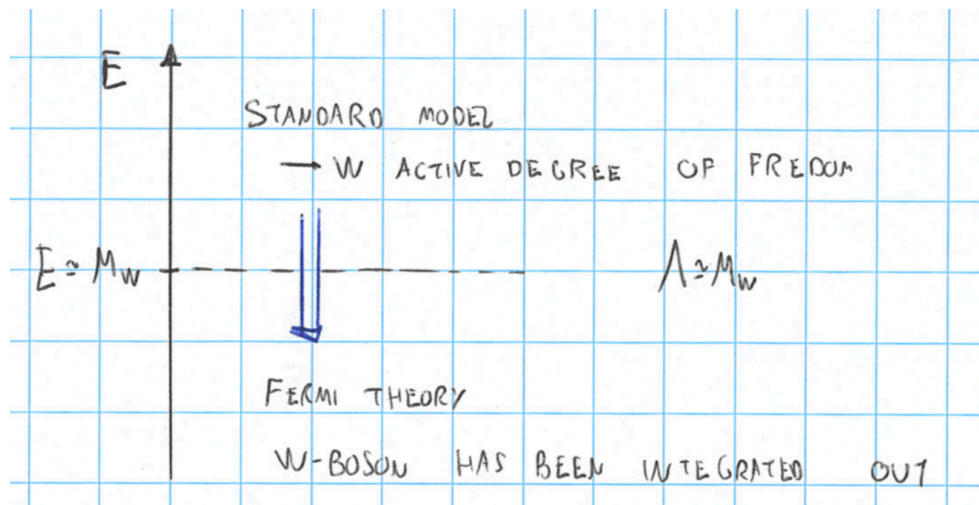
Figure 2.3. Fermi Theory the low-energy EFT of the weak interactions of the Standard Model.

We now have the answer to one of the questions that were raised above: the theory is well-behaved at large energies since there one has only renormalizable interactions where the W boson couples to two fermions. To answer the other question about Fermi theory, namely what determines the value of G_F, we need to introduce an important concept of the EFT framework.

EFT Matching

By *matching* we denote the mechanism whereby some of the parameters of the effective theory are computed in terms of those of the full theory. The matching procedure can be carried out in different ways, for example by computing scattering amplitudes fir the same process in the full and in the effective theory and then imposing that the two amplitudes are equivalent.

Let us illustrate how the concept of matching works in the case of Fermi theory. As we have mentioned above, in terms of couplings the scattering amplitude will be proportional to

$$\mathcal{A}_{\text{fermi}}(\nu_e + n \to e^- + p) \propto G_F, \tag{2.5}$$

while in the full theory, the Standard Model, as indicated by the diagram below the scattering amplitude
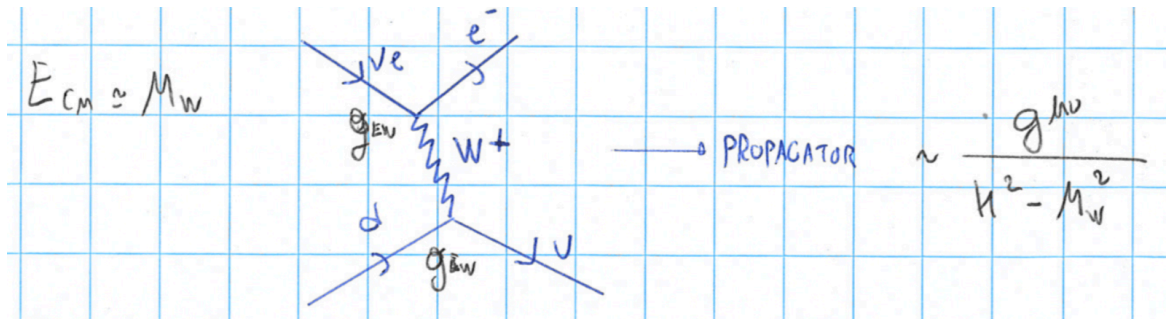
Figure 2.4.

for the same process will be proportional to the square of the electroweak $f\bar{f}W$ coupling, g_{EW}, and to the W boson propagator

$$\mathcal{A}_{\text{SM}}(\nu_e + n \to e^- + p) \propto \frac{g_{\text{EW}}^2}{k_W^2 - M_W^2}. \tag{2.6}$$

In general $\mathcal{A}_{\text{fermi}}$ and $\mathcal{A}_{\text{SM}}$ will be different, since the two theories are not valid in the same kinematic range (in particular the EFT is only valid at low energies). But in the region of validity of Fermi theory, $E \ll M_W$, then the two amplitudes should indeed be identical since they correspond to the same process. So in the regime of validity of the EFT the amplitude in the full theory reduces to

$$\mathcal{A}_{\text{SM}}(\nu_e + n \to e^- + p) \to \propto \frac{g_{\text{EW}}^2}{M_W^2}, \tag{2.7}$$

since all momentum transfers are well below the cutoff, $k^2 \ll M_W^2$. By imposing the equivalence between $\mathcal{A}_{\text{fermi}}$ and $\mathcal{A}_{\text{SM}}$ in the appropriate kinematical regime we thus find the *matching conditions* of the theory

$$G_F \propto \frac{g_{EW}^2}{M_W^2}, \tag{2.8}$$

whereby the parameters of the EFT (l.h.s.) are expressed in terms of the couplings and masses of integrated-out particles of the full theory (r.h.s.). Note also that for $g_{\text{EW}} = \mathcal{O}(1)$ we see that $M_W = \mathcal{O}\left(G_F^{-1/2}\right)$, in agreement with our previous estimate of the EFT cutoff scale Λ. The relation in Eq. (2.8) illustrates a

crucial feature of EFTs: the information about the integrated-out particles (in this case the W boson: its coupling to fermions g_{EW} and its mass) are still present in the low-energy EFT by means of its parameters.

2.3   An scalar EFT example

The same considerations that we have drawn concerning Fermi Theory can be extended to other EFTs. In one of the exercises of the first tutorial session we will consider a renormalizable theory composed by two scalar fields and defined by the Lagrangian:

$$\mathcal{L}_{\mathrm{full}} = \mathcal{L}_{\mathrm{kin}} + \mathcal{L}_{\mathrm{int}} \,, \tag{2.9}$$

where the kinetic term is given by

$$\mathcal{L}_{\mathrm{kin}} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 + \frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi - \frac{1}{2}M^2\Phi^2 \,, \tag{2.10}$$

and with $\mathcal{L}_{\mathrm{int}}$ indicating the interaction term. We will there construct an EFT valid in the region where $m \ll M$, and show how the EFT couplings can be expressed in terms of the masses and couplings of the full theory by means of the matching procedure. In general, the Lagrangian corresponding to this EFT will be written as

$$\mathcal{L}_{\mathrm{eft}} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 + c_4\phi^4 + c_6\phi^6 + c_8\phi^8 + \dots \,, \tag{2.11}$$

where the coefficients c_i are computed by matching scattering amplitudes in the full and effective theories (and can also vanish, depending on in the interactions in the full theory).

For example, let us assume that the interaction term in the full theory has the form

$$\mathcal{L}_{\mathrm{int}} = \rho\phi^3\Phi \,, \tag{2.12}$$

where ρ a dimensionless coupling. Can we then determine what will be the corresponding interactions in the EFT? We can consider the six-point scattering amplitude in the full and effective theories. In the full theory, see Fig. 2.5, the scattering amplitude will scale as

$$\mathcal{M}_{\mathrm{full}}\left(\phi\phi\phi \to \phi\phi\phi\right) \simeq \frac{\rho^2}{k^2 + M^2} \,, \tag{2.13}$$

while in the corresponding EFT description the same scattering amplitude reads

$$\mathcal{M}_{\mathrm{full}}\left(\phi\phi\phi \to \phi\phi\phi\right) \simeq c_6 \,, \tag{2.14}$$

since it is generated by the operator $c_6\phi^6$. Requiring that the two amplitudes coincide in the low-energy limit where the EFT description is valid, one finds that the matching conditions imply

$$c_6 \propto \frac{\rho^2}{M^2} \,, \tag{2.15}$$

Full Theory   **Effective Theory**
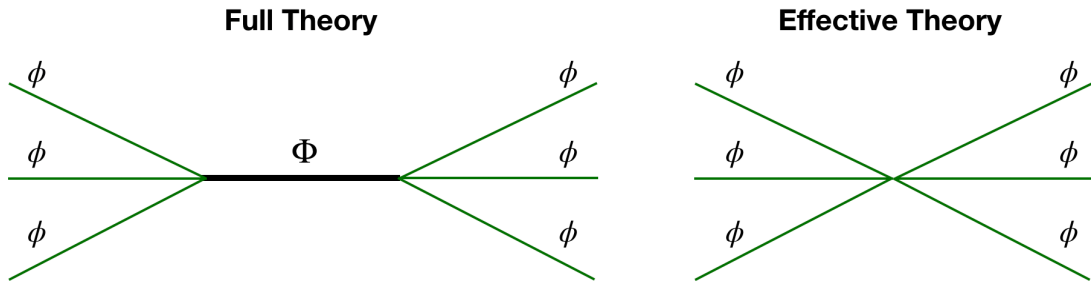
Figure 2.5. Contribution to the six-point scattering amplitude in a QFT composed with two scalar fields coupled via an interaction term $\mathcal{L}_{\text{int}} = \rho\phi^3\Phi$, and its counterpart in the effective theory where the heavy field Φ has been integrated out. Note that the latter description is only valid when $E \ll M$.

and thus in this case we can write the EFT Lagrangian entirely in terms of the parameters of the full theory

$$\mathcal{L}_{\text{eft}} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 + a\frac{\rho^2}{M^2}\phi^6\,, \tag{2.16}$$

with a being an $\mathcal{O}(1)$ coefficient. A similar derivation can be carried out for other forms of the full-theory interaction term $\mathcal{L}_{\text{int}}$.

2.4   General renormalizability conditions of QFTs

In Sect. 1 we have explicitly demonstrated how the $\lambda\phi^4$ theory can be renormalised at the one loop level. We have found that if we express the four-point $\mathcal{M}(\phi\phi \to \phi\phi)$ scattering amplitude in terms of the physical coupling λ_R, the dependence on the cutoff Λ disappears (or equivalently on the regulator ϵ if one is using dimensional regularisation), rendering all scattering amplitude finite. A similar analysis can be performed at higher-orders, and eventually demonstrated for all perturbative orders.

> In *renormalizable theories* all physical observables can be rendered well-defined by introducing a *finite number of counterterms*. Other theories don't satisfy this requirement, and go under the name of non-renormalizable QFTs. For example, the naive quantization of gravity leads to a QFT requiring an infinite number of counterterms and thus losing any predictive character.

It is possible to determine from first principles whether of not a given QFT is renormalizable, or this can only be found by trial an error? If we manage to renormalize a quantum field theory at one loop, which guarantee do we have that further divergences will not appear when computing higher orders? Note that non-renormalizable QFT they can still provide sensible predictions as an effective low energy theory, provided we don't extrapolate to large energies above its regime of validity.

In the following, we make use of *dimensional analysis* and *power counting* to provide some arguments about when a quantum field theory is renormalizable or not, and try to assess what is the degree of divergence of an arbitrary high-order scattering diagrams. The idea is to understand how many counter-terms are required to render physical observables finite. For instance, in massless $\lambda\phi^4$ theory after renormalisation,

defined by the Lagrangian

$$\mathcal{L} = -\frac{1}{2} Z_\phi \partial_\mu \phi \partial^\mu \phi - \frac{\lambda_R}{4!} Z_\lambda \, \phi^4 \tag{2.17}$$

is enough the presence of Z_ϕ and Z_λ to render all possible physical observables finite? Or would we need additional counter-terms in the Lagrangian? If additional divergences appear at higher orders requiring new counterterms to regulate them, this would indicate that the theory is not renormalizable. In this case we already know the answer, but it is instructive to find the same answer from different methods. As we will show, it is possible to (mostly) answer this question without actually computing a single diagram, just from general theory considerations about the structure of the perturbative expansion.

Dimensional analysis. The interactions between quantum fields can be classified into three different types, depending on the *mass dimensions* of the relevant coupling constant:

- *Relevant interactions*, when the coupling constant has positive mass dimensions $[g] > 0$.

  An example of this is the mass term, $-m^2\phi^2/2$, where the "coupling" has mass dimensions of $[g] = 2$.

- *Marginal interactions*, when the coupling constant is dimensionless $[g] = 0$.

  This is the case for example of the quartic interaction in $\lambda\phi^4$ theory in $d = 4$ dimensions.

- *Irrelevant interactions*, when the coupling constant has negative mass dimensions $[g] < 0$.

  An example of this is Fermi's theory of beta decay, whose Lagrangian includes a four-fermion interaction term of the form

  $$\mathcal{L}_F \subset G_F \bar{\psi}\psi\bar{\psi}\psi \,, \tag{2.18}$$

  and thus a massive coupling constant with negative mass dimensions.

Here the symbol $[\;]$ indicates the mass dimensions of given quantity, for example $[E] = 1$ and $[x] = -1$.

The renormalizability or not of a theory depends on the types of operators that it includes, namely:

- A *renormalizable theory* includes only *relevant* or *marginal* operators.

  This implies that a finite set of counterterms if enough to remove all the UV divergences.

- A *non-renormalizable theory* contains at least *one irrelevant operator*.

  This implies that to remove all infinities it would be necessary to include an infinite number counterterms.

Let us now give a bit more motivation to these statements.

Power counting. We have seen in our study of loop corrections to the scattering processes in $\lambda\phi^4$ theory that at the one-loop level we can absorb all the ultraviolet divergences into a redefinition of the parameters of the Lagrangian. In the case of the massless theory at one loop, expressing the results of the calculation in terms of the physical coupling λ_R renders all results finite. However, this does not automatically imply that the theory is renormalizable, since we have considered only a subset of all the possible higher-order diagrams.
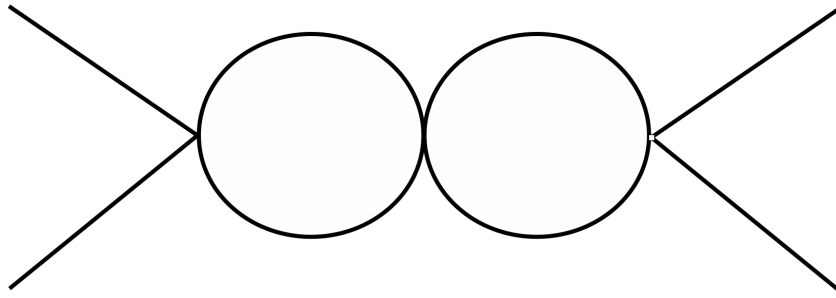
Figure 2.6. Feynman diagram corresponding to a two-loop calculation to the $\phi\phi \to \phi\phi$ scattering amplitude in $\lambda\phi^4$ theory.

> But what about other diagrams that might also exhibit UV divergences, for instance higher-order diagrams? Do we need to include additional modifications to the Lagrangian beyond those already added? How can one ensure that if we absorb all infinities in the Lagrangian renormalization at the one-loop level, we will not need to introduce additional counterterms to deal with higher-loop diagrams?

As a related question, how can we determine if a Feynman diagram is divergent and how bad this divergence is? Is there a general way to answer this question or it is only by explicit calculations that we can determine the answer?

Let us consider a generic Feynman diagram for an scattering amplitude. Such diagram will be characterized by a number of characteristic features, being in particular the most important ones:

- V: the number of interaction vertices.

- E: the number of external lines.

- I: the number of internal lines (propagators).

- L: the number of undetermined momenta, which is equivalent to the number of loops.

For simplicity, we assume that this diagram contains only scalar fields. We note that the internal momenta satisfy $(V - 1)$ relations among themselves, from the V vertices where momentum conservation is imposed and where where the -1 terms arises from the global conservation of four momenta. Therefore one has the following relation

$$L = I - (V - 1) \ . \tag{2.19}$$

In other words, the number of loops in the diagram is the same as the number of undetermined momenta, and this is equal to the number of internal lines minus the $V - 1$ constraints that arise in the interaction vertices from four-momentum conservation. Note that the -1 in the $(V - 1)$ term reflects that fact that, for the first interaction vertex, local and global momentum conservation are the same (as can be seen for example from the Born scattering amplitude). Note that this equation does not depend on E, the number of external lines in the diagram.

As we will shown, the usefulness of this relation is that it enables us to compute the naive counting of powers of k^μ in a given diagram. Let us give three examples showing how this relation holds:

- In the one-loop Feynman diagrams of Fig. 1.3, one has $L = 1$ (one loop), $V = 2$ (two interaction vertices), and $I = 2$ (two internal lines), so indeed Eq. (2.19) holds.

- In the first diagram for the two-loop correction of the propagator, shown in Fig. 1.4. we have $I = 3$ internal lines (propagators), $V = 2$ interaction vertices, $E = 2$ external lines. So applying Eq. (2.19) leads to $L = 3 - (2 - 1) = 2$, corresponding to the number of loops in this diagram.

- In Fig 2.6 we show one of the Feynman diagram that contribute to the two-loop calculation to the $\phi\phi \rightarrow \phi\phi$ scattering amplitude in $\lambda\phi^4$ theory.

  In this diagram we have three interaction vertices ($V = 3$), four external lines ($E = 4$), four internal lines ($I = 4$), and two loops ($L = 2$), which correspond to the two undetermined momenta running through the diagram. Therefore, according to equation Eq. (2.19), the following relation should be satisfied:

  $$L = 2 = 4 - (3 - 1) = 2 \,, \tag{2.20}$$

  so again here this equation is validated.

Note that the importance of relations of the form of Eq. (2.19) is that they are valid for *any Feynman diagram*, in particular they are valid for any number of loops. So if we can say anything about our theory based on this kind of general considerations, it will be very useful to understand its all-order behaviour as opposed to being restricted to explicit one-loop calculations as we have done up to now.

Next, one defines D as the *superficial (apparent) degree of divergence* of a given Feynman diagram. To do this, we note that:

- Each loop that appears in a diagram involves an integral over an undetermined momenta in d-dimensions, $\sim \int d^d k$

- Each internal line involves a propagator with an inverse power of the four-momenta, $\sim k^{-2}$ (recall that here we are working in the massless approximation).

You can convince yourselves that the superficial degree of divergence D of a Feynman diagram is given by:

$$D = d \times L - 2 \times I \,, \tag{2.21}$$

if we take into account the following considerations:

- the number of loops times the number of space-time dimensions d since each loop integration leads to d powers of k (given that the integration measure is $d^d k$);

- minus twice the number of propagators (internal lines), since each propagator contributes with two negative powers of k to the total scaling of the diagram.

- Note that in a theory like $\lambda\phi^4$, where the interaction vertices are *independent of the momentum k*, the number of vertices does not contribute to the overall k-scaling of the diagram.

This result implies that a Feynman diagram with a superficial degree of divergence D will scale with momentum as $\sim k^D$. This is already useful information, since without doing any calculation we can already determine the overall scaling of the Feynman diagram. We should be careful since this still does not necessarily tell us if a diagram is convergent or divergent, we might need some additional information for this.

