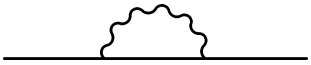
2012
Matthew Schwartz

III-4: Mass renormalization

1 Introduction

In this lecture we will study the following 1-loop Feynman diagram:



known as the **electron self-energy graph**. You may recall we encountered this diagram way back in Lecture I-4 in the context of Oppenheimer's Lamb shift calculation using old-fashioned perturbation theory. Indeed, this graph is important for the Lamb shift. However, rather than compute the Lamb shift (which is rather tedious and mostly of historical interest for us), we will use this graph to segue to a more general understanding of renormalization. You may also recall Oppenheimer's frustrated comment, quoted at the end of Lecture I-4, where he suggested that the resolution of these infinities would require an "adequate theory of the masses of the electron and proton". In this lecture, we will provide such an adequate theory.

The electron self-energy graph corrects the electron propagator in the same way that the photon self-energy graph corrects the photon propagator. Recall from Lecture III-2 that the photon-self energy graph could be interpreted as a vacuum polarization effect which generated a logarithmic weakening of the Coulomb potential at large distances. Thus by measuring $r_1V(r_1) - r_2V(r_2)$ with two different values of r one could measure vacuum polarization and compare it to the theoretical prediction. In particular, we were able to renormalize the divergent vacuum polarization graph by relating it to something (the potential) which can be directly connected to observables (e.g. the force between two currents or the energy levels of Hydrogen).

Proceeding in the same way, the electron self-energy graph would correct the effect generated by the exchange of an electron. However, since the electron is a fermion, and charged, this exchange cannot be interpreted as generating a potential in any useful way. Thus it is not clear what exactly one would measure to test whatever result we find by evaluating the self-energy diagram.

Thus, for the self-energy graph, and many other divergent graphs we will evaluate, it is helpful to navigate away from observables like the Lamb shift or the Coulomb potential which are particular to one type of correction, to thinking of general observables. Unfortunately the question of what is observable and what is not is extremely subtle and has no precise definition in quantum field theory. For example, one might imagine that S-matrix elements are observable; in many cases they are actually infinite due to infrared divergences, as we will see in Lecture III-6. Luckily, one does not need a precise definition of an observable to understand renormalization, since even non-observable quantities can be renormalized. We will therefore consider the renormalization of general time-ordered correlation functions or **Green's functions**

$$G(x_1, \cdots, x_n) = \langle \Omega | T\{\phi_1(x_1) \cdots \phi_n(x_n)\} | \Omega \rangle \tag{1}$$

where ϕ_i can be any type of field (scalars, electrons, photons, etc.). These Green's functions are in general *not* observable. In fact, they are in general not even gauge invariant. We will nevertheless show within a few lectures that all ultraviolet divergences can be removed from all Green's functions in *any* local quantum field theory through a systematic process of renormalization. Once the Green's functions are UV finite, S-matrix elements constructed from them using the LSZ reduction formula will also be UV finite. Infrared divergences and what can actually be observed is another matter.

One advantage of renormalizing general Green's functions rather than S-matrix elements is that the Green's functions can appear as internal sub-graphs in many different S-matrix calculations. In particular, we will find that in QED while there are an infinite number of divergent graphs contributing to the S-matrix, the divergences can be efficiently categorized and renormalized through the **one-particle irreducible** sub-graphs (defined as graphs which cannot be cut in two by cutting a single propagator). As we will see, these one-particle irreducible graphs compose the minimal basis of Green's functions out of which any S-matrix can be built. Organizing the discussion in terms of Green's functions and one-particle irreducible diagrams will vastly simplify the proof of renormalizability in QED (in Lecture III-7) and is critical to a general understanding of how renormalization works in various quantum field theories.

In this lecture, we abbreviate $\langle\Omega|T\{\phi_1(x_1)\cdots\phi_n(x_n)\}|\Omega\rangle$ with $\langle\cdots\rangle$ for simplicity.

2 Vacuum expectation values

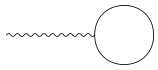
We begin our consideration of the renormalization of general Green's functions by considering the simplest Green's functions, the one-point functions:

$$\langle\phi(x)\rangle, \quad \langle\psi(x)\rangle, \quad \langle A_\mu(x)\rangle, \cdots \tag{2}$$

These give the expectation value of fields in the vacuum, also known as **vacuum expectation values**.

