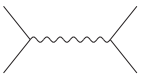
12 Introduction to Renormalization

12.1 Ultraviolet (UV) Singularities

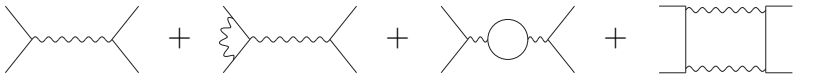
So far, everything was computed at tree-level, that is, at the lowest nontrivial order in perturbation theory. Very often, a more precise determination of a cross section is desirable and we are thus led to consider loop diagrams. In order to illustrate this, consider the example $e^+e^- \to \mu^+\mu^-$. The perturbative expansion of the corresponding amplitude is written as

$$\mathcal{M} = \alpha\mathcal{M}_0 + \alpha^2\mathcal{M}_1 + \alpha^3\mathcal{M}_2 + \mathcal{O}(\alpha^4), \tag{12.1}$$

where $\alpha \sim 1/137$. When we computed the corresponding amplitude in section 11.3 we only computed the leading order term

$\alpha\mathcal{M}_0 =$  $\sim e^2 \sim \alpha$ (12.2)

Using this expression for the amplitude, we will get the leading-order cross section $\sigma_0 \sim \alpha^2|\mathcal{M}_0|^2$. If we want to compute corrections of the order $\sim \alpha^3$ to this result, we will have to compute the amplitude to an accuracy of the order $\sim \alpha^2$.

$\mathcal{M} =$  $+ \ \ldots$ (12.3)

The one-loop correction to the cross section is related to the interference term of $\mathcal{M}_0$ and $\mathcal{M}_1$

$$\sigma_1 \sim |\alpha\mathcal{M}_0 + \alpha^2\mathcal{M}_1 + \mathcal{O}(\alpha^3)|^2 = \alpha^2|\mathcal{M}_0|^2 + 2\alpha^3\mathrm{Re}(\mathcal{M}_0\mathcal{M}_1^*) + \mathcal{O}(\alpha^4) \tag{12.4}$$

The whole procedure looks pretty straightforward. However, if we try to compute a loop diagram, we run into trouble. Consider as an example the vertex correction $\mathcal{V}$, depicted in figure 7.

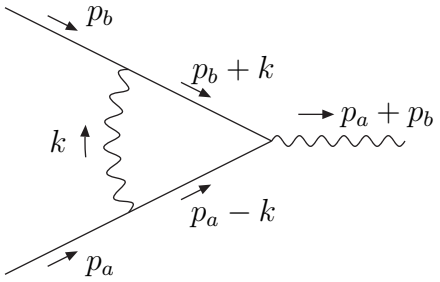


Figure 7: Vertex correction for $e^+e^- \to \mu^+\mu^-$ scattering.

Using the Feynman rules listed in section 10.4 we end up with an expression of the form

$$\mathcal{V} \sim \int \frac{d^4k}{(2\pi)^4} \frac{k\ k}{k^2((p_b+k)^2 - m^2)((p_a-k)^2 - m^2)} \tag{12.5}$$

where we did not bother to write down the full algebraic expression resulting from the spinor and Lorentz algebra but only the terms involving k. The two factors of k in the numerator stem

from the two fermion propagators. The important point is that this integral diverges. Indeed, considering the limit $k \to \infty$ we can neglect p_a, p_b and m and find

$$\mathcal{V} \sim \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^4} \sim \int \frac{dk}{(2\pi)} \frac{1}{k} = \infty \tag{12.6}$$

where we used $d^4k \sim k^3 dk$. These singularities are called ultraviolet (UV) singularities because they come from the region $k \to \infty$.

Similar problems are encountered if we try to compute some of the other one-loop diagrams and our final answer for the cross section at next-to-leading order seems to be infinity.

12.2 Infrared (IR) Singularities

There is another class of singularities that shows up in QED and QCD. As we saw in section 12.1 the UV singularities are related to the region of large k. However, there is also a potential danger of getting singularities from the region $k \sim 0$ or more generally, from zeros in the denominators of the integrand. These singularities are called infrared (IR) singularities. These occur if some (massless) particle becomes very soft or two become very collinear. These singularities have nothing to do with the UV singularities. The solution to the problem is completely different in the two cases. In fact, you already should have encountered an IR singularity. When you tried to compute the total cross section for Compton scattering in section 11.4 you should have found that the total cross section diverges. This is due to an IR singularity. Indeed, the final state photon can become arbitrarily soft, in which case the electron-photon pair becomes indistinguishable from a single electron. One possibility to get a well defined finite answer, for example, is to require that the final state photon has some minimal energy.