Note that in this derivation we have exploited the fact that in $\lambda\phi^4$ theory the Feynman rule for the vertex does not depend on momenta. In theories where this is not the case, the definition of Eq. (2.21) would be different. QFTs with derivative interactions will have associated powers of the momentum at each vertex and thus would contribute to the overall scaling with k.

Using again the same three Feynman diagrams as above as an example, we can see that:

- The one-loop diagram in Fig. 1.3 for the scattering amplitude has $d = 4$, $L = 1$, and $I = 2$, which corresponds to an apparent degree of divergence of $D = 0$. This is consistent with the *logarithmic divergence* which was found from the explicit computation of the Feynman diagram. Note however that this would have also been consistent with a constant finite result.

- The two-loop diagram to the propagator in Fig. 1.4 has $d = 4$, $L = 2$, and $I = 3$, which corresponds to $D = 2$. And indeed an explicit calculation of this Feynman diagram shows that it has a *quadratic divergence*, similarly to the corresponding one-loop diagram.

- For the two-loop diagram contributing to the four-point scattering amplitude, Fig. 2.6, we have that in $d = 4$ space-time dimensions it has $L = 2$ loops and $I = 4$ internal lines, and therefore its naive or apparent degree of divergence is $D = 4 \times 2 - 2 \times 4 = 0$.

  Note that this result is D is consistent either with a logarithmic divergence, since dk/k has $D = 0$, but it could also imply a finite result, that is, without any dependence on the undetermined momentum k.

  Interestingly, we find that the apparent degree of divergence D is the same at one loop that at two loops for the $\phi\phi \to \phi\phi$ scattering amplitude. Actually for these types of diagrams this relation holds for all loops, since $D = 4 \times L - 2 \times I$ but for every new loop, $L \to L + 1$, we need two more internal lines, $I \to I + 2$, so if $D = 0$ at the one loop level, it is also zero at all other orders.

  Note that this is also true at tree level: $L = 0$ and $I = 0$, so $D = 0$ trivially. And in this case the result was finite, showing that $D = 0$ does not necessarily imply a logarithmic divergence.

In addition to the above results, we need one more relation in order to write the superficial degree of divergence D in terms of the number of vertices V and of external lines E that a given Feynman diagram has. This is necessary in order to assess how the interactions affect the degree of divergences of a the diagram. For this purpose, suppose that V_N stands for the number of vertices with N legs. Note that in general a QFT can have vertices with three, four, five, or even more legs N. In a diagram with V_N such vertices, we have NV_N lines which are either internal or external, so one has the relation

$$NV_N = E + 2I \, . \tag{2.22}$$

Note that an internal line counts, twice since it both originates and ends at a vertex.

Combining all the relations that we have derived up to here, we can find that:

$$
\begin{aligned}
L &= I - V_N + 1 \,, \\
D &= d \times L - 2 \times I \,, \\
NV_N &= E + 2I \,,
\end{aligned}
\tag{2.23}
$$

and therefore the expression for the apparent degree of divergence of a Feynman diagram D in terms of N, V_N and the number of external legs E reads:

$$
D = d - \frac{1}{2}(d-2)E + V_N \left(\frac{N-2}{2}d - N \right) \,.
\tag{2.24}
$$

In the explicit case of $d = 4$ space-time dimensions. the apparent degree of divergence of the diagram is given by:

$$
D = 4 - E + (N-4)V_N \,.
\tag{2.25}
$$

The crucial property of this result is that

> In a theory without derivative interactions, the apparent degree of divergence D of a Feynman diagram depends only on the number of external legs E and the number of vertices V_N with N legs, but *not* on the number of loops or of internal lines.
>
> In particular, we find that in the case of $\lambda\phi^4$, where the only interaction vertex is V_4, then $D = 4 - E$ depends *only on the number of external legs*, and is therefore the same for any number of loops provided that E is fixed. So the value of D for the one-loop scattering amplitude $\phi\phi \to \phi\phi$ is the same than say in the 20 loop calculation.

Let's verify that this relation holds for the same three diagrams that we are considering in this section:

- The one-loop diagram in Fig. 1.3 for the scattering amplitude has $E = 4$, $N = 4$ and $V_4 = 2$, which corresponds to an apparent degree of divergence of $D = 4 - 4 + (4-4) \times 2 = 0$, consistent with the previous result.

- The two-loop diagram to the propagator in Fig. 1.4 has $E = 2$, $N = 4$, and $V_4 = 2$, which corresponds to $D = 4 - 2 + (4-4) \times 2 = 2$, again consistent with the above result, which implies a quadratic divergence of the diagram.

- For the two-loop diagram of Fig 2.6, we have that $E = 4$, $N = 4$ and $V_4 = 3$, which corresponds to an apparent degree of divergence of $D = 4 - 4 + (4-4) \times 3 = 0$, consistent with the previous result and with the general conclusion that in $\lambda\phi^4$ the value of D depends on on the number of external legs E.

We can put together now all this information in order to construct a possible classification of the degree of divergence of a given Feynman diagram.

Classification of divergences. With these results it is possible to construct the following *naive classification* about whether or not a given Feynman diagram is UV divergent:
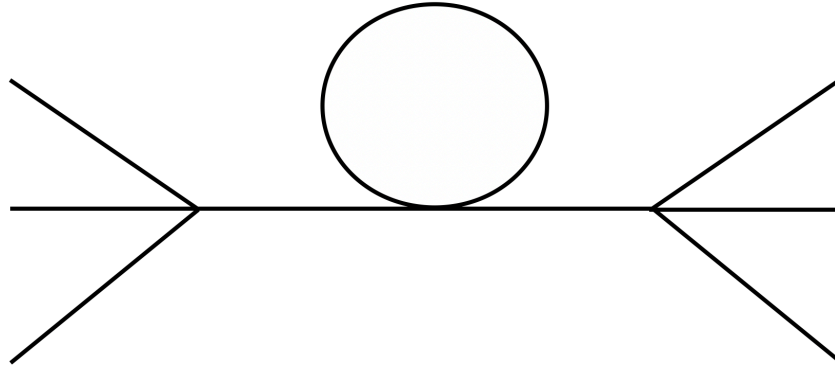
Figure 2.7. The one-loop correction to the six-point scattering amplitude $\mathcal{M}\left(\phi\phi\phi \to \phi\phi\phi\right)$ in $\lambda\phi^4$ theory. Without computing this diagram we can state that it will not contain divergences, since its apparent degree of divergence is $D = 4 - E = 4 - 6 = -2 < 0$.

- $D > 0$: diagram is UV *divergent*.

- $D = 0$: diagram is likely to be *logarithmically UV divergent*, but could also be finite (constant result).

- $D < 0$: diagram is likely to be UV *finite* (but in some cases can also be divergent).

Consider specifically the case of the $\lambda\phi^4$ theory. In this case $N = 4$, since the only interaction possible in this theory involves four legs. Then we have that $D = 4 - E$. The interesting result is that the degree of divergence only depends on the number of external legs, but it is actually *independent of the number of vertices* that we have in the diagram. Therefore we have only two classes of *problematic* diagrams in this theory, namely:

- Diagrams with 2 external legs, for which $D = 2$, corresponding to loop corrections to the propagator.

- Diagrams with 4 external legs, for which $D = 0$, corresponding to loop corrections to the $2 \to 2$ scattering amplitude.

Note that diagrams with a higher number of legs, for example those contributing to the six point scattering amplitude $\mathcal{M}\left(\phi\phi\phi \to \phi\phi\phi\right)$, are automatically finite. Indeed, we don't need for example to explicitly compute the one-loop correction to the six-point scattering amplitude $\mathcal{M}\left(\phi\phi\phi \to \phi\phi\phi\right)$ shown in Fig. 2.7, since we can determine that its degree of divergence to be $D = -2$ and therefore the calculation would yield a finite result. This highlights the usefulness of relations such as Eq. (2.25): we can determine whether or not they might be finite or divergent without explicitly evaluating the diagram.

This result also implies that, for the case of $\lambda\phi^4$ theory, *all* the counterterms that we need to renormalize the theory are those required to renormalize both the two-point function (the propagator) and the four-point function (the $\phi\phi \to \phi\phi$ scattering amplitude). Therefore, we can conclude that with the counterterms that have already been added above to the Lagrangian of the $\lambda\phi^4$ theory, Eq. (2.17), all possible diagrams that one can compute *will be finite*. In other words, we only need a finite number of terms to make finite the physical predictions of the theory, and therefore we can conclude that it is *renormalizable*. And as we compute more and more loops, the counterterms Z_λ, Z_ϕ and Z_m will get modified with higher order corrections, but no more counterterms will be required.

Before concluding this discussion, we need to clarify one further point. Why we call D a *naive* degree of divergence? The reason is that a diagram with $D = 0$ will *not* necessarily have a logarithmic divergence. Such divergence is possible and compatible with the diagram's features, but is not necessary. For example, *tree-level diagrams* have $D = 0$ in $\lambda\phi^4$ theory (trivially from Eq. (2.21) since $L = I = 0$), yet they are obviously finite. Also, $D < 0$ does not necessarily imply that a given diagram is finite. Understanding the actual degree of divergence of such a diagram is more delicate that this naive analysis, and in general requires to compute the diagram.

What this analysis can indicate is whether or not a given theory exhibits *UV divergences*, but it does not allow us to prove that it is renormalizable. For example in $d = 4$ we had

$$D = 4 - E + (N - 4)V_N \,. \tag{2.26}$$

For a theory with interactions of the form ϕ^5 or even with a higher power, there are infinitely many diagrams that are divergent with arbitrary number of external lines for which we will potentially need a different counterterm to cancel it. This is not inconsistent with our previous discussion: a theory that contains an interaction of the form $\lambda\phi^5$ has a coupling constant with mass dimensions $[\lambda] = -1$ and therefore this is an *irrelevant interaction* which in general will not be renormalizable. And for this case the naive degree of divergence reads

$$D = 4 - E + V_N \,, \tag{2.27}$$

so for a given number of external lines E, there are an infinitely large number of diagrams with $D > 0$ and therefore divergent. Each of this divergences will have a different power, and therefore a different counterterm needs to be associated to it. This is another way to illustrate how such interaction corresponds to a non-renormalizable theory.

Similar considerations as the ones presented here can be applied to other QFTs to understand the all-order behaviour of their scattering amplitudes. However in general the construction of the apparent degree of divergence D will different theory by theory, for instance in the presence of different interaction vertices.

Degree of divergence in scalar QED

In one of the exercises we will repeat the calculation of D presented here for the case of scalar QED,

$$\mathcal{L}_{\text{sQED}} \subset +ieA^\mu \left[(\partial_\mu \psi^*) \psi - \psi^* (\partial_\mu \psi) \right] - e^2 A_\mu A^\mu \psi^* \psi \,,$$

where the power counting of Feynman diagrams will differ from the case of $\lambda\phi^4$ theory: in particular, diagrams with a vertex ψAA will scale as k^μ, given the derivative nature of the corresponding interaction term in the Lagrangian.

3 Scalar QED and the role of symmetries in QFTs

We now move to explore a number of interesting topics in QFTs that involve Abelian gauge fields, such as Quantum Electrodynamics (QED) and its scalar counterpart, scalar QED (sQED). The latter is particularly useful to illustrate the principal concepts of the QFT of electromagnetism in a simplified setup. We will briefly review some of the aspects of the quantization of Abelian gauge symmetries that have already been discussed in the *Quantum Field Theory* course, and then introduce a number of new concepts focusing on the role that classical symmetries have at the QFT level.

First of all, we will review the covariant formulation of Maxwell's electromagnetism, with emphasis on the classical symmetries and related properties. This will include a discussion of the coupling between matter and gauge fields. We will then move to describe the canonical quantization of Maxwell's equations and derive the Feynman rules for the interaction between electrons and photons. We will use these rules to compute some simple processes in scalar QED, and highlight the deep connection between the theories of the symmetry and the corresponding physical results. At the end of this part, we will connect with the contents of the first part of the course and consider the behaviour of scalar QED at the one-loop level.

3.1 Maxwell's electromagnetism in the covariant formalism

The equations of motion of electromagnetism are Maxwell's equations, which when expressed in terms of the electric and magnetic field read

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot \mathbf{E} &= \rho\,, \\
\boldsymbol{\nabla} \times \mathbf{B} - \frac{\partial}{\partial \mathbf{t}} \mathbf{E} &= \mathbf{J}\,, \\
\boldsymbol{\nabla} \times \mathbf{E} + \frac{\partial}{\partial \mathbf{t}} \mathbf{B} &= 0\,, \\
\boldsymbol{\nabla} \cdot \mathbf{B} &= 0\,,
\end{aligned}
\tag{3.1}
$$

where quantifies in boldface refer to (three-dimensional) vectors. In these equations, ρ and $\mathbf{J}$ are the charge and current densities respectively, that in general depend on the space time coordinates $(t, \mathbf{x})$, and that we will take as external inputs for the time being.

The last two equations of Eq. (3.1), which do not have external sources, can be solved by introducing a vector potential $\mathbf{A}$ and a scalar potential ϕ so that

$$
\begin{aligned}
\mathbf{E} &= -\boldsymbol{\nabla}\phi - \frac{\partial}{\partial t} \mathbf{A}\,, \\
\mathbf{B} &= \boldsymbol{\nabla} \times \mathbf{A}\,.
\end{aligned}
\tag{3.2}
$$

You can verify that the two homogeneous Maxwell's equations are always satisfied if $\mathbf{E}$ and $\mathbf{B}$ are defined as in Eq. (3.2), using the fact that the divergence of the curl of any vector field $\mathbf{X}$ vanishes, as does the curt of the gradient of any scalar field ϕ,

$$
\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{X}) = 0\,, \qquad \boldsymbol{\nabla} \times (\boldsymbol{\nabla}\phi) = 0\,.
\tag{3.3}
$$

The fact that we can replace the electric and magnetic fields $(\mathbf{E}, \mathbf{B})$ by the scalar and vector potentials

$(\phi, \mathbf{A})$ reveals something rather deep about the structure of the theory, since the former has *6 independent components* while the latter only has only *4 components*. Therefore, Maxwell's electromagnetism expressed in terms of $(\mathbf{E}, \mathbf{B})$ has an inherent redundancy, with important consequences as we will now discuss.

Maxwell's theory, described by Eq. (3.1) admits a very elegant description in terms of the covariant formalism of special relativity, which will also be useful in order to later promote it to the quantum field theory level. With this motivation, we define a four-vector potential

$$A^\mu \equiv (\phi, \mathbf{A}) \,, \tag{3.4}$$

which is known as the *gauge field*, and the corresponding tensor

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu \,, \tag{3.5}$$

which goes under the name of the *field strength* tensor. Note that by construction this tensor is antisymmetric, $F^{\mu\nu} = -F^{\nu\mu}$. By comparing with Eq. (3.1), we can identity

$$F^{0i} = E^i \,, \qquad F^{ij} = \epsilon^{ijk} B_k \,, \tag{3.6}$$

and this way write the first two equations of motion (the inhomogeneous Maxwell's equations) as

$$\partial_\nu F^{\mu\nu} = J^\mu \,, \tag{3.7}$$

where the four-current is naturally defined as $J^\mu \equiv (\rho, \mathbf{J})$. The last two of Maxwell's equations (the homogeneous ones) can be also written in terms of the field strength tensor as

$$\epsilon_{\mu\nu\rho\sigma} \partial^\rho F^{\mu\nu} = 0 \,, \tag{3.8}$$

which is known as the *Bianchi identity*, and $\epsilon_{\mu\nu\rho\sigma}$ is the maximally antisymmetric tensor. In the absence of external currents, $J = (0, \mathbf{0})$, the Lagrangian of Maxwell's electromagnetism is given by

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} = \frac{1}{2} \left(\mathbf{E}^2 - \mathbf{B}^2 \right) \,, \tag{3.9}$$

from which the equations of motion of the theory, namely the Maxwell's equations given in Eq. (3.1), can be derived using the Euler-Lagrange equations. This can be demonstrated by evaluating the Euler-Lagrange equation for the four-potential A_μ

$$\frac{\partial L}{\partial A_\mu} - \partial_\nu \left(\frac{\partial L}{\partial (\partial_\nu A_\mu)} \right) = 0 \,, \tag{3.10}$$

which leads to the classical equations of motion.

Properties of the four-potential A_μ.   The vector field A_μ has of four components, so one could conclude that the gauge field represented by it has 4 physical degrees of freedom. However, we know that from electromagnetism that physical (massless) photons have only two degrees of freedom, namely its two possible polarization states. To understand this apparent discrepancy, let us express the Lagrangian as a function of

the gauge field A_μ restricting ourselves to the purely gauge part. In this case one finds

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu + \frac{1}{2}\partial^\mu A^\nu \partial_\nu A_\mu \tag{3.11}$$

If we inspect this equation, we observe that the Lagrangian does not contain time derivatives of A_0, since these cancel out due to its symmetry structure. Therefore, we know from classical mechanics that is the Lagrangian contains no $\partial A_0 / \partial t$ terms, then this means that the degree of freedom A_0 is not dynamical (in other words, $\partial_t A_0 = 0$). This conclusion can also be derived from the Euler-Lagrange relation, Eq. (3.10).

The same result, namely that the number of physical components of A_μ is smaller than 4, can also be derived from inspecting the equations of motion, where we have that

$$\boldsymbol{\nabla} \cdot \mathbf{E} = 0 \quad \to \quad \boldsymbol{\nabla}^2 A_0 + \boldsymbol{\nabla} \cdot \frac{\partial \mathbf{A}}{\partial t} = 0 \,, \tag{3.12}$$

where we have used the definition of the vector potential in terms of the electric field in Eq. 3.2. This equation can be solved for the time-component of the four vector $A_0(\mathbf{x})$ to give

$$A_0(\mathbf{x}) = \int d^3 x' \frac{1}{4\pi|\mathbf{x}-\mathbf{x}'|}\left(\boldsymbol{\nabla} \cdot \frac{\partial \mathbf{A}}{\partial t}\right)(\mathbf{x}') \,, \tag{3.13}$$

which illustrates that given some boundary conditions for the spatial components of the gauge field $\mathbf{A}$, then the non-dynamical time component $A_0(\mathbf{x})$ is fully specified. Therefore, we can conclude that:

> A_0 is not an independent degrees of freedom, and we are left with *3 independent degrees of freedom.* This is better than before, but still one too many degree of freedom.

In order to eliminate the other unphysical degree of freedom, we note another symmetry that we have not discussed up to now. Indeed, the Lagrangian of our theory, Eq. (3.9), has a large additional symmetry group, since it is fully invariant under transformations of the form

$$A_\mu \to A_\mu + \partial_\mu \lambda(x) \,, \tag{3.14}$$

with $\lambda(x)$ an arbitrary scalar function. Note that not only the Lagrangian, but also the field strength $F_{\mu\nu}$ itself is invariant under this transformation. We call this kind of invariance a *gauge symmetry*, for historical reasons that are not really relevant here.

Crucially, $\lambda(x)$ in general depends on the space-time position x. A gauge symmetry is an example of a *local symmetry*, namely a symmetry transformation which is different in each space-time point x^μ as opposed to a global transformation which is the same irrespective of the value of x^μ. This symmetry is so important that theories that exhibit it are denoted as *gauge theories*: this is almost its defining factor.

The existence of this additional *local* symmetry implies that A_μ provides a redundant description of the electromagnetic field $F_{\mu\nu}$. One possible way to show that gauge theories are *redundancies of the theory*, and here are on a quite different conceptual footing as other types invariance such a *global* symmetries, is

obtained by looking again to Maxwell's equations in the covariant formulation, and noting that

$$\partial_\mu F^{\mu\nu} = \left(\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu\right) A^\nu = 0\,. \tag{3.15}$$

Crucially, the operator that appears in Eq. (3.15), $\left(\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu\right)$ is *non-invertible*. This means that Eq. (3.15) is satisfied by any function of the form $\partial_\mu\lambda(x)$, which is a total differential. In turn, this property implies that we cannot really distinguish between $A_\mu(x)$ and $A_\mu(x) + \partial_\mu\lambda(x)$ as solutions of the Maxwell's equations: they are physically the same. However, if we impose the formal equivalence

$$A_\mu \approx A_\mu + \partial_\mu\lambda\,, \tag{3.16}$$

then the whole problem disappears, and Maxwell's equations become invertible.

Gauge choices. As a result of this discussion, we find that in classical electrodynamics all states that are related by a gauge transformation of the form

$$A_\mu \to A_\mu + \partial_\mu\lambda(x)\,, \tag{3.17}$$

are considered to be *physically equivalent*. By making a *gauge choice*, which is a specific choice of Eq. (3.17), we can select a representative of each family of states related by gauge transformation. Differences choices of gauge are more suitable than others for specific problems, but at the end of the day results should be the same irrespective of the specific gauge choice. In other words, we say that physical observables should be *gauge independent*. As we will see, this provides a very important constraint on QED calculations: we can keep explicit the dependence on the gauge-fixing parameters, and we should find that they disappear at the end of the calculation.

There are a number of widely used gauge choices, each one defined by a *gauge condition*. Some of the most frequent choices are:

- The *Lorenz gauge*.

  The Lorenz choice of gauge is defined by the following condition:[3]

$$\partial_\mu A^\mu = 0\,, \tag{3.18}$$

  which in terms of the vector $\mathbf{A}$ and scalar potential ϕ reads

$$\boldsymbol{\nabla}\cdot\mathbf{A} + \frac{\partial\phi}{\partial t} = 0\,. \tag{3.19}$$

  This choice does not remove all the redundancies in the theory, since any gauge transformation that satisfies $\partial_\mu\partial^\mu\lambda(x) = 0$ will still preserve the Lorentz gauge condition. Indeed, one can see that

$$\partial_\mu A^\mu = 0 \quad \Rightarrow \quad \partial_\mu\left(A^\mu + \partial^\mu\lambda(x)\right) = 0 \tag{3.20}$$

  for this specific subset of gauge transformations.

---

[3]Note that the Lorentz gauge is named after the Danish physicist Ludwig *Lorenz*, not after the Dutch physicist Hendrik *Lorentz* of the eponymous transformations.

The main advantage of Eq. (3.18) is that it is Lorentz invariant (valid in any reference frame), but the drawback is that since there is still some redundancy, it is sometimes difficult to keep track of the physical degrees of freedom. So other options for the gauge fixing might be advantageous for specific calculations.

- The *Coulomb gauge.*

  Starting from the Lorentz gauge, Eq. (3.18), we can make use of the further residual freedom that is still available in the gauge transformation, which allows to use any function that $\partial_\mu \partial^\mu \lambda(x) = 0$, to set

  $$\boldsymbol{\nabla} \cdot \mathbf{A} = 0 \,. \tag{3.21}$$

  In this case, we find by comparing with Eq. (3.13) that the time component of the gauge field vanishes, $A_0 = 0$. While the Coulomb gauge Eq. (3.21) is not Lorentz invariant, it allows exhibit the physical degrees of theory in an explicit way.

In order to illustrate the usefulness of the Coulomb gauge, let us write Maxwell's equations explicitly in this gauge. Using Eq. (3.21) and setting $A_0 = 0$, we can write

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} = -\frac{1}{2} \partial_\mu A_\nu \partial^\mu A^\nu + \frac{1}{2} \partial^\mu A^\nu \partial_\nu A_\mu = -\frac{1}{2} \partial_t A_i \partial_t A^i - \frac{1}{2} \partial_j A_i \partial^j A^i \,. \tag{3.22}$$

Next, if we now if we compute the Euler-Lagrange equation for the spatial (non-zero) components of the gauge field A_i,

$$\partial_\nu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\nu A_i)} \right) - \frac{\partial \mathcal{L}}{\partial A_i} = 0 \,, \quad i = 1, 2, 3 \,, \tag{3.23}$$

we find that the classical equation of motion for the dynamical components of the gauge field A_i are

$$\Box A_i = \left(-\frac{\partial^2}{\partial t^2} + \boldsymbol{\nabla}^2 \right) A_i = 0 \,, \tag{3.24}$$

where $\Box$ denotes the D'Alambertian operator.

The resulting Eq. (3.24) for the spatial components of the four-potential A_i should be familiar to us, since they correspond to the *Klein-Gordon equation* for the three scalar fields A_1, A_2, and A_3. Using the known solutions of the Klein-Gordon equation, we find that the general solution for $\mathbf{A}$ can be written as

$$\mathbf{A}(\mathbf{x}, t) = \sum_\lambda \int \frac{d^3 k}{2\omega(2\pi)^3} \left[(\vec{\epsilon}_\lambda)^*(k) a_\lambda(k) e^{ik \cdot x} + \vec{\epsilon}_\lambda(k) a_\lambda^\dagger(k) e^{-ik \cdot x} \right] \,, \tag{3.25}$$

where $\vec{\epsilon}_\lambda$ are the gauge field polarization vectors, $A_0 = 0$ in the Coulomb gauge, and the sum runs over $\lambda = 1, 2$ the two physical polarizations of the photon. As can be seen from the applying Eqn. (3.24) and to Eqn. (3.25), the solutions of the massless Klein-Gordon equation satisfy the mass-shell condition $(k^\mu)^2 = 0$, reflecting the fact that classically photons are massless. At the classical level, $a_\lambda(k)$ are just functions, and they will be later promoted to operators once we quantize the photon field, specifically to the creation and annihilation operators. This coefficients $a_\lambda(k)$ determine the relative strength of the gauge field modes associated to the wave vector $\vec{k}$.

The Klein-Gordon equation

The Klein-Gordon equation is the relativistic equivalent of the Schroedinger wave function equation in quantum mechanics. It is given in natural units by

$$\left(\Box + m^2\right)\phi(x) = 0 \tag{3.26}$$

where $\phi(x)$ is the wave function, m is the particle mass and the d'Alembert operator is defined by

$$\Box \equiv -\frac{\partial^2}{\partial t^2} + \boldsymbol{\nabla}\,. \tag{3.27}$$

While the Klein-Gordon equation is Lorentz invariant, it has unphysical solutions with negative energy. Therefore it is not the suitable relativistic generalization of Schroedinger equation, which is instead provided by the Dirac equation. For a free particle, the solution of the Klein-Gordon equation is given by plane waves of the form

$$\phi(\mathbf{x}, t) = \exp\left(i(\mathbf{k}\cdot\mathbf{r} - \omega t)\right)\,, \tag{3.28}$$

which have associated the following dispersion relation:

$$E = \sqrt{m^2 + \mathbf{p}^2}\,. \tag{3.29}$$

The general solution of the Klein-Gordon equation, assuming finite boundary conditions, is

$$\varphi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega}\left[a(\mathbf{k})e^{ikx} + a^*(\mathbf{k})e^{-ikx}\right]\,, \tag{3.30}$$

where the phase-space term $d^3k/((2\pi)^3 2\omega)$ is Lorentz invariant. The quantization of the Klein-Gordon field $\varphi(x)$ can be achieved by promoting the coefficients $a(\mathbf{k})$ to operators satisfying a suitable set of commutation relations.

The polarisation vectors. Let us now determine the form of the gauge field polarization vectors $\vec{\epsilon}_\lambda$ that appear in Eq. (3.25). From the Coulomb gauge condition, Eq. (3.21), it follows that

$$\vec{\epsilon}_\lambda(k)\cdot\mathbf{k} = 0\,, \quad \lambda = 1, 2\,, \tag{3.31}$$

that is, the polarization vectors $\vec{\epsilon}_\lambda(k)$ are perpendicular to the direction of motion $\mathbf{k}$. To see this, note that if we apply the Coulomb gauge-fixing condition for the expression we have just derived for the spatial components of the four-potential, Eq. (3.25), then we get that

$$0 = \boldsymbol{\nabla}\cdot\vec{A} = i\sum_\lambda \int \frac{d^3k}{2\omega(2\pi)^3}\left[\left((\vec{\epsilon}_\lambda)^*(k)\cdot\vec{k}\right)a_\lambda(k)e^{ik\cdot x} - \left(\vec{\epsilon}_\lambda(k)\cdot\vec{k}\right)a_\lambda^*(k)e^{-ik\cdot x}\right]\,, \tag{3.32}$$

which requires the photon polarization to be perpendicular to the direction of propagation. This requirement, Eq. (3.31), has two independent solutions, which we label for convenience as $\lambda = \pm$. Assuming that the

photon propagates in the z direction, $\mathbf{k} = (0, 0, |k|)$, a suitable choice of polarization vectors corresponds to

$$
\begin{aligned}
\vec{\epsilon}_+(k) &= \frac{1}{\sqrt{2}} (1, -i, 0) \,, \\
\vec{\epsilon}_-(k) &= \frac{1}{\sqrt{2}} (1, +i, 0) \,,
\end{aligned}
\tag{3.33}
$$

which physically corresponds to *circularly polarized* photons, and where the overall prefactor follows from the normalization condition. This explicitly shows how the two physical polarization states of the photon are *orthogonal to each other*, namely

$$
\vec{\epsilon}_+(k) \cdot \vec{\epsilon}_-^*(k) = 0 \,.
\tag{3.34}
$$

For a generic direction of the propagation of the photon, these *polarization vectors* form an orthonormal and complete set, and satisfy the following conditions

$$
\begin{aligned}
\epsilon_\lambda(k) \cdot \epsilon_{\lambda'}^*(k) &= \delta_{\lambda\lambda'} \,, \\
\sum_{\lambda=\pm} \epsilon_{i,\lambda}^*(k) \epsilon_{j,\lambda}(k) &= \delta_{ij} - \frac{k_i k_j}{|\mathbf{k}|^2} \,,
\end{aligned}
\tag{3.35}
$$

where you can easily check that the orthogonality requirement Eq. (3.31) that follows from the Coulomb gauge fixing condition holds now for a generic vector $\mathbf{k}$.

It is also worth mentioning here that relations of the form of Eq. (3.35) are useful when evaluating Feynman diagrams that include photons in the initial and final states (thus real, on mass-shell, photons). The reason is that in most cases the photon polarisation is not experimentally accessible, and therefore in most calculations we will be interested in summing over final-state photon polarisation states and averaging over the initial ones.

3.2   Coupling gauge and matter fields

The discussion so far has been restricted to the case of electromagnetism *without external charges or currents*, in other words, without accounting for the effects of the coupling to matter. Such situation is known as the *pure gauge theory*. Let us now study how the quantum field theory of electromagnetism is modified in the presence of a non-zero four-current $j^\mu \neq 0$. For the time being we are still reviewing the theory at the classical level, and then afterwards we will study the consequences of its quantisation. In this case, the equations of motion of the theory read

$$
\begin{aligned}
\partial_\mu F^{\mu\nu} &= j^\nu \,, \\
\epsilon_{\mu\nu\lambda\rho} \partial^\nu F^{\lambda\rho} &= 0 \,.
\end{aligned}
\tag{3.36}
$$

The antisymmetry property of the field strength tensor $F_{\mu\nu}$ together with the first of Eq. (3.36) implies that

$$
\partial_\nu \partial_\mu F^{\mu\nu} = \partial_\nu j^\nu = 0 \,,
\tag{3.37}
$$

and therefore one finds the important result that:

only conserved currents ($\partial_\nu j^\nu = 0$) are consistent with Maxwell's electromagnetism.

This property, that only conserved four-currents have a place in classical electrodynamics, can be understood as a direct consequence of the gauge invariance exhibited by the theory. Indeed, adding a current to the equations of motion of the theory can be achieved by means of a covariant coupling of the current to the gauge field in the Lagrangian, leading to an action

$$S = \int d^4 x \left(-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + j^\mu A_\mu \right) . \tag{3.38}$$

Now let us see how the action is affected if we perform a gauge transformation of the form Eq. (3.17). While $F_{\mu\nu}$ is gauge invariant, the matter-gauge coupling term is modified:

$$S \to S = \int d^4 x \left(-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + j^\mu \left(A_\mu + \partial_\mu \lambda \right)) \right) , \tag{3.39}$$

This means that the action transforms under a gauge transformation as

$$S \to S + \int d^4 x \, \partial_\mu \left(j^\mu \lambda(x) \right) - \int d^4 x \left(\partial_\mu j^\mu \right) \lambda(x) , \tag{3.40}$$

where we have integrated by parts. Since the four-current goes to zero at infinity for all sensible boundary conditions, the second term in Eq. (3.40) vanishes, being a total differential. Therefore we find that the action under a gauge transformation varies as

$$S \to S - \int d^4 x \left(\partial_\mu j^\mu \right) \lambda(x) , \tag{3.41}$$

and therefore S is only invariant under gauge transformations if the four-current is conserved, $\partial_\mu j^\mu = 0$. So the conservation of the four-current is a direct consequence of the gauge invariance of electrodynamics.

Having realized that only conserved currents are allowed, let us now see which further information we can derive on the form of the coupling between matter and gauge fields. In this respect, gauge invariance is extremely useful as a guiding principle. The reason is that if one *ensures gauge invariance in the action*, we know the the resulting equations of motion will be unchanged and that the number of degrees of freedom will not be modified.

To move forward, we need to assume a specific scenario for the matter particles that define the external current. Instead of a current based on fermion fields, as is the case in Quantum Electrodynamics (QED), we will consider instead currents constructed from complex scalar fields ϕ, that define the theory known as Scalar Quantum Electrodynamics (sQED).

Scalar QED. To illustrate how we can couple photons with matter fields in the case of scalar particles, assume that we want to add the following term to the Maxwell Lagrangian

$$\mathcal{L} \to \mathcal{L} + A_\mu \phi \partial^\mu \phi , \tag{3.42}$$

which would describe the coupling between the photon and a real scalar field ϕ. One can verify that this term is not gauge invariant and thus not allowed if we want to keep our theory invariant under gauge transformations. To see this, note that due to a gauge transformation we have

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \lambda(x) \,, \tag{3.43}$$

so the additional term in Maxwell's Lagrangian would transform as

$$A_\mu \phi \partial^\mu \phi \to A_\mu \phi \partial^\mu \phi + (\partial_\mu \lambda(x)) \phi \partial^\mu \phi \,, \tag{3.44}$$

which depends on the gauge-fixing function and is thus not gauge invariant. Note that the real scalar ϕ does not vary upon a gauge transformation.

> You can convince yourselves that any coupling between the photon and a real scalar field will never satisfy gauge invariance. The reason is that gauge invariance will only be satisfied if the new field ϕ transforms in a way that exactly cancels the gauge transformation of the photon. This restricts the ways in which matter fields can couple to photon fields: only couplings that satisfy gauge invariance will be allowed.

Therefore, while coupling a photon to a real scalar field seems to be difficult to reconcile with gauge invariance, the situation is more interesting with a complex scalar field, which we will denote by ψ (not to confuse with a Dirac fermion field - only scalar fields will appear in our study of sQED). This complex field is composed by two independent real scalar fields:

$$\psi = \phi_1 + i\phi_2 \,. \tag{3.45}$$

Then we will assume that under a gauge transformation this complex field changes as

$$\psi(x) \to e^{-i\alpha(x)} \psi(x) \,, \tag{3.46}$$

with $\alpha(x)$ being an arbitrary real function. Note that this kind of transformation (a rotation in the complex plane) was not possible for the couplings between the photon and a real scalar field as that given by Eq. (3.42). Being a scalar field, the kinetic term associated to ψ will be the Klein-Gordon Lagrangian

$$\mathcal{L}_{\psi,\text{kin}} = -\partial_\mu \psi^* \partial^\mu \psi - m^2 \psi^* \psi \,, \tag{3.47}$$

which under a gauge transformation of the form Eq. (3.46) transforms as

$$\mathcal{L}_{\psi,\text{kin}} \to \mathcal{L}_{\psi,\text{kin}} + (\partial_\mu \alpha \partial^\mu \alpha) \; \psi^* \psi \,, \tag{3.48}$$

which is not gauge invariant. In order to ensure that the Lagrangian Eq. (3.47) is invariant under gauge transformations, we can promote its partial derivative into an object known as the *covariant derivative*, defined as

$$D_\mu \equiv \partial_\mu + ieA_\mu \,, \tag{3.49}$$

with e some unspecified constant. It is now possible to verify that under a gauge transformation of the form $A_\mu \to A_\mu + (\partial_\mu \alpha)/e$ the covariant derivative transforms as

$$D_\mu \psi \to e^{-i\alpha(x)} D_\mu \psi \,, \tag{3.50}$$

that is, exactly in the same way as the field ψ itself, see Eq. (3.46). Therefore, a Lagrangian of the form

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (D_\mu \psi)^* (D_\mu \psi) - m^2 \psi^* \psi \,, \tag{3.51}$$

can be shown to be fully gauge invariant. It therefore represents a physically consistent way to model the interactions between a photon and a complex scalar field. This theory is know as *scalar QED* (sQED), since it has the same structure as the usual quantum electrodynamics but with a scalar field replacing the Dirac fermion. If instead of coupling the photon to a complex scalar we couple it to a fermion field governed by Dirac's equation, then we obtain the QED Lagrangian:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^2 + \bar{\psi} i \gamma^\mu D_\mu \psi - m \bar{\psi} \psi \,. \tag{3.52}$$

Note that the gauge invariance requirement fixes completely the coupling between matter and gauge fields, irrespective of the specific type of the former. This is a powerful consequence of gauge invariance: it constraints very significantly the possible ways in which gauge and matter fields can couple among themselves.

Gauge symmetries and conserved currents.   In general, the symmetries of a theory can be classified into two main types:

- *Local symmetries.*

  These symmetries are parametrized by a continuous function $\lambda(x)$ which depends on the space-time position x. Gauge symmetry is the most representative example of this family.

- *Global symmetries.*

  These are instead given in terms of constant (independent of the space-time position) parameters (numbers, vectors, matrices). Examples of global symmetries include translations, rotations, and Lorentz boosts.

Note that any local symmetry will have also associated a global symmetry, corresponding to the particular case where the function λ is constant.

From *Noether's theorem* we know that for each global (continuous) symmetry has associated a conserved charge. Does this also apply to local symmetries such as gauge transformations? In other words, has Maxwell's Lagrangian (without external currents) a conserved charge associated to gauge invariance? Actually, it can be shown however that there are no conserved charges associated to local symmetries such as gauge invariance, and that only global symmetries have associated conserved charges.