I will not discuss the IR singularities any further and will simply ignore them, safe in the knowledge that they can be dealt with in a manner totally different to that for the UV singularities. Thus in what follows I will call a cross section finite if it has no UV singularities, but it might well have IR singularities. Strictly speaking, we should replace every 'finite' below by 'UV-finite'.

12.3 Renormalization

It is important to realise that renormalization is not really about the removal of divergences, but simply an expression of the fact that in quantum field theories the value of certain parameters, e.g. the coupling constants, change with the energy scale used in a process. The infinities we encounter are then just a consequence of our ignorance of what is happening as $E \to \infty$ although we integrate up to this limit in any loop diagrams. We will demonstrate this below, and show how results do turn out to be finite after all.

To obtain a prediction for any measurable quantity $\mathcal{S}$, say a cross section, we start with a Lagrangian $\mathcal{L}$ from which we deduced the Feynman rules, which in turn were used to compute $\mathcal{S}$. The Lagrangian of QED eq. (10.33) has some parameters. So far, we denoted them by e, m and referred to them as mass and charge of the electron. Therefore, our result $\mathcal{S}$ will depend

on these parameters. However, the crucial point to observe is that the parameter m in the Lagrangian is *not* the real mass of the electron, nor is e its charge. The identification of the parameter in the Lagrangian and the measurable quantity is only justified at tree level, because beyond this level the parameters themselves receive corrections, i.e. the propagator and vertex diagram which define the mass and coupling strength are themselves corrected. Therefore, from now on we will be more precise and denote the parameters in $\mathcal{L}$ by m_0 and e_0 and call them the *bare* mass and *bare* charge respectively. Note that the bare parameters are not measurable. The (measurable) physical mass and charge of the electron will be denoted (as always) by m and e. $\mathcal{L}$ also depends on the fields, which we denoted so far by ψ and A. From now on, we denote them by ψ_0 and A_0 and call them the bare fields.

We are now ready to reformulate the problem we encountered in section 12.1. If we try to compute a measurable quantity in terms of the unmeasurable bare quantities as a perturbative expansion in the coupling constant we generally encounter divergences. That is, if we compute

$$\mathcal{S}(e_0, m_0, \psi_0, A_0) = \mathcal{S}_0(e_0, m_0, \psi_0, A_0) + e_0^2\ \mathcal{S}_1(e_0, m_0, \psi_0, A_0) + \mathcal{O}(e_0^4) \tag{12.7}$$

then we may find that $\mathcal{S}_1(e_0, m_0, \psi_0, A_0) = \infty$. In particular, this is true for two special physical quantities, namely the mass and the charge of the electron,

$$\begin{aligned} m &= m_0 + e_0^2\ m_1(e_0, m_0, \psi_0, A_0) + \mathcal{O}(e_0^4) \\ e &= e_0 + e_0^3\ e_1(e_0, m_0, \psi_0, A_0) + \mathcal{O}(e_0^5). \end{aligned} \tag{12.8}$$

But this is an expression for two measurable quantities in terms of unknown parameters. If the unknowns m_0 and e_0 are finite then we would get divergences in m_1 and e_1 and hence in m and e. Since m and e are finite quantities we conclude that the bare quantities are infinite. This is the root of the problem. UV divergences in our perturbative calculations show up if we try to express our results in terms of the unmeasurable, unphysical bare parameters, i.e. the parameters of the original Lagrangian.

In order to save the situation, we have to find new parameters such that the result of any physical quantity expressed in these new parameters — at any order in perturbation theory — is finite. Is this possible? Generally, the answer is no. However, for some special theories (and luckily QED is one of them) it is possible. Such theories are called renormalizable theories. The new parameters are called the *renormalized* quantities and are denoted by e_R, m_R and ψ_R, A_R. They are related to the bare quantities as follows:

$$\begin{aligned} \psi_0 &= Z_2^{1/2}\psi_R \\ A_0 &= Z_3^{1/2}A_R \\ m_0 &= Z_m m_R \\ e_0 &= Z_1 Z_2^{-1} Z_3^{-1/2} e_R \end{aligned} \tag{12.9}$$

This is simply a definition of the renormalization factors Z_1, Z_2, Z_3 and Z_m. Since the renormalization factors relate finite and divergent quantities, they have to be divergent themselves. More precisely, they can be written as a perturbative series with divergent coefficients.

To summarize, if we express the perturbative series for our physical quantity in terms of the renormalized quantities

$$\mathcal{S}(e_R, m_R, \psi_R, A_R) = \mathcal{S}_0(e_R, m_R, \psi_R, A_R) + e_R^2\ \mathcal{S}_1(e_R, m_R, \psi_R, A_R) + \mathcal{O}(e_R^4) \tag{12.10}$$

there will be no UV-divergences at any order in perturbation theory. Some people refer to this as 'hiding the infinities'. What is meant by this statement is that if we have a small number of input values ($m_R, e_R \ldots$) and express all results in terms of these input values we get finite answers for all measurable quantities. Thus, renormalizing QED enables us to relate any measurable quantity to a small number of measurable input values.