Noether's Theorem

Assume a theory which describes some fields $\{\phi_i\}$ in terms of a Lagrangian $\mathcal{L}(\phi_a, \phi_b, \ldots)$. These fields can be either scalars, spinors, or vector fields. Assume we have a *global symmetry* that leaves the *action invariant*. At the infinitesimal level, this means that a transformation of the fields of the form

$$\delta\phi_a(x) = X_a(\{\phi_i\}),\tag{3.53}$$

leaves the Lagrangian invariant up to a total derivative, $\delta\mathcal{L} = \partial_\mu G^\mu$, so that the action itself is invariant under such transformation. Now, an arbitrary transformation of the fields (not necessarily the one associated to the symmetry that we are considering here) leads to:

$$
\begin{aligned}
\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial_\mu(\delta\phi_a), \\
&= \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\frac{\partial L}{\partial(\partial_\mu\phi_a)}\right]\delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right).
\end{aligned}\tag{3.54}
$$

Now, the terms in brackets vanishes due to the Euler-Lagrange equations of motion. Therefore we find that the total variation of the Lagrangian is given by

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right),\tag{3.55}$$

and for the transformation of the specific form of Eq. (3.53), this Lagrangian variation should a total differential,

$$\delta\mathcal{L} = \partial_\mu G^\mu.\tag{3.56}$$

Therefore, we obtained the important result that the existence of this global symmetry has associated a conserved current

$$\partial_\mu j^\mu = 0,\tag{3.57}$$

given by

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}X_a(\phi) - G^\mu(\phi),\tag{3.58}$$

which is the result of Noether's theorem. Eq. (3.58) is known as *Noether's current*, and illustrates the deep connection between the symmetries of a theory and the associated conservation laws.

To show this, we note that if we apply Noether's theorem to gauge symmetry, one finds that in the case of a gauge transformation the conserved four-current would be

$$j^\mu = \frac{\delta\mathcal{L}}{\delta\partial_\mu A_\nu}\delta A_\nu = -F^{\mu\nu}\partial_\nu\lambda(x) = -\partial_\nu\left(F^{\mu\nu}\lambda(x)\right),\tag{3.59}$$

where in the last step we have used the chain rule for differentiation and the fact that the equations of motion of the theory require $\partial_\mu F^{\mu\nu} = 0$. Therefore the *conserved charge* associated to a gauge transformation is:

$$Q = \int d^3x\, j^0 = \int d^3x\, \partial_i\left(F^{0i}\lambda(x)\right) = 0,\tag{3.60}$$

---

since it is a total derivative, and the photon fields vanish at infinity. Therefore, there are no conserved charges associated to local symmetries such as gauge invariance.

Despite this result, when we coupled A_μ to either a complex scalar or a fermion field, we obtained a non-trivial global symmetry, corresponding to the case of a constant gauge transformation. The application of Noether's theorem to the scalar field ψ in sQED Eq. (3.52), where $\delta\psi = -i\alpha\psi$, see Eq. (3.46), where now α is a constant, gives the following result

$$
\begin{aligned}
j_s^\mu &= \frac{\delta\mathcal{L}}{\delta\partial_\mu\psi}\delta\psi + \frac{\delta\mathcal{L}}{\delta\partial_\mu\psi^*}\delta\psi^* \\
&= i\alpha\left[\psi\partial^\mu\psi^* - \psi^*\partial^\mu\psi + i2eA^\mu\psi^*\psi\right],
\end{aligned}
\tag{3.61}
$$

while in the case of fermions the conserved current is instead

$$
j_f^\mu = \alpha\bar{\psi}\gamma^\mu\psi\,.
\tag{3.62}
$$

We observe that in both cases Noether's current j^μ is equal to the term the multiplies the photon field A_μ in the Lagrangian. This conserved current turns out to correspond the *electric charge* carried by the matter fields, as we will discuss in more detail below.

> Using Noether's theorem, we can illustrate how electric charge conservation is related to a specific case of gauge transformations.

3.3   Canonical quantization of the photon field

At this point we are ready to discuss the quantization of the gauge field A_μ using the canonical formalism. It is possible to also derive the same quantisation conditions using the path-integral formalism - the interested student can consult the textbooks for this. Also, since we are more interested in the implications of gauge symmetry to scattering amplitudes than to formal aspects, we will keep this discussion relatively light.

As we have mentioned, in the following we will l work in the Coulomb gauge, discussed in Sect. 3.1, and defined by Eq. (3.21). The Lagrangian of Maxwell's electromagnetism in the Coulomb gauge (in the absence of additional matter fields) is given by

$$
\mathcal{L} = \frac{1}{2}\dot{A}_i\dot{A}_i - \frac{1}{2}\partial_j A_i\partial_j A_i\,,
\tag{3.63}
$$

where we have used that $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$ and thus the time component of the four-potential vanishes, $A_0 = 0$, in this gauge. The canonical momenta of the theory are given by

$$
\begin{aligned}
\pi^0 &= \frac{\delta\mathcal{L}}{\delta\dot{A}_0} = 0\,, \\
\pi^i &= \frac{\delta\mathcal{L}}{\delta\dot{A}_i} = \dot{A}_i = \frac{\partial A_i}{\partial t}\,,
\end{aligned}
\tag{3.64}
$$

where we note that the fact that $\boldsymbol{\nabla}\cdot\mathbf{A} = \mathbf{0}$ also implies that $\boldsymbol{\nabla}\cdot\boldsymbol{\pi} = \mathbf{0}$ since the derivatives commute. The canonical momentum associated to A_0, π^0, also vanishes as a consequence of the Coulomb gauge-

fixing condition, implying that the spatial components are the only non-trivial ones. Now we can write the corresponding Hamiltonian density, finding

$$\mathcal{H} = \pi_i \dot{A}_i - \mathcal{L} = \frac{1}{2}\pi_i \pi_i + \frac{1}{2}\nabla_j A_i \nabla_j A_i \,, \tag{3.65}$$

in terms of the spatial components of the photon field A_i and their corresponding conjugate momenta π_i.

Commutation relations.   Next, in order to quantize the photon field, we need to impose suitable *canonical commutation relations* between the field and its conjugate momentum. A naive condition would be the following:

$$[A_i(\vec{x}, t), \pi_j(\vec{y}, t)] = i\delta_{ij}\delta^{(3)}(\vec{x} - \vec{y})\,, \tag{3.66}$$

but this option is not consistent with the gauge condition. To see this, one should note that in the Coulomb gauge we still want to have $\boldsymbol{\nabla} \cdot \mathbf{A} = \mathbf{0}$ at the level of operators (in addition to at the level of fields as is already the case in the classical picture), but if we take the divergence of Eq. (3.66) we find that

$$[\boldsymbol{\nabla} \cdot \mathbf{A}(\vec{x}, t), \boldsymbol{\nabla} \cdot \vec{\pi}(\vec{y}, t)] = i\nabla^2 \delta^{(3)}(\vec{x} - \vec{y}) \neq 0\,, \tag{3.67}$$

which is inconsistent with the Coulomb gauge fixing condition. So Eq. (3.66) is internally inconsistent since does not correspond to the same gauge in which we are quantising the theory.

The problem with the naive commutation relations of Eq. (3.66) is that we have not properly taken into account the *constraints from the gauge fixing condition*. Without deriving a formal proof, here we just state that the correct set of commutation relations for the gauge field are given by

$$[A_i(\vec{x}, t), \pi_j(\vec{y}, t)] = i\left(\delta_{ij} - \frac{\partial_i \partial_j}{\nabla^2}\right)\delta^{(3)}(\vec{x} - \vec{y})\,, \tag{3.68}$$

and we show that this commutation relation is consistent with the Coulomb gauge constraints. To prove this, we note that the RHS of Eq. (3.68) can be rewritten in momentum space as follows:

$$[A_i(\vec{x}, t), \pi_j(\vec{y}, t)] = i\int \frac{d^3p}{(2\pi)^3}\left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2}\right)\exp\left(i\vec{p}\cdot(\vec{x} - \vec{y})\right)\,, \tag{3.69}$$

and if we take a derivative of this commutation relation, we find that

$$\begin{aligned}[\partial_i A_i(\vec{x}, t), \pi_j(\vec{y}, t)] &= -\int \frac{d^3p}{(2\pi)^3}\left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2}\right)p_i \exp\left(i\vec{p}\cdot(\vec{x} - \vec{y})\right)\\ &= -\int \frac{d^3p}{(2\pi)^3}\left(p_j - p_j\right)\exp\left(i\vec{p}\cdot(\vec{x} - \vec{y})\right) = 0\,,\end{aligned} \tag{3.70}$$

and therefore the Coulomb gauge condition $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$ is still satisfied by the commutation relations Eq. (3.68), as we wanted to demonstrate. Therefore, the gauge fixing condition is valid both at the level of classical fields as well as the level of quantised operators as relevant for the QFT description.

Taking all this information into account, the complete set commutation relations for the gauge field in

the quantised version Maxwell's theory will be given by

$$
\begin{aligned}
[A_i(\vec{x}, t), \pi_j(\vec{y}, t)] &= i \left(\delta_{ij} - \frac{\partial_i \partial_j}{\nabla^2} \right) \delta^{(3)}(\vec{x} - \vec{y}) \,, \\
[A_i(\vec{x}, t), A_j(\vec{y}, t)] &= 0 \,, \\
[\pi_i(\vec{x}, t), \pi_j(\vec{y}, t)] &= 0 \,.
\end{aligned}
\tag{3.71}
$$

At this point, we are ready to use the solution for $\vec{A}$ in the Coulomb gauge found using the Klein-Gordon equation, Eq. (3.25), namely

$$
\mathbf{A}(\mathbf{x}, t) = \sum_{\lambda = \pm} \int \frac{d^3 k}{2\omega (2\pi)^3} \left[(\vec{\epsilon}_\lambda)^*(k) a_\lambda(k) e^{ikx} + \vec{\epsilon}_\lambda(k) a_\lambda^*(k) e^{-ikx} \right] \,,
\tag{3.72}
$$

and combine it with the commutation relations in Eq. (3.71). The idea is to be able to express the commutation relations for the photon field and its conjugate momentum in terms of the corresponding creating and annihilation operators. By combining this information, we end up with the following commutation relations for the creating and annihilation operators associated to the photon field with polarisation λ and momentum k^μ:

$$
\begin{aligned}
\left[a_\lambda(k), a_{\lambda'}(k') \right] &= 0 \,, \\
\left[a_\lambda^\dagger(k), a_{\lambda'}^\dagger(k') \right] &= 0 \,, \\
\left[a_\lambda(k), a_{\lambda'}^\dagger(k') \right] &= (2\pi)^3 2\omega \delta^{(3)}(k - k') \delta_{\lambda\lambda'} \,,
\end{aligned}
\tag{3.73}
$$

as can be checked explicitly, using the fact that the canonical momentum is given by

$$
\vec{\pi}(\mathbf{x}, t) = \frac{\partial \vec{A}}{\partial t} = \sum_{\lambda = \pm} \int \frac{d^3 k}{2\omega (2\pi)^3} k_0 \left[(\vec{\epsilon}_\lambda)^*(k) a_\lambda(k) e^{ikx} - \vec{\epsilon}_\lambda(k) a_\lambda^*(k) e^{-ikx} \right] \,,
\tag{3.74}
$$

where we have used Eq. (3.64). In a similar way as what is done with the quantization of the fermion field from the Dirac Lagrangian, we can physically identify the operators in Eq. (3.72) as follows:

- $a_\lambda^\dagger(k)$: *creation operator.*

  When applied to a given quantum state, $a_\lambda^\dagger(k)$ creates a mode of the photon field with momentum k and polarization λ.

- $a_\lambda(k)$: *annihilation operator.*

  When applied to a given quantum state, $a_\lambda(k)$ annihilates a mode of the photon field with momentum k and polarization λ.

The two physical polarizations of the photon field can be identified as:

- $\lambda = +$: right-circular polarization.

- $\lambda = -$: left-circular polarization.

Finally, we can write the Hamiltonian Eq. (3.65) in terms of the creation and annihilation operators to find

$$H = \sum_{\lambda=\pm} \int \frac{d^3k}{(2\pi)^3} \frac{1}{2} a_\lambda^\dagger(k) a_\lambda(k) + 2E_0 V \,, \tag{3.75}$$

where the zero point energy is defined as

$$E_0 = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3} \omega \,. \tag{3.76}$$

and V is the integration volume. As we know, the fact that the zero point energy is infinite is not relevant, since only differences in energy are physically meaningful.

3.4   Interactions in gauge theories: the scalar QED case

Following this discussion about the canonical quantization of the free photon field in electrodynamics, Eq. (3.72), we are now at a stage where we can formulate the full theory of quantum electrodynamics, the quantum field theory that governs the interactions between electrons and photons. This will require the derivation of Feynman rules that allow us eventually to compute the amplitudes for simple processes. Since the fermion case is a bit more complicated, to make our lives easier in this part of the course we will focus on the *scalar QED case*: the theory of the interactions between the charged scalar fields by means of photons. Recall that scalar QED at the classical level was introduced above in Eq. (3.51). This process will include several steps, with the ultimate goal of performing some simple tree-level computations of scattering processes using this theory.

From the formal point of view, the Feynman rules of both scalar QED and full QED can be derived in a similar way as those for Dirac's fermion and for the scalar field in $\lambda\phi^4$ theory. In particular, we need to use the Lehmann-Symanzik-Zimmermann (LSZ) reduction formula, which allows us to construct appropriate initial and final states for scattering experiments.

The LSZ reduction formula

The LSZ reduction formula allows us to construct appropriate initial and final states for scattering experiments. For a real scalar free field, the one-particle states are given by

$$|k\rangle = a^\dagger(\mathbf{k})|0\rangle\,, \tag{3.77}$$

where the creation operator in terms of the field $\varphi(x)$ is given by

$$a^\dagger(\mathbf{k}) = -i\int d^3x\, e^{ikx}\overline{\partial}_0\varphi(x)\,, \tag{3.78}$$

and it is normalized as follows

$$\langle k||k'\rangle = (2\pi)^3 2\omega\delta^{(3)}(\mathbf{k}-\mathbf{k}')\,. \tag{3.79}$$

In the interacting theory, the creation and annihilation operators will in general depend on the time t. Then a suitable choice for the initial and final states of scattering experiments is

$$|i\rangle = \lim_{t\to-\infty} a_1^\dagger(t)a_2^\dagger(t)|0\rangle\,, \qquad |f\rangle = \lim_{t\to+\infty} a_{1'}^\dagger(t)a_{2'}^\dagger(t)|0\rangle\,, \tag{3.80}$$

and one can show that using appropriately defined normalizations, the matrix element between the initial and final states will be given by in the general case of n incoming and n' outgoing particles

$$\langle i||f\rangle = i^{n+n'}\int d^4x_1 e^{ik_1 x_1}(-\partial_1^2 + m^2)\ldots d^4x_{1'} e^{ik_1' x_1'}(-\partial_{1'}^2 + m^2)\ldots \times \langle 0|T\left(\varphi(x_1)\ldots\varphi(x_1')\ldots\right)|0\rangle\,. \tag{3.81}$$

At this point we are ready to derive the LSZ reduction formula for photons. In the Coulomb gauge, the spatial components of the photon field, $\vec{A}$, are given by Eq. (3.72), where the gauge field polarization vectors $\vec{\epsilon}_\lambda$ being perpendicular to the direction of propagation, $\vec{k}\cdot\vec{\epsilon}_\lambda(k) = 0$, and satisfying the orthonormality conditions of Eq. (3.35). In the same way as what is done for scalars and fermions, it is possible to invert the relation between the creation and annihilation operators and the photon field itself to yield

$$
\begin{aligned}
a_\lambda(k) &= i\vec{\epsilon}_\lambda(k)\cdot\int d^3x\, e^{-ikx}\overline{\partial}_t\vec{A}(x)\,,\\
a_\lambda^\dagger(k) &= -i\vec{\epsilon}_\lambda^*(k)\cdot\int d^3x\, e^{ikx}\overline{\partial}_t\vec{A}(x)\,,
\end{aligned} \tag{3.82}
$$

where we have defined the following symbol

$$f\overline{\partial}g \equiv f(\partial g) - (\partial f)g\,. \tag{3.83}$$

To verify these properties, you can plug Eq. (3.82) in Eq. (3.72) can check that the identity is satisfied. From here, the LSZ framework applies in the same way as in the scalar and fermion cases. The only subtlety here

is that Note that for an incoming and an outgoing photon we need to make the corresponding replacements:

$$
\begin{aligned}
a_\lambda^\dagger(k)_{\text{in}} &= i\vec{\epsilon}_\lambda^*(k) \cdot \int d^4x\, e^{+ikx}(-\partial^2)\vec{A}_\mu(x)\,, \\
a_\lambda^\dagger(k)_{\text{out}} &= i\vec{\epsilon}_\lambda(k) \cdot \int d^4x\, e^{-ikx}(-\partial^2)\vec{A}_\mu(x)\,.
\end{aligned}
\tag{3.84}
$$

Feynman rules for scalar QED.   We now have all the ingredients that we need to derive the Feynman rules for scalar QED. The full Lagrangian of the theory is given by

$$
\begin{aligned}
\mathcal{L} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - D^\mu\psi^\dagger D_\mu\psi - m^2\psi^\dagger\psi \\
&= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \partial^\mu\psi^*\partial_\mu\psi - m^2\psi^*\psi + ieA^\mu\left[(\partial_\mu\psi^*)\,\psi - \psi^*\,(\partial_\mu\psi)\right] - e^2 A_\mu A^\mu\psi^*\psi\,,
\end{aligned}
\tag{3.85}
$$

where we have expanded the covariant derivatives D_μ to make explicit the couplings between the vector field A_μ and the complex scalar field ψ. From this expression, we see that there are two *interaction terms*, meaning two terms that couple the photon with the scalar field. As we have seen, the structure of these terms is fixed by the gauge invariance requirements of the theory. These terms lead to the two interaction vertices shown in Fig. 3.1:

- The first interaction corresponds to a vertex involving two scalars and one photon (that is, a three-point vertex), given by

$$
ieA^\mu\left[(\partial_\mu\psi^*)\,\psi - \psi^*\,(\partial_\mu\psi)\right]\,,
\tag{3.86}
$$

  which involves also the derivatives of the scalar field, and is proportional to the electric charge e of the scalar field.

  Note that this term corresponds to a *derivative coupling*, namely a coupling terms which involves derivatives over the scalar field ψ. As we will show later, this property implies a Feynman rule that depends on the momenta of the scalar fields that enter in this interaction vertex.

  Note also that this term is different from the fermion-fermion-photon coupling that appears in QED, since there the coupling does involve derivatives of the fermion field (as one can determine from purely dimensional reasons).

  In this vertex, the coupling constant e is dimensionless, as can be verified from dimensional counting.

- The second interaction term is a four-point vertex

$$
- e^2 A_\mu A^\mu\psi^*\psi\,,
\tag{3.87}
$$

  and leads to a vertex with two photons and two scalar fields, with no dependence on neither the photon or the scalar field momenta.

  Note that such four-point vertex is absent in QED, and is thus a genuine new feature of its scalar version. Note also that in this case the coupling constant associated to the vertex is dimensionless.