It is a highly non-trivial exercise to show that QED is indeed a renormalizable theory. Ultimately, it is related to the fact that all divergences in diagrams originate from divergences in the 2 and 3 point functions, and these divergences are removed by renormalisation of the mass coupling and fields. But once we know that we can find a set of renormalized parameters e_R, m_R, ψ_R, A_R such that eq. (12.10) has finite coefficients at each order, it is clear that we can find as many other sets as we like. Indeed, if we chose $e'_R, m'_R, \psi'_R, A'_R$ such that m_R and m'_R (and all other parameters) are related by a finite series, then

$$\mathcal{S}'(e'_R, m'_R, \psi'_R, A'_R) = \mathcal{S}'_0(e'_R, m'_R, \psi'_R, A'_R) + (e'_R)^2\ \mathcal{S}'_1(e'_R, m'_R, \psi'_R, A'_R) + \mathcal{O}((e'_R)^4) \tag{12.11}$$

is also finite at each order in perturbation theory. In other words, the divergent pieces of the renormalization factors in eq. (12.9) are uniquely determined by requiring that the divergences cancel. However, we are completely free to fix the finite pieces to whatever we want. Choosing a particular set of renormalized quantities, that is, giving some prescription on how to fix the finite pieces of the renormalization factors, is called choosing the *renormalization scheme*. It is possible in QED that $m_R = m$ and $e_R = e$, i.e. the renormalized coupling is determined by real electron photon scattering. The renormalization scheme that satisfies these constraints is called the on-shell scheme. Alternatively, the renormalized coupling may be determined by scattering with, for example, a virtual photon. In this case the value of e_R will depend on the scale of the scattering, i.e. the coupling will "run" with the renormalization scale. To be precise let me also mention that one more constraint is needed to fix the scheme completely. Naively you would expect that four constraints are needed, since we have four renormalization factors to fix. However, two of them are related, $Z_1 = Z_2$. This identity is due to gauge invariance and is called the Ward identity. As a result, we only need three constraints to fix the renormalization scheme completely.

Of course, the result of our calculation has to be independent of the renormalization scheme. This remark is not quite as innocuous as it looks. In fact, it is only true up to the order to which we decided to compute. If we decide to include the $\mathcal{O}(e_R^2)$ but not the higher order terms in our calculation, we have

$$\mathcal{S}(e_R, m_R, \psi_R, A_R) - \mathcal{S}'(e'_R, m'_R, \psi'_R, A'_R) = \mathcal{O}(e_R^4) \tag{12.12}$$

The numerical result for our prediction will depend on the renormalization scheme! Even though the difference is formally of higher order it still can be numerically significant, in particular in QCD.

12.4 Regularization

What we have learnt so far is that we have to express the result of our calculation in terms of renormalized quantities rather than the bare ones. But since the starting point of any calculation is the Lagrangian, the first step in any calculation is to get the results in terms of bare quantities. Only then, we replace the bare quantities by the renormalized quantities, using eq. (12.9) and get a finite result. In intermediate steps we will have to deal with divergent expressions.

In order to give a mathematical meaning to these intermediate expressions, we will have to regularize the integrals. That is, we have to change them in a systematic way, such that they become finite. By doing so, we change the value of the integrals. However, at the end of our calculation, we are able to undo this change. Since the final result is finite, this step will not introduce a singularity.

There are — at least in principle — many different possibilities for regularizing the integrals. To illustrate the idea of regularization I will discuss first the method of introducing a cutoff, even though in practice this method is not really used. Consider again the vertex correction in eq. (12.5). As we saw, we got the UV singularity from the region $k \to \infty$. To regularize this expression, we introduce a cutoff Λ

$$\mathcal{V} \to \mathcal{V}_{\rm reg} \sim \int^{\Lambda} \frac{d^4k}{(2\pi)^4} \frac{k\ k}{k^2((p_b+k)^2-m^2)((p_a-k)^2-m^2)} \tag{12.13}$$

Of course, by doing so we changed the value of the integral. At the end of our calculation we will have to let $\Lambda \to \infty$. Introducing this cutoff, however, gives us the possibility to deal with such intermediate expressions.

Let me illustrate the interplay between renormalization and regularization with an oversimplified example. Assume that with the cutoff regularization we get as a result of our calculation of some physical quantity, say a cross section

$$\mathcal{S} = e_0^2 A + e_0^4 \left(B \ln \frac{\Lambda}{\mu_R} + F_{\mathcal{S}}(\mu_R) \right) + \mathcal{O}(e_0^6) \tag{12.14}$$

where A, B and $F_{\mathcal{S}}$ are some *finite* terms. μ_R is called the "renormalisation scale", and is introduced as a representation of the physical scale at which the physics happens and separates the loop correction out into finite and divergent terms (we could, for example choose μ_R equal to the physical mass of a particle). The originally divergent expression for $\mathcal{S}$ has been rendered finite by regularization. At this point we cannot let $\Lambda \to \infty$ since we would get $\mathcal{S} \to \infty$. However, we learnt that we have to express our results in terms of e_R and not e_0 (For simplicity, I ignore the mass renormalization). This step is renormalization (not regularization). Computing the relation between e_0 and e_R, using the same regularization, we would find

$$e_R(\mu_R) = e_0 + e_0^3 \left(C \ln \frac{\Lambda}{\mu_R} + F_e(\mu_R) \right) + \mathcal{O}(e_0^5) \tag{12.15}$$

and reversing this

$$e_0 = e_R(\mu_R) - e_R^3(\mu_R) \left(C \ln \frac{\Lambda}{\mu_R} + F_e(\mu_R) \right) + \mathcal{O}(e_R^5) \tag{12.16}$$

where C and F_e are also finite. Plugging in eq. (12.16) into eq. (12.14) we get

$$\mathcal{S} = e_R^2(\mu_R)A + e_R^4(\mu_R)\left((B-2AC)\ln\frac{\Lambda}{\mu_R} + F_{\mathcal{S}}(\mu_R) - 2AF_e(\mu_R)\right) + \mathcal{O}(e_R^6) \tag{12.17}$$

and we would find $(B-2AC)=0$. Since QED is a renormalizable theory this 'miracle' would happen for any measurable quantity. In this example the divergences in $\mathcal{S}$ would be coming from subdiagrams which already lead to divergences, and hence renormalisation for e. Finally, in the expression

$$\mathcal{S} = e_R^2(\mu_R)A + e_R^4(\mu_R)\left(F_{\mathcal{S}}(\mu_R) - 2AF_e(\mu_R)\right) + \mathcal{O}(e_R^6) \tag{12.18}$$

we can let $\Lambda \to \infty$ and 'undo' the regularization.

To summarize, regularization enables us to work with divergent intermediate expressions. In the example above, instead of writing ∞ we write $\log\Lambda$ and have in mind $\Lambda \to \infty$. Renormalization, on the other hand removes the (would be) singularities, i.e. it removes the $\log\Lambda$ terms. Therefore, after renormalization we can (and have to) undo the regularization.

Note that we could have defined a different renormalized coupling

$$\tilde{e}_R = e_0 + e_0^3\left(C\ln\frac{\Lambda}{\tilde{\mu_R}} + G_e(\tilde{\mu_R})\right) + \mathcal{O}(e_0^5) \tag{12.19}$$

and this would have lead to

$$\mathcal{S} = \tilde{e}_R^2(\tilde{\mu_R})A + \tilde{e}_R^4(\tilde{\mu_R})(F_{\mathcal{S}}(\tilde{\mu_R}) - 2AG_e(\tilde{\mu_R})) + \mathcal{O}(\tilde{e}_R^6) \tag{12.20}$$

and we would have a different expression in terms of a different coupling - both equally valid, and identical up to the $\mathcal{O}(\tilde{e}_R^6)$ corrections.

As mentioned above, the method of introducing a cutoff for regularization is hardly ever used in actual calculations. The by far most popular method is to use dimensional regularization. The basic idea is to do the calculation not in 4 space-time dimensions but rather in D dimensions. Why does this help?

Consider once more our initial example of the vertex correction in eq. (12.5), which has an UV singularity in $D=4$ space-time dimensions (see eq. (12.6)). For arbitrary D, using $d^Dk \sim k^{D-1}dk$ we get

$$\mathcal{V} \sim \int \frac{d^Dk}{(2\pi)^4}\frac{1}{k^4} \sim \int \frac{dk}{(2\pi)}k^{D-5} \tag{12.21}$$

and the integral is UV-finite for say $D \leq 4$. Thus changing the dimension can regulate integrals. It is important to note that this is only a technicality. There is no Physics associated with $D \neq 4$ and at the end of the calculation we have to let $D \to 4$. If we did renormalize our theory properly this last step will not lead to UV divergences.

The reason why dimensional regularization is so popular is that it preserves gauge invariance and is technically relatively simple. Another very important issue is that this regularization not only regulates UV singularities, but also IR singularities. As mentioned in section 12.2, theories like QED or QCD are very often plagued by IR singularities. It is therefore very convenient if we do not have to introduce another regularization for IR singularities. Only after all UV and IR singularities have been removed, we can let $D \to 4$ and finally obtain a finite result.