The diagrams in Eq. (3.1) are the only possible ways that photons and scalars can interact being consistent with the requirements of gauge symmetry: as discussed above, other possible interaction terms would

$$ieA^\mu \left(\partial_\mu \psi^* \psi - \psi^* \partial_\mu \psi\right) \qquad\qquad -e^2 A_\mu A^\mu \psi^* \psi$$
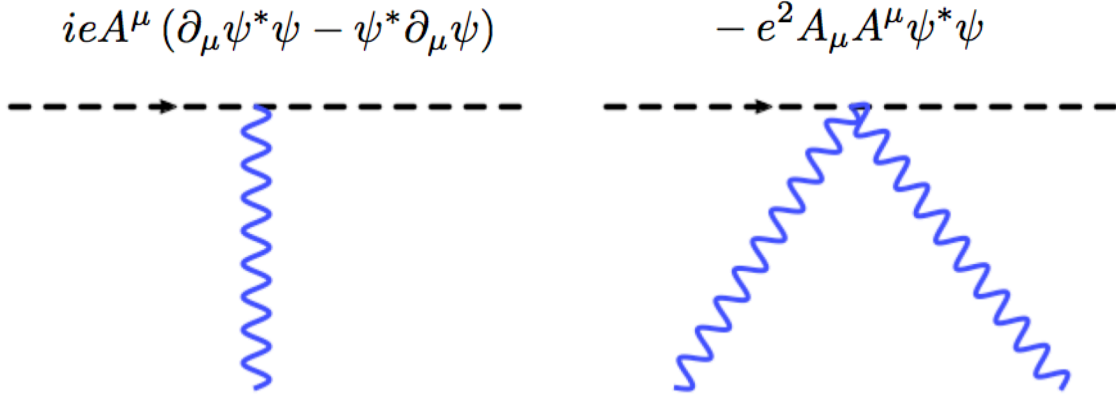
Figure 3.1. The photon-scalar interaction vertices in scalar QED, see Eq. (3.85) and the corresponding discussion in the text more more details.

not satisfy gauge invariance. Interestingly, from dimensional reasons one could have considered the two interaction terms separately each with an associated coupling constant,

$$ig_1 A^\mu \left[\left(\partial_\mu \psi^*\right) \psi - \psi^* \left(\partial_\mu \psi\right)\right] , \qquad -g_2 A_\mu A^\mu \psi^* \psi , \tag{3.88}$$

but if we impose gauge invariance to be a symmetry of our theory then these two coupling constants cannot be independent and must satisfy the relation $g_1^2 = g_2$.

From the discussion of the symmetry properties of scalar QED in Sect. 3.2, we found that a global gauge transformation had associated a conserved current (via Noether's theorem) of the form of Eq. (3.61), which can also be expressed in terms of the covariant derivative as

$$j^\mu = i \left(\psi^* D^\mu \psi - (D^\mu \psi)^* \psi\right) . \tag{3.89}$$

Note that in sQED the electric current also depends on the photon field itself, as a consequence of the $\psi^* \psi A_\mu A^\mu$ coupling. The opposite sign between the first and second term implies that the charged scalar field ψ contains states with both *positive* and *negative charge*, as opposed to states with only one type of charge. With some abuse of notation we can write the scalar field in terms of creation and annihilation operators as

$$\psi(\mathbf{x}, t) = \int \frac{d^3 p}{(2\pi)^3 2\omega} \left(a_s(k)e^{ikx} + b_s^\dagger(k)e^{-ikx}\right) , \tag{3.90}$$

so the state created by $a_s^\dagger$ will be called the "positron" e^+ and the state created by $b_s^\dagger$ will be called the "electron" e^-. Needless to say, electron and positron are the anti-particles of each other. Recall that ψ represents scalar fields in sQED, so it would be more appropriate to denote the particles created by it "scalar electron" and "scalar positron".

Combining the various ingredients together, is it possible to derive the Feynman rules of scalar QED allowing to construct the corresponding S matrix elements. Skipping due to lack of time their formal derivation, here we state the results and then study that are their implications at the level of scattering amplitudes.

- For an incoming positron or electron e^+ or e^-, one uses the same rules as in the free scalar theory.

- For an incoming photon: add a factor of $\epsilon_{\lambda_i}^{*\mu}(k_i)$

- For an outgoing photon: add a factor of $\epsilon_{\lambda_i}^{\mu}(k_i)$

- For the interaction vertices, using the sign convention in Fig. 3.2, we add for the $\psi\psi^* A$ vertex a factor

$$ie(k_1 + k_2)_\mu (2\pi)^4 \delta\left(\sum k_i\right) , \tag{3.91}$$

while for the $\psi\psi^* A^2$ vertex we have instead to include a term

$$- 2ie^2 \eta_{\mu\nu} (2\pi)^4 \delta\left(\sum k_i\right) , \tag{3.92}$$

accounting for all relevant symmetry factors, and where the delta functions ensure momentum propagation at the vertices. The Lorentz indices need to be properly contracted either with a photon propagator, another photon vertex, or with the polarization vectors of an incoming/outgoing photon, ensuring that the final scattering amplitude is a Lorentz scalar.

In the Feynman rule for the three point vertex, Eq. (3.91), the dependence on the momenta $(k_1 + k_2)$ of the scalar fields is a direct consequence of the derivative nature of the corresponding interaction term in the Lagrangian, Eq. (3.86).

Note that the second interaction term is formally suppressed by a factor e as compared to the first one. On the other hand, the three-point vertex becomes suppressed where the momenta of the incoming scalar fields is very small.

- For the internal lines (propagators) we have

$$\frac{-i}{k^2 + i\epsilon} \left(\eta^{\mu\nu} - (1 - \xi)\frac{k_\mu k_\nu}{k^2}\right) , \tag{3.93}$$

for the photon propagator, and

$$\frac{-i}{k^2 + m^2 - i\epsilon} , \tag{3.94}$$

for the scalar electron (positron) propagator.

The photon propagator has been derived in the generalised R_ξ gauge. As will be discussed below, this gauge is a generalisation of the Lorenz gauge that has a number of advantages. The factor ξ is a parameter of this gauge fixing criterion, and therefore scattering amplitudes should not depend on it.

- Integrate over all internal (loop) momenta with a factor $\int d^4k/(2\pi)^4$.

- Sum over all diagrams that contribute to the same final and initial state, including the relevant averages in the initial state case.

In most cases, we will not be interested in the polarisation of the initial and final state photons. In these cases, we need to *average* over all possible initial state photon polarisations, and *sum* over all possible final state photon polarisations. When performing such calculations, the expressions in Eq. (3.35) allow simplifying the scattering amplitudes by replacing the polarization vectors by delta functions and powers of the photons' momenta k^μ.

$$ie(k_1 + k_2)_\mu (2\pi)^4 \delta \left(\sum k_i \right) \qquad - 2ie^2 \eta_{\mu\nu} (2\pi)^4 \delta \left(\sum k_i \right)$$
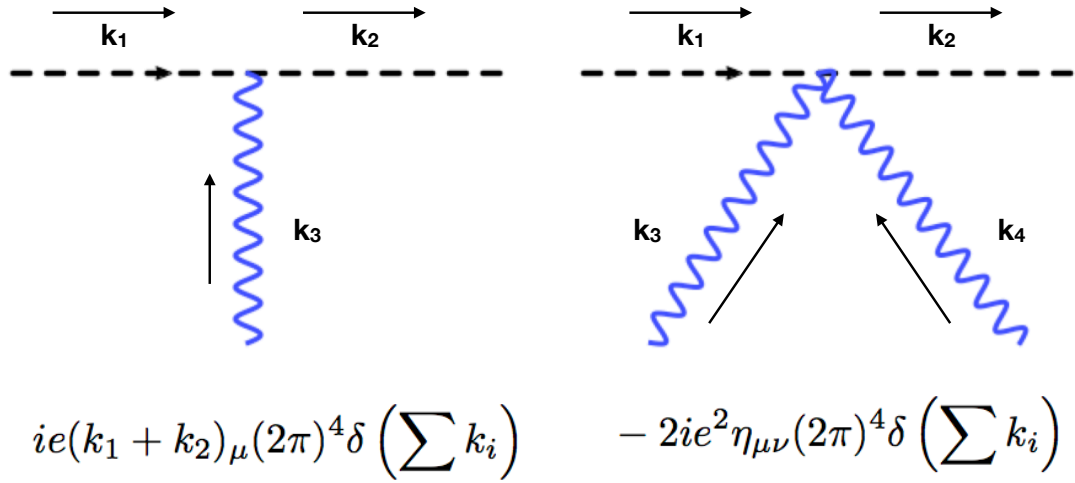
Figure 3.2. The Feynman rules for the interaction vertices in scalar QED.

With these Feynman rules, it is possible to evaluate any process in scalar QED, first at tree level, and then eventually at the loop level using a generalization of the renormalization procedure that has been described in Sect. 1 in the case of $\lambda \phi^4$ theory. Indeed, no major conceptual differences arise when computing one-loop diagrams in scalar QED as compared to what we have studied for $\lambda \phi^4$ theory.

The R_ξ gauge. The photon propagator in Eq. (3.93) has been derived using the so-called *generalized Feynman gauge*, also known as the R_ξ gauge. This family of gauges are a generalisation of the Lorenz gauge that we have been using so far, namely $\partial_\mu A^\mu = 0$. Setting $\xi = 1$ leads to the *Feynman gauge*, while setting $\xi = 0$ gives the Lorenz gauge (also known as the *Landau gauge*). Given that the choice for the value of ξ is gauge-dependent, it should drop from all physical quantities, such as scattering amplitudes. We will demonstrate that this is actually the case in explicit calculations.

The basic idea of the R_ξ gauge fixing method is to replace the auxiliary equation that determines the specific gauge condition by adding an explicitly gauge-fixing term to the Lagrangian

$$\Delta \mathcal{L} = -\frac{(\partial_\mu A^\mu)^2}{2\xi} \,. \tag{3.95}$$

If we take the limit $\xi \to 0$ the recover the Lorenz gauge, since then minimizing the action is equivalent to requiring $\partial_\mu A^\mu = 0$. In this case, the expression for the photon propagator in the Lorentz is given by

$$\frac{-i}{k^2 + i\epsilon} \left(\eta^{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right) \,, \tag{3.96}$$

which is related to the commutation relations for the photon that were derived above. For calculational reasons, the Feynman gauge $\xi = 1$ is most useful, the reason is that there the second term of the propagator cancels out and one ends up with a simplified expression for the photon propagator.

$$\frac{-i\eta^{\mu\nu}}{k^2 + i\epsilon} \,. \tag{3.97}$$

In the lectures we will keep in most cases the explicit dependence on the gauge fixing parameter ξ in our calculations of scattering amplitudes. The reason is that this way we can show explicitly how terms proportional to ξ cancel out in the final result, demonstrating that indeed the calculation does not depend on the specific gauge choice that has been made.

3.5   Scattering processes in scalar QED.

As a first application of these Feynman rules, let us calculate one representative tree-level scattering amplitude in scalar QED, the so-called *Moller scattering* process, namely the scattering between two electrons:

$$ e^- + e^- \to e^+ - e^- \,. \tag{3.98} $$

The two diagrams that contribute to this processes are those shown in Fig. 3.3, which correspond to t-channel and u-channel scattering respectively. Applying the Feynman rules for this theory, we see that the amplitude for this processes at tree level is obtained by adding the contribution of the s-channel and t-channel diagrams yielding

$$
\begin{aligned}
\mathcal{M} &= \left[ie(k_1 + k_3)_\mu \frac{-i}{p^2} \left(\eta^{\mu\nu} - (1-\xi)\frac{p^\mu p^\nu}{p^2} \right) ie(k_2 + k_4)_\nu \right. \\
&+ \left. ie(k_1 + k_4)_\mu \frac{-i}{\widetilde{p}^2} \left(\eta^{\mu\nu} - (1-\xi)\frac{\widetilde{p}^\mu \widetilde{p}^\nu}{\widetilde{p}^2} \right) ie(k_2 + k_3)_\nu \right] (2\pi)^4 \delta(k_1 + k_2 - k_3 - k_4) \,.
\end{aligned}
\tag{3.99}
$$

with $p = k_1 - k_3$ and $\widetilde{p} = k_1 - k_4$ for the $t-$ and $s-$ channel respectively, and the overall delta functions ensure global momentum conservation between the initial and final state particles.

Note also that here the $+i\epsilon$ prescription in the photon propagator Eq. (3.93) is not needed since the kinematics of the process ensures that the photon will always be *off-shell*. We can also observe how the dependence on the external momenta of the scalar electron and position appears in the scattering amplitude as a consequence of their derivative coupling with the photon field. As mentioned above, for this type of calculations it is instructive to maintain the explicit dependence on the gauge-fixing parameter ξ to demonstrate that it cancels out at the end of the calculation. The way this cancellation takes place is also helpful to illustrates the constraints that gauge invariance imposes at the level of scattering amplitudes.

We can further simplify this expression by noting that for the t-channel diagram we have that

$$ p^\mu (k_1 + k_3)_\mu = (k_1 - k_3)^\mu (k_1 + k_3)_\mu = k_1^2 - k_3^2 = m^2 - m^2 = 0 \,, \tag{3.100} $$

since both scalar electrons have the same mass, and similarly for the u-channel diagram. Taking this property into account, the scattering amplitude simplifies to

$$ \mathcal{M}\left(e^- e^- \to e^- e^- \right) = e^2 \frac{(k_1 + k_3)_\mu (k_2 + k_4)^\mu}{t} + e^2 \frac{(k_1 + k_4)_\mu (k_2 + k_3)^\mu}{u} \,, \tag{3.101} $$

where as usual the Mandelstam variables are defined as

$$ s \equiv -(k_1 + k_2)^2 \,, \quad t \equiv -(k_1 - k_3)^2 \,, \quad u \equiv -(k_1 - k_4)^2 \,, \tag{3.102} $$

since we are using the $(-+++)$ signature for the Minkowski metric.
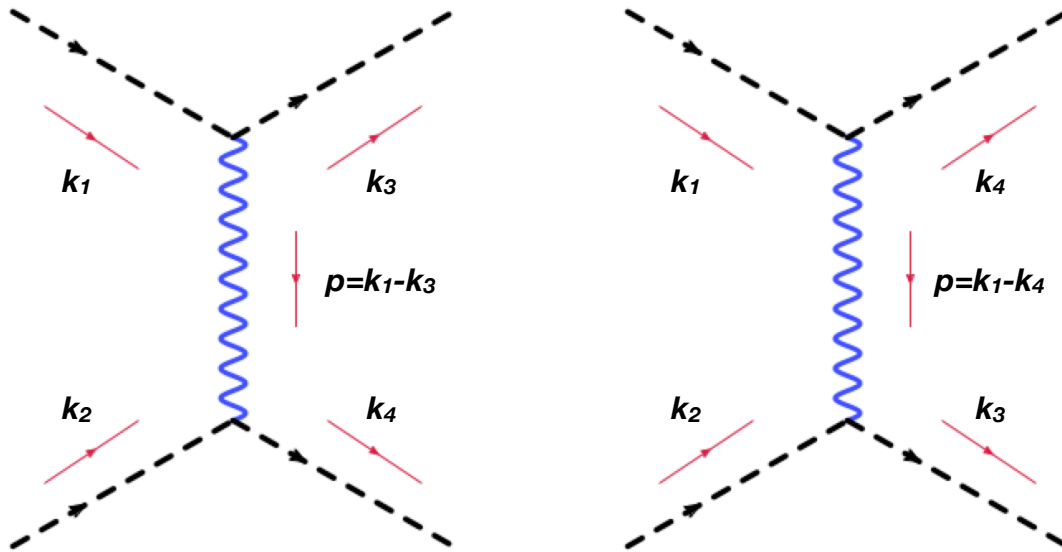
Figure 3.3. Feynman diagrams relevant for the calculation of Moeller scattering, $e^- e^- \to e^- e^-$, at tree level in scalar QED. See text for more details.

> As predicted, the gauge parameter ξ of the photon propagator does not appear in the final result Eq. (3.101). If it had appeared, it would mean that the result depends on the specific choice of gauge, which is of course unphysical. The cancellation takes place due to the specific form of the interaction vertices, which as we have seen have an structure which is precisely dictated by gauge invariance.

From the result Eq. (3.101) we see that the amplitude is the largest when either t or s are small: we say that Moeller scattering in scalar QED has both a *t-channel and a u-channel singularity*. By squaring this scattering amplitude and integrating over the available phase space, we would find the final result for the cross-section of this process.

Similar results apply for related scattering processes: the key message here is that the term which depends on the gauge fixing always cancels out, as required by the gauge invariance of the theory.

3.6   Ward identities in scalar QED

In the Coulomb gauge, we found that the photon polarization vectors are perpendicular to the photon propagation direction, $\vec{k} \cdot \vec{\epsilon}_\lambda(k) = 0$, and satisfied the orthonormality conditions in Eq. (3.35). However, upon a closer inspection, one realizes that there might be an inconsistency somewhere in our derivation. To see this, consider the following reasoning. Assume that now we carry out a Lorentz transformation and change our reference frame, then the four-momentum k^μ associated to the photon is transformed according to the Lorentz rotation matrix:

$$\Lambda^\mu_\nu k^\nu = k'^\mu \,. \tag{3.103}$$

On the other hand, in general a Lorentz transformation of the polarization states will mix the physical polarization states with the momentum (since the polarisation four-vectors do not form a closed group under Lorentz transformations) and we find:

$$\Lambda^{\mu}_{\nu}\epsilon^{\nu}_{+} = \epsilon^{'\nu}_{+} = c_1\epsilon^{\nu}_{+}(k^{'\mu}) + c_2\epsilon^{\nu}_{-}(k^{'\mu}) + c_3 k^{'\mu}\,, \tag{3.104}$$

where the coefficients c_i depend on the specific choice of Lorentz transformation matrix Λ. Therefore, a Lorentz transformation of the photon polarisation vector cannot be written as a linear combination of the two photon polarisation vectors ϵ^{μ}_{λ}.

To illustrate why this is a problem, consider the matrix element that we would obtain from computing the emission of a photon from an arbitrary system composed by photons and scalar particles. Given the Feynman rules of the theory, we know that the matrix element will look like

$$\mathcal{M} = \epsilon_{\mu}M^{\mu}\,, \tag{3.105}$$

since the photon is outgoing, and where ϵ_{μ} is some linear combination of $\epsilon_{\pm}$ and M_{μ} transforms as a vector under a Lorentz transformation

$$M_{\mu} \rightarrow M'_{\mu} = \Lambda^{\nu}_{\mu}M_{\nu}\,. \tag{3.106}$$

Now, a Lorentz transformation of the matrix element $\mathcal{M}$ will lead to

$$\mathcal{M} \rightarrow \left(c_1\epsilon^{\mu}_{+} + c_2\epsilon^{\mu}_{-} + c_3 k^{'\mu}\right)M'_{\mu}\,, \tag{3.107}$$

which is a problem, since:

> the new polarization vector is *not* part of the physical Hilbert space of the theory, where only linear combinations of $\epsilon^{\mu}_{\pm}$ are allowed. Therefore it seems that upon a Lorentz transformation we end up with an unphysical state.

To avoid this problem, one would need to ensure that the product $\epsilon_{\mu}M^{\mu}$ in a given reference frame becomes $\epsilon'_{\mu}M^{'\mu}$ is some other reference frame, with both ϵ_{μ} and ϵ'_{μ} being linear combinations of the physical polarisation vectors. In other words, we want that the structure $\epsilon_{\mu}M^{\mu}$ of physical matrix elements holds in any reference frame. In order to ensure this condition, namely that

$$\mathcal{M} = \epsilon_{\mu}M^{\mu} \rightarrow \epsilon'_{\mu}M^{'\mu}\,, \tag{3.108}$$

with $\epsilon'_{\mu} = c_1\epsilon^{\mu}_{+} + c_2\epsilon^{\mu}_{-}$ a physical polarisation vector, we need to impose that

$$\mathcal{M}^{'\mu}k'_{\mu} = \mathcal{M}^{\mu}k_{\mu} = 0\,, \tag{3.109}$$

a property which is known as the *Ward identity* and that is an important feature of Quantum Electrodynamics (both with scalars and fermions as matter particles). The validity of Ward identities is closely related both for the Lorentz and to the gauge invariance of (s)QED. Also, while here we focus mostly on tree-level processes, it can be shown that Eq. (3.109) also holds when loop corrections are included. In other words, the Ward

identities need to be satisfied by scattering amplitudes at all orders in perturbation theory.

In order to highlight the interplay between the Ward identities and the gauge invariance of our theory, one should recall that the propagator for the photon $\Delta_{\mu\mu}(k)$ in (scalar) QED Eq. (3.93) included an explicit dependence on the unphysical gauge fixing parameter ξ,

$$i\Delta_{\mu\nu}(k) = \frac{-i}{k^2 + i\epsilon} \left(\eta_{\mu\nu} - (1 - \xi)\frac{k_\mu k_\nu}{k^2} \right) , \tag{3.110}$$

which should cancel out in physical scattering amplitudes due to the gauge symmetry of the theory. Since a general matrix element that contains a photon propagator must, due to the Lorentz invariance requirement, have the general form

$$\mathcal{M} = \mathcal{M}_{\mu\nu}\Delta^{\mu\nu}(k) , \tag{3.111}$$

then we see that gauge invariance, which demands the independence of $\mathcal{M}$ on the value of ξ, must lead to the following condition:

$$\mathcal{M}_{\mu\nu}k^\mu k^\nu = 0 , \tag{3.112}$$

which is a property closely related to the original Ward identity in Eq. (3.109), showing how these have deep connections with the gauge invariance of the theory.

Ward identities in sQED

- In a general scattering amplitude $\mathcal{M}$ with an outgoing photon with polarisation λ and momentum k^μ, $\mathcal{M} = \mathcal{M}_\mu \epsilon^\mu_\lambda(k)$, replacing the polarisation vector by the corresponding momentum leads to $\mathcal{M}_\mu k^\mu = 0$.

- In a general scattering amplitude $\mathcal{M}$ with an internal photon propagator, $\mathcal{M} = \mathcal{M}_{\mu\nu}\Delta^{\mu\nu}(k)$, replacing the propagator by the product $k^\mu k^\nu$ leads to the vanishing of the scattering amplitude, $\mathcal{M}_{\mu\nu}k^\mu k^\nu = 0$.

In addition, since we have not made any assumption about the perturbative order at which the scattering amplitudes $\mathcal{M}$ are being computed, it follows that these Ward identities will be valid both at tree level and at the loop order: these relations need to be obeyed separately for each term in the perturbative expansion in the electron charge.

Let us now illustrate the implications of the Ward identify with an explicit calculation. We consider electron-positron scattering (or alternatively, $\pi^+ + \pi^-$ scattering) into a pair of photons in scalar QED, namely

$$e^- + e^+ \to \gamma + \gamma . \tag{3.113}$$

At the tree level, this process receives contributions from both t-channel and u-channel diagrams, as represented in the Feynman diagram of Fig. 3.4. We now want to compute these diagrams and verify that indeed the Ward identities defined above are applicable.

To being with, starting with the t-channel diagram and applying the Feynman rules of scalar QED, we
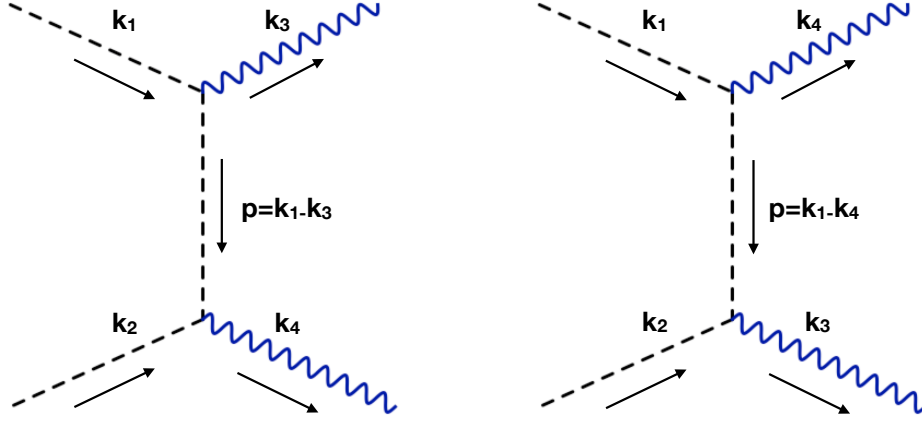
Figure 3.4. Feynman diagrams relevant for the calculation of the tree level scattering process $e^- + e^+ \to \gamma + \gamma$ in scalar QED. See text for more details.

find that the amplitude reads

$$i\mathcal{M}_t = ie^2 \frac{(2k_1 - k_3)_\mu (k_4 - 2k_2)_\nu}{(k_1 - k_3)^2 + m^2} \epsilon_3^{*\mu} \epsilon_4^{*\nu}, \tag{3.114}$$

that after some algebra and imposing the mass-shell condition $k_1^2 = -m^2$ for the scalar electrons can be written as

$$i\mathcal{M}_t = -ie^2 \frac{(k_3 \cdot \epsilon_3^* - 2k_1 \cdot \epsilon_3^*)(k_4 \cdot \epsilon_4^* - 2k_2 \cdot \epsilon_4^*)}{k_3^2 - 2k_1 \cdot k_3}, \tag{3.115}$$

with the corresponding u-channel diagram giving a closely related result

$$i\mathcal{M}_u = -ie^2 \frac{(k_4 \cdot \epsilon_4^* - 2k_1 \cdot \epsilon_3^*)(k_3 \cdot \epsilon_3^* - 2k_2 \cdot \epsilon_4^*)}{k_3^2 - 2k_2 \cdot k_3}. \tag{3.116}$$

The full matrix element is of course given by the sum of the two diagrams,

$$\mathcal{M}\left(e^- + e^- \to \gamma + \gamma\right) = \mathcal{M}_t\left(e^- + e^- \to \gamma + \gamma\right) + \mathcal{M}_u\left(e^- + e^- \to \gamma + \gamma\right). \tag{3.117}$$

For this amplitude to be physically consistent, as discussed above, it needs to satisfy the Ward identity, Eq. (3.109). From the practical point of view, this corresponds to replacing the polarization vector ϵ_3^* with the corresponding four-momentum k_3 (or the same with $3 \to 4$). Doing this replacement, some terms cancel out between the numerator and the denominator, and we find

$$(\mathcal{M}_t + \mathcal{M}_u)_{\epsilon_3 \to k_3} = -e^2 \left[k_4 \cdot \epsilon_4^* - 2k_2 \cdot \epsilon_4^* + k_4 \cdot \epsilon_4^* - 2k_1 \cdot \epsilon_4^*\right] = -2e^2 \left[k_4 - k_2 - k_1\right] \times \epsilon_4^* \neq 0, \tag{3.118}$$

which is different from zero, and thus the Ward identity is not satisfied. Does it mean that the theory is internally inconsistent? Or perhaps we have missed some ingredient in the calculation?

The reason why the Ward identity seems to fail is that the diagrams shown in Fig. 3.4 are not the only diagrams that contribute to the $e^- e^+ \to \gamma\gamma$ scattering amplitude at this order in the perturbative expansion. Indeed, if we aim to compute the scattering amplitude for this process consistently at $\mathcal{O}(e^2)$, we need to
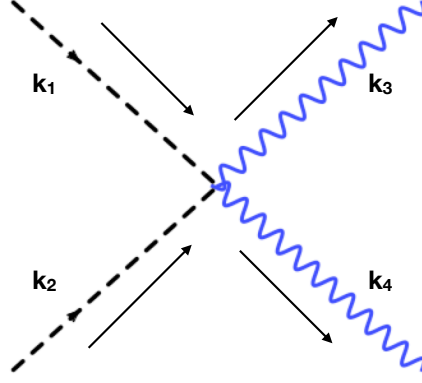
Figure 3.5. The other Feynman diagram that contributes to the scattering amplitude of the $e^- e^- \to \gamma\gamma$ in scalar QED at $\mathcal{O}\left(e^2\right)$, in addition to the two diagrams already shown in Fig. 3.5.

also take into account the interaction vertex with two gauge fields and two scalar fields (right diagram in Fig. 3.2), which contributes at the same order in the expansion in e^n. Using the convention for the sign of the momenta shown in Fig. 3.5 and the corresponding Feynman rule, the amplitude corresponding to this additional diagram turns out to be

$$i\mathcal{M}_4 = -2ie^2\eta_{\mu\nu}\epsilon_3^{*\mu}\epsilon_4^{*\nu}\,, \tag{3.119}$$

and doing the same replacement as above we get

$$i\mathcal{M}_4\Big|_{\epsilon_3 \to k_3} = -2ie^2\eta_{\mu\nu}k_3^{\mu}\epsilon_4^{*\nu} = -2ie^2 k_3 \cdot \epsilon_4^*\,. \tag{3.120}$$

Now if we add the contributions from the three diagrams we get

$$(\mathcal{M}_t + \mathcal{M}_u + M_4)_{\epsilon_3 \to k_3} = -2e^2\left[k_4 + k_3 - k_2 - k_1\right] \times \epsilon_4^* = 0\,, \tag{3.121}$$

due to overall momentum conservation, since $k_1 + k_2 = k_3 + k_4$. So reassuringly enough, the Ward identity is explicitly verified, as expected from the general symmetries of the theory, in particular from the Lorenz and gauge invariance requirements. One can check that Ward identities also hold for all tree-level processes in scalar QED, as well as at higher orders as well.

Ward identities with internal photon lines. Crucially, the above *derivation* of the Ward identities did not require at any point to assume that the outgoing photons were *on-shell* and thus *massless*. indeed, at no point we have used that $k_3^2 = 0$ or $k_4^2 = 0$, which would imply that the outgoing photons are on-shell. For example, in Eq. (3.115) we did not set $k_3^2 = 0$, and therefore our results are also valid when the photon lines are virtual (off-shell). This has the very important consequence that Ward identities should be valid even if the photon is off-shell, for example, a photon in a *loop* or a photon appearing as an *internal propagator* in a Feynman diagram.

In order to give an explicit example of the second of the relations relation introduced above related to the Ward identities, Eq. (3.112), which as we showed was directly related to the gauge invariance of the theory, we need to consider diagrams with internal photons (since only there the photon propagator, including the
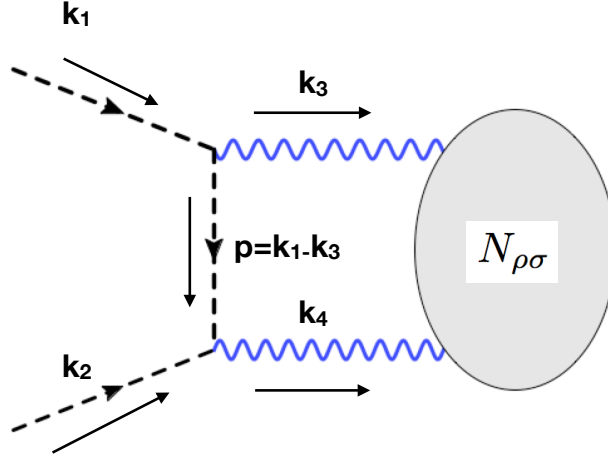
Figure 3.6. The Feynman diagram for a similar scattering amplitude as that of Fig. 3.4 but where now the two photons are internal lines (so off-shell) rather than outgoing particles, and where we denote in full generality the rest of the diagram by $N^{\rho\sigma}$.

gauge fixing parameter ξ, appears). In order to achieve this, let us consider the diagram in Fig. 3.6, which is similar to that of $e^- e^+ \to \gamma\gamma$ scattering in Fig. 3.4 but now the two photons are internal lines, and we denote by $N_{\rho\sigma}$ the remainder of the matrix element. We will keep $N_{\rho\sigma}$ unspecified since the following derivation does not rely on its specific form. Note that $N_{\rho\sigma}$ could include both external lines and loops, showing the general results we are aiming to demonstrate here related to the gauge invariance of the theory independent of the specific perturbative order of the calculation.

Let us now compute the loop diagram in Fig. 3.6 by using the Feynman rules of scalar QED. Recall that the two photons in the diagram, with momenta k_3 and k_4 respectively, are internal lines and thus we need to integrate over all their possible values (in other words, these momenta are unspecified), taking into account overall momentum conservation. Then one finds the following result for the scattering amplitude:

$$\mathcal{M} = e^2 \int \frac{d^4 k_3}{(2\pi)^4} \int \frac{d^4 k_4}{(2\pi)^4} \delta^{(4)}(k_1 + k_2 - k_3 - k_4) \frac{(k_3^\mu - 2k_1^\mu)(k_4^\nu - 2k_2^\nu)}{k_3^2 - 2k_3 \cdot k_1} \Delta_{\mu\rho}(k_3) \Delta_{\nu\sigma}(k_4) N^{\rho\sigma} . \quad (3.122)$$

This result is clearly not gauge invariant, since contracting the momentum dependence from the photon vertices with the propagators does not cancel out the terms proportional to ξ, the gauge fixing terms.

Of course, this is per se not surprising: individual diagrams do not need to be gauge invariant, and only the sum of all diagrams that contribute to a physical scattering amplitude has to be gauge invariant. This is the same property as that illustrated by the previous example: only when adding the two diagrams of Fig. 3.4 with that of Fig. 3.5 the Ward identify was satisfied. In this case, we need to supplement the diagram of Fig. 3.6 with the corresponding versions of the diagrams in Fig. 3.4 (right) and Fig. 3.5 where the photons being internal lines attached to the rest of the diagram, which is denoted generically by $N^{\rho\sigma}$.

Accounting for these two additional diagrams, the result of Eq. (3.122) now reads

$$
\begin{aligned}
\mathcal{M} &= e^2 \int \frac{d^4 k_3}{(2\pi)^4} \int \frac{d^4 k_4}{(2\pi)^4} \delta^{(4)}(k_1 + k_2 - k_3 - k_4) \left[\frac{(k_3^\mu - 2k_1^\mu)(k_4^\nu - 2k_2^\nu)}{k_3^2 - 2k_3 \cdot k_1} \right. \\
&+ \left. \frac{(k_3^\mu - 2k_2^\mu)(k_4^\nu - 2k_1^\nu)}{k_3^2 - 2k_3 \cdot k_2} + 2\eta^{\mu\nu} \right] \Delta_{\mu\rho}(k_3) \Delta_{\nu\sigma}(k_4) N^{\rho\sigma} .
\end{aligned}
\tag{3.123}
$$

Now, the term which is not gauge invariance in the photon propagator $\Delta_{\mu\nu}(k_3)$ is proportional to $k_3^\mu k_3^\nu$, recall the result for the propagator in the R_ξ gauge that we are using, Eq. (3.93). Now replacing $\Delta_{\mu\nu}(k_3)$ by $k_3^\mu k_3^\nu$ in Eq. (3.123), in order to isolate the gauge-dependent contribution to the total scattering amplitude, leads to the following result:

$$
\begin{aligned}
\mathcal{M}_{\text{gauge-dep}} &\propto e^2 \int \frac{d^4 k_3}{(2\pi)^4} \int \frac{d^4 k_4}{(2\pi)^4} \delta^{(4)}(k_1 + k_2 - k_3 - k_4) \left[\frac{(k_3^\mu - 2k_1^\mu)(k_4^\nu - 2k_2^\nu)}{k_3^2 - 2k_3 \cdot k_1} \right. \\
&+ \left. \frac{(k_3^\mu - 2k_2^\mu)(k_4^\nu - 2k_1^\nu)}{k_3^2 - 2k_3 \cdot k_2} + 2\eta^{\mu\nu} \right] k_3^\mu k_3^\nu \Delta_{\nu\sigma}(k_4) N^{\rho\sigma} = 0 .
\end{aligned}
\tag{3.124}
$$

which vanishes due to the overall four-momentum conservation imposed by the delta function, as can be shown explicitly. This exercise demonstrates that:

> Irrespective of the specific form that the remainder of the diagram $N^{\rho\sigma}$ has in the scattering reaction illustrated by Fig. 3.6, the part of the diagram with the two internal photon lines will be independent of the gauge fixing parameter ξ. It can be checked using higher order diagrams that this property always holds, as required from the Lorenz and gauge invariance of the theory.

3.7   Lorentz invariance and charge conservation

Moving forward in our study of the features and properties of quantum electrodynamics, we would like now to discuss a rather elegant illustration of the direct connection between Lorentz invariance and electric charge conservation, which is independent of the gauge symmetry requirements. This can be achieved by showing:

> a theory that includes a massless spin-one particle, such as QED or scalar QED, has associated automatically a *conserved charge*, irrespective of whether or not gauge symmetry is present.

To demonstrate this remarkable property, consider a scattering amplitude $\mathcal{M}_0(\{p_i\})$ with n external legs (which can be either incoming or outgoing) in scalar QED, such as that represented in Fig. 3.7. These legs correspond to a fermion field in QED and a scalar case in sQED, we will assume the latter in this exercise. The amplitude will depend in general on the n four-momenta of the external legs, denoted by $(\{p_i\})$. Next, we want to attach an outgoing photon to one of the external legs of the diagram, and consider all possible ways in which this attachment can be done. To begin with, one can attach the photon to an incoming electron/positron leg, as shown in Fig. 3.7. This leads to the following amplitude, denoted $\mathcal{M}_1(\{p_i\}, q)$, and

$$\mathcal{M}_0\left(\{p_i\}\right) \qquad\qquad\qquad \mathcal{M}_1\left(\{p_i\}, q\right)$$
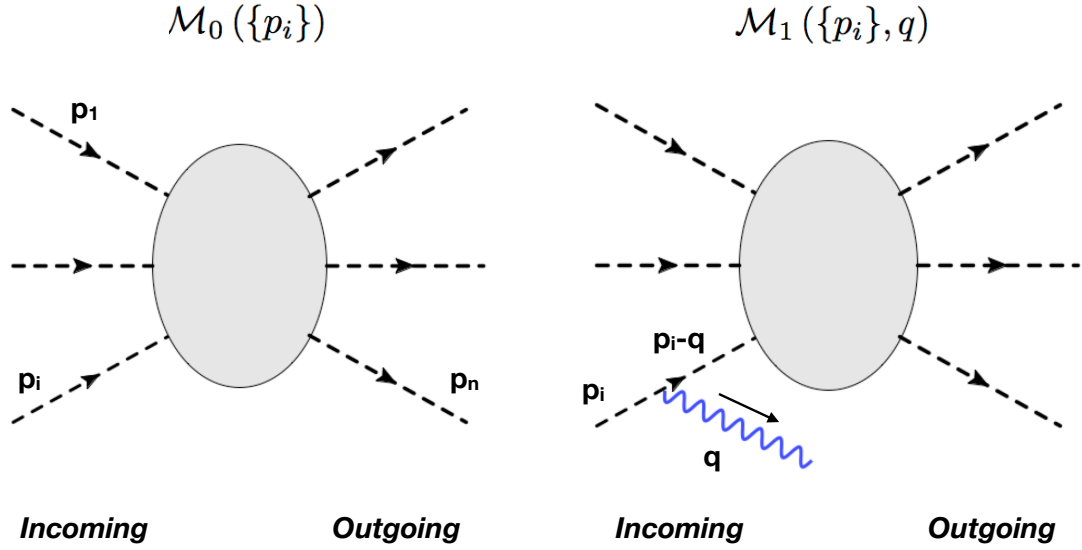
Figure 3.7. Left plot: a generic scattering amplitude in scalar QED, denoted by $\mathcal{M}\left(\{p_i\}\right)$, consisting of a number of incoming and outgoing (scalar) electrons or positrons. Right plot: the same diagram, now with an outgoing photon with momentum q attached to the incoming scalar leg i. In this case the full scattering amplitude is denoted by $\mathcal{M}\left(\{p_i\}, q\right)$, see text for more details.

which now depends in addition on q, the photon momentum,

$$\mathcal{M}_1\left(\{p_i\}, q\right) = e\frac{(2p_i^\mu - q^\mu)}{(p_i - q)^2 + m^2}\epsilon_\mu(q)\mathcal{M}_0(p_1, p_2, \ldots, p_i - q, \ldots, p_n)\,, \tag{3.125}$$

where $\mathcal{M}_0(p_1, p_2, \ldots, p_i - q, \ldots, p_n)$ is the original scattering amplitude $\mathcal{M}_0\left(\{p_i\}\right)$ but with the replacement $p_i \to p_i - q$ due to the effects of attaching the outgoing photon to the i-th incoming scalar leg. When computing Eq. (3.125), we have used the Feynman rules of scalar QED, in particular Eq. (3.94) for the scalar propagator and Eq. (3.91) for the scalar-photon vertex.

Next, by using the mass-shell relation $p_i^2 = -m^2$ for the (scalar) electrons (which we can do since p_i corresponds to an external line), the orthogonality between polarization and momentum vectors of the photon, $q \cdot \epsilon = 0$, and the fact that on-shell external photons are massless, $q^2 = 0$, we can write the amplitude as

$$\mathcal{M}_{1,\text{in},-} = -e\frac{p_i \cdot \epsilon}{p_i \cdot q}\mathcal{M}_0(p_1, p_2, \ldots, p_i - q, \ldots, p_n)\,, \tag{3.126}$$

where the subindex indicates that the photon has been attached to an incoming electron. On the other hand, we could also attach the outgoing photon to a leg j with opposite charge than leg i, and then one would get

$$\mathcal{M}_{1,\text{in},+} = +e\frac{p_j \cdot \epsilon}{p_j \cdot q}\mathcal{M}_0(p_1, p_2, \ldots, p_j - q, \ldots, p_n)\,, \tag{3.127}$$

which is the same as for electrons but with an overall change of sign. We can do that because in scalar QED the two degrees of freedom of the complex scalar field can be made to correspond to "scalar electrons" and "scalar positrons", see the discussion around Eq. (3.90).

Using similar considerations, if the photon is attached to an outgoing scalar leg labelled by the index k,

one would then find for the amplitude,

$$\mathcal{M}_{1,\text{out},\pm} = \mp e \frac{p_k \cdot \epsilon}{p_k \cdot q} \mathcal{M}_0(p_1, p_2, \ldots, p_k + q, \ldots, p_n), \tag{3.128}$$

where note that now the replacement is $p_k \to p_k + q$ since the four-momentum p_k is outgoing. The overall sign in front of $\pm e$ is determined by whether or not this particular line corresponds to an scalar electron or an scalar positron. Finally, we can attach the outgoing photon to some of the internal lines of the diagram, leading to an amplitude denoted by $\mathcal{M}_{1,\text{int}}(\{p_i\}, q)$. In this case, note that the external scalar lines are not modified due to the emission of the photon with momentum q.

Note that so far we have not used the fact the the theory is gauge invariant at all, just that there is a scalar-fermion interaction with a dimensionless coupling. In order to derive the connection between the *Lorentz invariance of the theory* and the fact that there is a conserved charge (that we will be able identify with the electric charge) we need to take the so-called *soft limit* for the outgoing photon:

> The soft limit of an scattering amplitude in a gauge theory is defined by the condition that all the components of the four-momentum associated to the emitted photon, q, are much smaller than any of the external momenta $\{p_i\}$. To be precise, we define the soft limit by the following condition:
>
> $$|q \cdot p_i| \ll |p_j \cdot p_k|. \tag{3.129}$$
>
> In this approximation, we say that the scattering amplitude emits a *soft photon*.

In the soft limit, we can approximate all terms where the photon momentum appears additively as follows

$$\mathcal{M}_0(p_1, p_2, \ldots, p_i \pm q, \ldots, p_n) \simeq \mathcal{M}_0(p_1, p_2, \ldots, p_i, \ldots, p_n), \tag{3.130}$$

that is, we can neglect q whenever it is added to any of the external momenta, and therefore the only relevant contributions in this limit are those where the photon is attached to the external scalar electron/positron legs, where the terms of the form $1/(p_i \cdot q)$ become large.

> We denote matrix elements that have terms such as $1/(p_i \cdot q)$, where a external photon becomes soft, as scattering amplitudes exhibiting a *soft singularity*. This means that that matrix element is formally divergent when the photon becomes infinitely soft, $q \to 0$.

Under these conditions, and taking into account that for $\mathcal{M}_{1,\text{int}}(\{p_i\}, q)$ there will be no terms of this form (so its contribution to the total matrix element is not enhanced in the soft limit), we can write the scattering amplitude with the extra outgoing photon in the soft limit, including all possible ways to attach this photon, as follows

$$\mathcal{M}_{1,\text{soft}} \simeq e \mathcal{M}_0\left(\{p_i\}\right) \left[\sum_{\text{incoming}} Q_i \frac{p_i \cdot \epsilon}{p_i \cdot q} - \sum_{\text{outgoing}} Q_i \frac{p_i \cdot \epsilon}{p_i \cdot q}, \right] \tag{3.131}$$

where $Q_i = \pm$ depending on the specific charge carried by the external leg.

Now the crucially observation is that under a Lorentz transformation any scattering amplitude $\mathcal{M}$ should be *Lorentz invariant*, since $\mathcal{M}$ is manifestly a Lorentz scalar. On the other hand, we know that the photon polarization states ϵ_μ themselves are not Lorentz invariant. So assume we have a Lorentz transformation such that

$$
\begin{aligned}
\epsilon_\mu &\rightarrow \epsilon_\mu + q_\mu \,, \\
q_\mu &\rightarrow q_\mu \,,
\end{aligned}
\tag{3.132}
$$

which defines the so-called *little group* of the Lorentz group. If we apply this Lorentz transformation on the matrix element Eq. (3.131), we find that

$$
\mathcal{M}_{1,\text{soft}} \rightarrow \mathcal{M}_{1,\text{soft}} + e\mathcal{M}_0 \left[\sum_{\text{incoming}} Q_i - \sum_{\text{outgoing}} Q_i \right]
\tag{3.133}
$$

where the original matrix element $\mathcal{M}_0 \left(\{p_i\} \right)$ is invariant under this Lorentz transformation since it does not contain any dependence on the photon momentum q.

Therefore, we find that the only way for the scattering amplitude Eq. (3.131) to be Lorentz invariant is to ensure that

$$
\sum_{\text{incoming}} Q_i = \sum_{\text{outgoing}} Q_i \,,
\tag{3.134}
$$

in other words, that electric charge is conserved in all interactions of the scalar QED theory. We emphasize that this argument does not reply on gauge symmetry, and highlights the close connection between the conservation of electric charge and the Lorentz invariance of (scalar) QED. A similar derivation can be carried out in the case of the spinor electrodynamics. Summarizing, the main take-away message that we can derive from this exercise is:

> In a theory that involves spin-1 fields (photons), one can show that the Lorentz invariance of the theory requires the conservation of the electric charge. This can be demonstrated in the soft limit, where a photon with small energy is emitted from one of the external legs of a general scattering amplitude.

3.8   Scalar QED at the one loop level

Up to now we have considered scalar QED for tree-level processes, in other words, we have restricted ourselves to the Born approximation. Now we can take a first look at the behaviour of scalar QED at the one-loop level, using some of the concepts that we developed in the first part of the course. Specifically, we will show how to compute some of the counterterms of scalar QED, using the same techniques that were used in the context of $\lambda\phi^4$ theory.

The starting point of the calculation is the Lagrangian of scalar QED at the tree level, namely

$$
\begin{aligned}
\mathcal{L} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - D^\mu\psi^\dagger D_\mu\psi - m^2\psi^\dagger\psi \\
&= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \partial^\mu\psi^\dagger\partial_\mu\psi - m^2\psi^*\psi + ieA^\mu\left[\left(\partial_\mu\psi^\dagger\right)\psi - \psi^\dagger\left(\partial_\mu\psi\right)\right] - e^2A_\mu A^\mu\psi^\dagger\psi, \quad (3.135)
\end{aligned}
$$

where we have expanded the covariant derivative in order to better highlight the different types of interaction operators that appear in the theory - recall that we have both three-point and four-point interactions. Note that this Lagrangian already includes all possible terms with positive or zero mass dimension coefficients and that satisfy Lorentz, U(1) gauge symmetry, parity, time reversal, and charge conjugation.

What we need to do now is to include the counterterms to account for the fact that beyond the tree level the bare parameters of the Lagrangian are not physical observables, and we can redefine them in order to renormalize the theory, absorbing this way the ultraviolet infinities that arise in the one-loop scattering amplitudes. This is the same program as the one that was discussed at some length in the case of $\lambda\phi^4$ theory. Once we account for these counterterms, the full Lagrangian of scalar QED now reads

$$
\begin{aligned}
\mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_1, &(3.136)\\
\mathcal{L}_0 &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \partial^\mu\psi^*\partial_\mu\psi - m^2\psi^*\psi, \\
\mathcal{L}_1 &= iZ_1e\left[\psi^\dagger\partial^\mu\psi - \left(\partial^\mu\psi^\dagger\right)\psi\right]A_\mu - Z_4e^2\psi^\dagger\psi A^\mu A_\mu + \mathcal{L}_{\mathrm{ct}}, \\
\mathcal{L}_{\mathrm{ct}} &= -(Z_2-1)\partial^\mu\phi^\dagger\partial_\mu\phi - (Z_m-1)m^2\phi^\dagger\phi - \frac{1}{4}(Z_3-1)F^{\mu\nu}F_{\mu\nu}.
\end{aligned}
$$

The structure of the Lagrangian in Eq. (3.136) is the following:

- First of all, $\mathcal{L}_0$ contains the kinetic terms, namely those corresponding to the free theory.

- The interaction terms are then contained in $\mathcal{L}_1$, which is divided into the same terms as in the original Lagrangian but with the renormalization factors Z_i included and the $\mathcal{L}_{\mathrm{ct}}$ part, which contain the actual counterterms. Note that at tree level one has that $Z_i = 1$, and thus one has that $\mathcal{L}_{\mathrm{ct}} = 0$.

- This Lagrangian is now expressed in terms of the physical coupling constant e and the physical scalar mass m

Although it is beyond the scope of these lectures to give a formal proof, scalar QED is a *renormalizable theory*, which means that once I compute the values of the counterterms $Z_1, Z_4, Z_\lambda, Z_2, Z_m, Z_3$ at any given order in perturbation theory, then any other general physical observable will also be finite at the same perturbative order.

Note that in Eq. (3.136) we have explicitly separated the free theory kinetic terms (contained in $\mathcal{L}_0$) from the interaction terms in $\mathcal{L}_1$. The latter include the pure counterterm operators, $\mathcal{L}_{\mathrm{ct}}$, which at this point are treated as interactions. The advantage of this way of writing the Lagrangian of the theory is that in Feynman diagrams we can always use the free theory propagators of the photon and scalar fields, and the counterterms in $\mathcal{L}_{\mathrm{ct}}$ are treated as *new interaction terms* with their associated Feynman rules.

In this way the renormalization process is somewhat more transparent, and the only thing we need to do is to determine the explicit values of the Z_i at any given order in perturbation theory. In addition, this way of expressing the theory beyond the Born approximation is handier from the computational point of view: we can use the propagators of the free theory supplemented by two types of renormalized interactions:

- The same ones as in the original theory, but now expressed in terms of the physical coupling, such as $-Z_4 e^2 \psi^\dagger \psi A^\mu A_\mu$.

- Genuinely new interaction terms, that arise from the counterterms associated to the kinetic terms, for example $-\frac{1}{4}(Z_3 - 1)F^{\mu\nu}F_{\mu\nu}$ results a new two-point interaction between an incoming and an outgoing photon field.

In order to show how scalar QED works beyond the Born approximation, let us illustrate the computation of the photon self-energy $\Pi^{\mu\nu}(k)$ at $\mathcal{O}(e^2)$ in the perturbative expansion. Recall that in the case of $\lambda\phi^4$ theory, the calculation of the one loop corrections to the scalar self-energy lead to the definition of the *physical mass*. A photon is massless and so this is not possible here, and what this calculation will lead us is to the explicit form of the renormalized (interacting) propagator. As in the case of $\lambda\phi^4$, this self-energy arises from the loop corrections to the photon propagator.

In the same way as in $\lambda\phi^4$ theory, the all-order expression of the photon propagator in momentum space can be written as follows:

$$\boldsymbol{\Delta}_{\mu\nu}(k) = \Delta_{\mu\nu}(k) + \Delta_{\mu\rho}(k)\Pi^{\rho\sigma}(k)\Delta_{\sigma\nu}(k) + \ldots \tag{3.137}$$

with $i\Pi^{\mu\nu}(k)$ given by the sum of all 1PI diagrams with two external photon lines (and with the photon propagators removed) and where the free-photon propagator has the standard expression in the R_ξ gauge,

$$\Delta_{\mu\nu}(k) = \frac{1}{k^2 - i\epsilon}\left(g_{\mu\nu} - (1-\xi)\frac{k_\mu k_\nu}{k^2}\right). \tag{3.138}$$

This series is a geometric one that can be summed to all orders once the expression for the self-energy $i\Pi^{\mu\nu}(k)$ is known at any given order in perturbation theory, in the same way as in the case of $\lambda\phi^4$ theory.

Recall also that in our discussion of the Ward identities in Sect. 3.6 we reasoned that the photon propagator should be *transverse*, namely the fact that in any amplitude containing a photon propagator should satisfy

$$\mathcal{M} = \mathcal{M}_{\mu\nu}\Delta^{\mu\nu}(k) \Rightarrow \mathcal{M}_{\mu\nu}k^\mu k^\nu = 0, \tag{3.139}$$

to ensure that the amplitude is independent of the unphysical gauge fixing parameter ξ. From Eq. (3.138) this has the consequence that the photon self-energy should be transverse as well, that is, it should satisfy the condition

$$\Pi^{\mu\nu}(k)k_\mu k_\nu = 0. \tag{3.140}$$

As we will show now, this condition is indeed satisfied by the explicit calculation of the self-energy at $\mathcal{O}(e^2)$.

> The Ward identities of QED tell us that the photon propagator should be transverse to all orders in the perturbative expansion, implying $\Pi^{\mu\nu}(k)k_\mu k_\nu = 0$. When performing loop calculations, it is important to verify this property order by order.

The Feynman diagrams that contribute to the photon self-energy at $\mathcal{O}(e^2)$ are shown in Fig. 3.8. Note that they contain both the one-loop diagrams and the contributions from the counterterms, which as we know are required in order to cancel the ultraviolet divergences of the theory. Specifically, the last diagram in Fig. 3.8 corresponds to the counterterm associated to the photon kinetic term in $\mathcal{L}_{\rm ct}$ in Eq. (3.136).

The contribution for these three diagrams can be computed using the Feynman rules of scalar QED, suitably modified in order to account for the counterterm, required once we start doing loop calculations. The result for the sum of the three diagrams is given by

$$
\begin{aligned}
i\Pi^{\mu\nu}(k) &= -(iZ_e e)^2 \int \frac{d^4 l}{(2\pi)^4} \frac{(2l+k)^\mu (2l+k)^\nu}{((l+k)^2+m^2)(l^2+m^2)} \\
&+ (-2iZ_4)e^2 \eta^{\mu\nu} \int \frac{d^4 l}{(2\pi)^4} \frac{1}{l^2+m^2} - i(Z_3-1)\left(k^2 \eta^{\mu\nu} - k^\mu k^\nu\right) ,
\end{aligned}
\tag{3.141}
$$

since it can be shown that the last diagram in Fig. 3.8 has associated a Feynman rule proportional to $\left(k^2 \eta^{\mu\nu} - k^\mu k^\nu\right)$. Note also the contribution from the counterterm does not include any loop integral. Note how the couplings now include the corresponding counterterms Z_e and Z_4: this is the only modification that the original Feynman rules of scalar QED experience in the case of the renormalized theory.

An important simplification of this expression is provided by the fact that, given that $Z_i = 1 + \mathcal{O}(e^2)$ (since all counterterms are one at the Born level), we can set $Z_1 = Z_4 = 1$ in the expression above since the subsequent terms would be formally higher order. We then see that Z_3 is the only counterterm relevant for the calculation of the photon self-energy at this perturbative order. In other words, for the renormalisation of the photon self-energy at the one loop level the effects of the renormalisation of the coupling constants are formally subleading.

For the next step, we can combine the two first terms in Eq. (3.141), those that involve an integral over the loop momentum l, into a single term, to have

$$
i\Pi^{\mu\nu}(k) = e^2 \int \frac{d^4 l}{(2\pi)^4} \frac{N^{\mu\nu}}{((l+k)^2+m^2)(l^2+m^2)} - i(Z_3-1)\left(k^2 \eta^{\mu\nu} - k^\mu k^\nu\right) ,
\tag{3.142}
$$

where by doing some simple algebra one can check that the numerator is given by

$$
N^{\mu\nu} = (2l+k)^\mu (2l+k)^\nu - 2\left[(l+k)^2 + m^2\right] \eta^{\mu\nu} .
\tag{3.143}
$$

The integral over l is clearly divergent in the ultraviolet, since at large momenta we have $\sim \int d^4 l \ (l^2/l^4) = \int l^3 dl \ (l^{-2}) \sim \Lambda_{\rm uv}^2$. Therefore in order to preempt the fact that we need to work in $d = 4 - \epsilon$ space-time dimensions in order to regulate the ultraviolet divergent that would arise from this integral we replace $e \to e\mu^{\epsilon/2}$ (so that the electron charge is still dimensionless in d dimensions) and then combine the two denominators using Feynman's formula. The result is given by

$$
i\Pi^{\mu\nu}(k) = e^2 \mu^\epsilon \int_0^1 dx \frac{d^d q}{(2\pi)^d} \frac{M^{\mu\nu}}{(q^2+D)^2} - i(Z_3-1)\left(k^2 \eta^{\mu\nu} - k^\mu k^\nu\right) ,
\tag{3.144}
$$

where we have defined $q = l + xk$, $D = x(1-x)k^2 + m^2$, and the numerator is given by

$$
M^{\mu\nu} = (2q + (1-2x)k)^\mu (2q + (1-2x)k)^\nu - 2\left[(q + (1-x)k)^2 + m^2\right] \eta^{\mu\nu} .
\tag{3.145}
$$
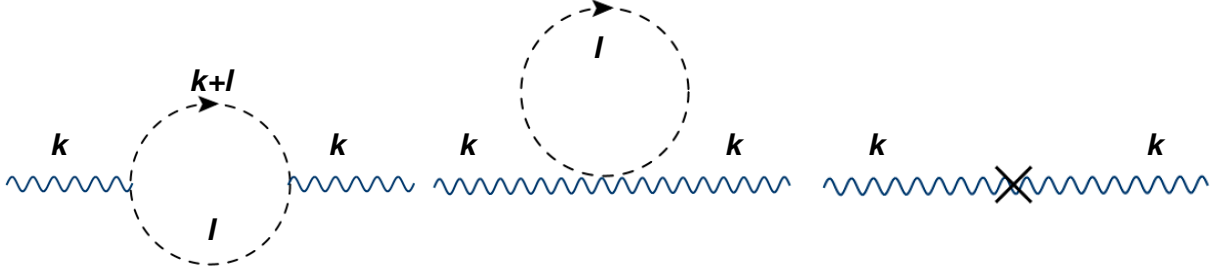
Figure 3.8. The three Feynman diagrams for the $\mathcal{O}\left(e^2\right)$ contributions to the photon self-energy in scalar QED. The last term corresponds to the counterterm associated to the photon propagator being treated as an interaction term.

At this point we have everything that we need in order to evaluate the photon self-energy $\Pi^{\mu\nu}$ in scalar electrodynamics at $\mathcal{O}\left(e^2\right)$. Doing some algebra with the numerator, including canceling terms linear in q and reshuffling the integration variables one finds that one can make the following replacement:

$$N^{\mu\nu} \to -4y^2(k^2 g^{\mu\nu} - k^\mu k^\nu)\,, \tag{3.146}$$

with $x = y + 1/2$. Therefore we find that the self-energy $\Pi^{\mu\nu}(k)$ is transverse at $\mathcal{O}(e^2)$, implying that

$$\Pi^{\mu\nu}(k)k_\mu k_\nu = 0\,, \tag{3.147}$$

as required from the gauge invariant of the theory, as discussed above in the context of the Ward identities. Is it possible to verify by explicit computation that the same property holds for higher perturbative orders.

Next, we can perform the integral over the rescaled loop momentum q, and focusing only in the divergent part, we can expand using similar methods as in the case of $\lambda\phi^4$ the result to get

$$\mu^\epsilon \int \frac{d^d q}{(2\pi)^d} \frac{1}{(q^2 + D)^2} = \frac{i}{8\pi^2} \frac{1}{\epsilon} + \mathcal{O}\left(\epsilon^0\right)\,. \tag{3.148}$$

Following this, we need to compute the integral over y to get

$$\int_{-1/2}^{1/2} dy N^{\mu\nu} = -\frac{1}{3}\left(k^2 \eta^{\mu\nu} - k^\mu k^\nu\right)\,, \tag{3.149}$$

and combining all the results obtained so far, we achieve the following results for the photon-self energy at the first non-trivial order in the perturbative expansion:

$$\Pi^{\mu\nu}(k) = \Pi(k^2)\left(k^2 \eta^{\mu\nu} - k^\mu k^\nu\right)\,, \tag{3.150}$$

where we have defined

$$\Pi(k^2) = -\frac{e^2}{24\pi^2} \frac{1}{\epsilon} + \text{finite} - (Z_3 - 1)\,. \tag{3.151}$$

Therefore, requesting the finiteness of the photon self-energy at one loop in perturbation theory, this calcu-

lation provides us with the answer for the first non-trivial correction to the Z_3 counterterm, namely

$$Z_3 = 1 - \frac{e^2}{24\pi^2}\frac{1}{\epsilon} + \dots , \qquad (3.152)$$

where the dots indicate possible constant terms. Recall that different renormalization choices will lead to different constant terms in the definition of the counterterms of the theory. We have thus demonstrated that the quadratic divergence that affects the one-loop diagrams in Fig. 3.8 for the photon self-energy is canceled out thanks to the counterterm Z_3 that redefines the physical photon propagator.

Using a similar approach one can evaluate all the other counterterms of the theory. Specifically, one needs to evaluate the one-loop corrections to the *scalar field propagator* as well as to the *three-point* and *four-point interaction vertices*. This way, the complete set of counterterms of scalar QED will be determined, allowing the computation of general finite physical observables at a given order in the perturbative expansion.

3.9   QED: spinor electrodynamics

So far we have considered the interactions of a spin-one gauge field, the photon, with a complex scalar field. We can now use the concepts that we have learned to introduce spinor QED, namely the Quantum Field Theory describing the interactions between (Dirac) fermions and photons. Our starting point of this discussion of QED is the realisation that

> In spinor electrodynamics, the electromagnetic current $j^\mu(x)$ is proportional to the *Noether current* corresponding to the U(1) global symmetry of the Dirac field upon a phase rotation. Therefore electric charge conservation is a direct consequence of this U(1) global symmetry of the theory.

This conserved current associated to the global U(1) symmetry can be computed using Noether's theorem and found to be:

$$j^\mu(x) = e\bar{\Psi}(x)\gamma^\mu\Psi(x) , \qquad (3.153)$$

where e is the electron charge in the Heaviside-Lorentz system of units. Note that in the following we use $\Psi(x)$ to denote the Dirac spinor fields, to avoid the confusion from the complex scalar field $\psi(x)$. Taking into account Eq. (3.153), and that the matter-photon coupling must be of the form $j_\mu(x)A^\mu(x)$, the full Lagrangian of spinor QED will be given by

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + i\bar{\Psi}\slashed{\partial}\Psi - m\bar{\Psi}\Psi + e\bar{\Psi}\gamma^\mu\Psi A_\mu . \qquad (3.154)$$

Here we will keep the discussion of QED restricted to tree-level processes, so we do not need to add counterterms of this Lagrangian. In other words, we have the kinetic terms of the photon and spinor fields plus one interaction term that couples them. Note that unlike in the case of scalar QED, a four-point interaction vertex of the form $A_\mu A^\mu\bar{\Psi}\Psi$ is absent (it corresponds to an irrelevant (non-renormalizable) interaction term).

This Lagrangian is invariant under a local U(1) gauge symmetry, as we saw in the case of scalar QED, namely:

$$A^\mu(x) \to A^\mu(x) - \partial^\mu\lambda(x) , \qquad (3.155)$$

$$\Psi(x) \rightarrow \exp\left(-ie\lambda(x)\right)\Psi(x)\,. \tag{3.156}$$

Note that the Lagrangian is invariant under this gauge transformation irrespective of whether of not the fields obey the equations of motion. This invariance of the theory under gauge transformations becomes more transparent once the Lagrangian is expressed in terms of the covariant derivative

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + i\bar{\Psi}\slashed{D}\Psi - m\bar{\Psi}\Psi\,, \tag{3.157}$$

which has been defined as

$$D_\mu \equiv \partial_\mu - ieA_\mu\,, \tag{3.158}$$

and which transforms in the same way (*covariantly*) as the Dirac field under a gauge transformation. Indeed, the transformation properties under U(1) of the covariant derivative are the same irrespective of the matter fields to which the fermion is coupling. A useful expression relates the covariant derivative with the field-strength tensor, namely

$$F^{\mu\nu} = \frac{i}{e}\left[D^\mu, D^\nu\right]\,, \tag{3.159}$$

where recall that $F^{\mu\nu}$ is invariant under a gauge transformation. This relation is particularly useful in the case of gauge theories based on non-Abelian gauge groups, since there the field-strength tensor $F^{\mu\nu}$ is not invariant anymore. To demonstrate this relation, note that from the definition of the commutator

$$[D^\mu, D^\nu] = \left(\partial^\mu - ieA^\mu\right)\left(\partial^\nu - ieA^\nu\right) - \left(\partial^\nu - ieA^\nu\right)\left(\partial^\mu - ieA^\mu\right) = -ie\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right)\,, \tag{3.160}$$

with all other terms canceling out.

Using similar methods as in the case of scalar QED, it is now possible to derive the Feynman rules of spinor QED:

- For each interaction vertex: $ie\gamma^\mu$,

- For each internal photon: $-ig^{\mu\nu}/(k^2 - i\epsilon)$,

- For each internal fermion: $-i(-\slashed{p} + m)/(p^2 + m^2 - i\epsilon)$,

where we are working for simplicity in the Feynman gauge $\xi = 1$. The rules for the external fermions and photons are the same ones as in the corresponding free theories. Recall that when working in spinor QED one has to be rather careful with the spinor order, since spinors do not commute among them: in particular, spinor indices should be contracted starting at one end of a fermion line.

Let us now show how we can use these Feynman rules in order to compute an scattering amplitude (and the corresponding spin-averaged square) in spinor QED. Specifically, we will consider the process of electron-positron annihilation into a photon pair,

$$e^+ + e^- \rightarrow \gamma + \gamma\,, \tag{3.161}$$

whose Feynman diagram are represented in Fig. 3.9. In the same way as with the fact that in general we are not interested in the photon polarisation and we will sum (average) over the final (initial) state polarisation vectors, since for fermions their helicity is not typically observable we will also sum (average) over the final (initial) state helicities of the outgoing (incoming) fermions.
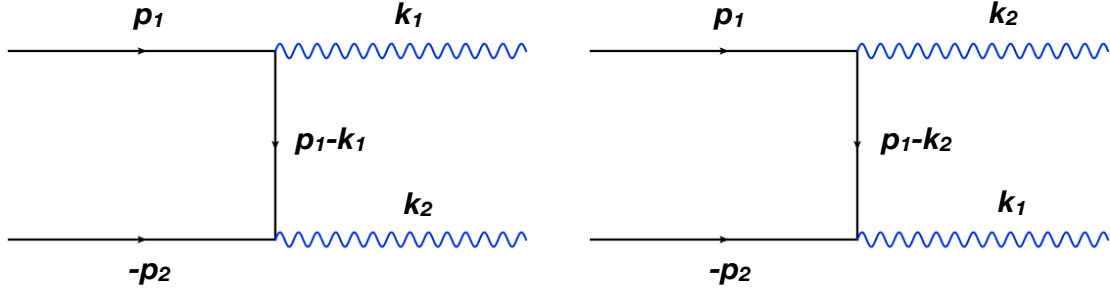
Figure 3.9. The Feynman diagram for the process of electron-positron annihilation into a photon pair in spinor QED, corresponding to t-channel and u-channel scattering respectively.

Using the Feynman rules of QED, we now have all the ingredients to evaluate these two diagrams, and one finds that the scattering amplitude of this process $\mathcal{M}$ will be given by

$$\mathcal{M} = e^2 \epsilon_{1'}^\mu \epsilon_{2'}^\nu \bar{v}_2 \left[\gamma_\nu \left(\frac{-\slashed{p}_1 + \slashed{k}_1 + m}{-t + m^2} \right) \gamma_\mu + \gamma_\mu \left(\frac{-\slashed{p}_1 + \slashed{k}_2 + m}{-u + m^2} \right) \gamma_\nu \right] u_1 \qquad (3.162)$$

where the primes indicate that a particle corresponds to the final state, and where we have defined the usual Mandelstam variables:

$$s = -(p_1 + p_2)^2, \quad t = -(p_1 - k_1)^2, \quad u = -(p_1 - k_2)^2. \qquad (3.163)$$

Recall that the three Mandelstam variables are related among them due to four-momentum conservation as $s + t + u = 2m^2$. As is clear from the structure of the amplitude $\mathcal{M}$, these two diagrams will have associated t-channel and u-channel *singularities* respectively. This means that the scattering amplitude will have a maximum when either $-t \simeq m^2$ or when $-u \simeq m^2$.

This scattering amplitude, Eq. (3.162), is not particularly illuminating since it depends on the specific spins of the incoming fermions and on the specific polarization of the outgoing photons. Let us now evaluate the square of the scattering amplitude, as needed in order to compute the cross-section for this process. To do this, we write our scattering amplitude as

$$\mathcal{M} \equiv \epsilon_{1'}^\mu \epsilon_{2'}^\nu \bar{v}_2 A_{\mu\nu} u_1, \qquad (3.164)$$

where $A_{\mu\nu}$ is a matrix in spinor space. Squaring this amplitude leads us, using the properties of the spinor algebra, to

$$|\mathcal{M}|^2 = \epsilon_{1'}^\mu \epsilon_{2'}^\nu \epsilon_{1'}^{\rho*} \epsilon_{2'}^{\sigma*} \left(\bar{v}_2 A_{\mu\nu} u_1 \right) \left(\bar{u}_1 A_{\sigma\rho} v_2 \right). \qquad (3.165)$$

In order to simplify this expression, we will use that fact that we are interested in the squared amplitude averaged over initial state spins of the electron and the positron. The calculation of this average, using the spinor algebra techniques, leads to

$$|\mathcal{M}|^2 = \frac{1}{4} \epsilon_{1'}^\mu \epsilon_{2'}^\nu \epsilon_{1'}^{\rho*} \epsilon_{2'}^{\sigma*} \mathrm{Tr} \left[A_{\mu\nu} \left(-\slashed{p}_1 + m \right) A_{\sigma\rho} \left(-\slashed{p}_2 + m \right) \right]. \qquad (3.166)$$

In addition to averaging over the spins of the electron and the positron, in general one is not able to measure the polarization of the final-state photons in a hard-scattering reaction, and therefore we wish to

sum over the final state photon polarizations.[4] From Eq. (3.166) is clear that we need to evaluate sums over the photon polarization vectors of the form

$$\sum_{\lambda=\pm} \epsilon_{\lambda}^{\mu}\left(\vec{k}\right) \epsilon_{\lambda}^{\rho*}\left(\vec{k}\right) \,, \qquad (3.167)$$

which can be computed *e.g.* in the Coulomb gauge, and where using the fact that photon is on-shell, $k^2 = 0$ and that the amplitudes involving external photons must satisfy the Ward identities, one can derive the following rather compact substitution rule

$$\sum_{\lambda=\pm} \epsilon_{\lambda}^{\mu}\left(\vec{k}\right) \epsilon_{\lambda}^{\nu*}\left(\vec{k}\right) \to g^{\mu\nu} \,. \qquad (3.168)$$

Note that if we multiply the above equation by $k_\mu k_\nu$ we find

$$\sum_{\lambda=\pm} \left(k_\mu \epsilon_{\lambda}^{\mu}\left(\vec{k}\right)\right)\left(k_\nu \epsilon_{\lambda}^{\nu*}\left(\vec{k}\right)\right) = k^2 = 0 \,, \qquad (3.169)$$

which is consistent with the fact that the polarization vector of a photon is orthogonal the corresponding four-momentum. So this is a useful cross-check of the internal consistency of this property of the sum over photon helicities.

Using the substitution rule, Eq. (3.168), we can perform the sum over final state polarizations to obtain

$$\frac{1}{4}\sum_{\lambda_1',\lambda2'}\sum_{s_1,s_2}|\mathcal{M}|^2 = e^4\left[\frac{\langle\Phi_{tt}\rangle}{(m^2-t)^2} + \frac{\langle\Phi_{uu}\rangle}{(m^2-u)^2} + \frac{\langle\Phi_{tu}+\Phi_{ut}\rangle}{(m^2-t)(m^2-u)}\cdot\right] \qquad (3.170)$$

Therefore, one finds that the squared amplitude is composed by three terms: the square of the t-channel amplitude, the square of the u channel amplitude, and the interference term. It is always possible to identify for instance the term that arises from the t channel scattering, since it is the one that diverges when $t \to -m^2$. Note that the t- and u-channel singularities are stronger in the squared contributions as compared to the interference term.

The various numerators in Eq. (3.170) are given by the traces of matrices in spinor space. For instance, the numerator of the square of the t-channel amplitude is given by

$$\langle\Phi_{tt}\rangle = \frac{1}{4}\mathrm{Tr}\left[\gamma_\nu\left(-\not{p}_1 + \not{k}_1' + m\right)\gamma_\mu\left(-\not{p}_1 + m\right)\gamma^\mu\left(-\not{p}_1 + \not{k}_1' + m\right)\gamma^\nu\left(-\not{p}_2 + m\right)\gamma^\mu\right] \,, \qquad (3.171)$$

which can be computed using properties of the algebra of the Dirac matrices, such as

$$\gamma^\mu \not{A}\gamma_\mu = 2\not{A} \,, \qquad (3.172)$$

and related expressions, as well as the formulae for the traces of matrices in spinor space. After a lengthy but straightforward calculation one gets

$$\langle\Phi_{tt}\rangle = 2\left[tu - m^2(3t + u) - m^4\right] \,, \qquad (3.173)$$

---

[4]Note that one *averages* over initial state polarizations and spins, but one instead *sums* for final state polarization and spin.

and likewise for the other contributions in Eq. (3.170). This example illustrates how from the conceptual point of view, scalar QED and spinor QED are closely related theories, but the use for the former in these lectures is motivated since carrying out the actual calculations is simpler with scalars than having to use the spinor algebra.

The same strategies that we have applied here for the $e^+ + e^- \rightarrow \gamma\gamma$ scattering amplitude also apply to all other tree level diagrams in QED, which thus have associated no further conceptual difficulties. Likewise, the computation of loop corrections in spinor electrodynamics follows the same strategy as that outlined in Sect. 3.8 in the case of scalar electrodynamics.

The helicity formalism. It is easy convince yourselves that the structure of the scattering amplitudes in QED simplify significantly in the massless limit, where spin coincides with helicity, and we can use the so-called *helicity formalism*. For example, for this specific process we have that

$$\frac{\langle \Phi_{tt} \rangle}{(m^2 - t)^2} = \frac{2 \left[tu - m^2(3t + u) - m^4 \right]}{(m^2 - t)^2} \rightarrow \frac{2tu}{t^2} = \frac{2u}{t} , \tag{3.174}$$

in the massless limit, and likewise for the other contributions to this process. Using the spinor helicity formalism, the calculations of scattering amplitudes in gauge theories become rather simpler than in the massive case. Another advantage of the helicity formalism is that a fermion helicity becomes a conserved quantity: for example, in the massless limit a right-handed electron will always be a right-handed electron and will never change into a left-handed one.