At tree-level, the vacuum expectation value of a field is the lowest energy configuration which satisfies the classical equations of motion. All Lagrangians we have considered so far begin at quadratic order in the fields, so that $\psi = A = \phi = 0$ are solutions to the equations of motion. Other solutions, such as plane waves in the free theory, contribute to the gradient terms in the energy density and thus have higher energy than the constant solution. Thus, $\psi = A = \phi = 0$ is the minimum energy solution and all the expectation values in Eq. (2) vanish at tree-level. More directly, we can see that $\langle\phi\rangle = \langle\psi\rangle = \langle A_\mu\rangle = 0$ at tree-level in the canonically quantized theory, since each quantum field has creation and annihilation operators which vanish in the vacuum.

At 1-loop, vacuum expectation values, for example for $\langle A_\mu\rangle$ could come from diagrams like



(3)

This is called a **tadpole** diagram. It and all higher-loop contributions to $\langle A_\mu\rangle$ vanish identically in QED. This is easy to see in perturbation theory since all fermion loops with an odd number of photons attached involve a trace over an odd number of γ matrices, which vanishes. It is also true that $\langle\psi\rangle = 0$ to all orders in QED, simply because one cannot draw any diagrams.

A somewhat simpler proof that $\langle A_\mu\rangle$ or $\langle\psi\rangle$ must vanish is that nonzero values would violate Lorentz invariance, and Lorentz invariance is a symmetry of the QED Lagrangian. However, it may sometimes happen that the vacuum does not in fact satisfy every symmetry of the Lagrangian, in which case we say **spontaneous symmetry breaking** has occurred. Spontaneous symmetry breaking is covered in depth in Lecture IV-4. A familiar example is the spontaneous breaking of rotational invariance by a ferromagnet when cooled below its Curie temperature. At low temperature, the magnet has a preferred spin direction which could equally well have pointed anywhere, but must point somewhere. Another example is the ground state of our universe which has a preferred frame, the rest frame of the cosmic microwave background. In both cases space-time symmetries are symmetries of the Lagrangian but not of the ground state.

Spontaneous symmetry breaking can also apply to internal symmetries, such as global or gauge symmetries of a theory. For example, in the Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity the $U(1)$ symmetry of QED is spontaneously broken in type II superconductors as they are cooled below their critical temperature. The attractive force between electrons due to phonon exchange becomes stronger than the repulsive Coulomb force and the vacuum becomes charged. Another important example is the Glashow-Weinberg-Salam (GWS) theory of weak interactions (Lecture IV-5). This theory embeds the low-energy theory of weak interactions into a larger theory which has an exact SU(2) symmetry which acts on the left-handed quarks and leptons.

Spontaneous symmetry breaking is an immensely important topic in quantum field theory, which we will systematically discuss beginning in Lecture IV-4, including more details of the above examples. Now, it is merely a distraction from our current task of understanding renormalization. Since $\langle A_\mu \rangle = \langle \psi \rangle = 0$ in QED to all orders in perturbation theory, there is nothing to renormalize and we can move on to two-point functions.

3 Electron self-energy

There are a number of two-point functions in QED. In Lecture II-2, we discussed the renormalization of the photon propagator which corresponds to $\langle A_\mu A_\nu \rangle$. Two-point functions like $\langle \psi A_\mu \rangle$ vanish identically in QED since there are simply no diagrams which could contribute to it. That leaves the fermion two-point function $\langle \psi \bar{\psi} \rangle$.

As with the photon, it is helpful to study $\langle \psi \bar{\psi} \rangle$ in momentum space. We define the momentum space Green's function by

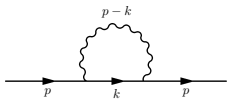
$$\langle \psi(x) \bar{\psi}(y) \rangle = \int \frac{d^4p}{(2\pi)^4} e^{-ip(x-y)} iG(\not{p}) \tag{4}$$

This is possible since the left-hand side can only depend on $x - y$ by translation invariance.

At tree level, $G(\not{p})$ is just the momentum space fermion propagator

$$iG_0(\not{p}) \equiv \frac{i}{\not{p} - m} \tag{5}$$

At 1-loop it gets a correction due to the self-energy graph

$iG_2(\not{p}) =$ 

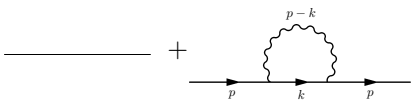
$= iG_0(\not{p})[i\Sigma_2(\not{p})]iG_0(\not{p})$

where, in Feynman gauge,

$$i\Sigma_2(\not{p}) = (-ie)^2 \int \frac{d^4k}{(2\pi)^4} \gamma^\mu \frac{i(\not{k} + m)}{k^2 - m^2 + i\varepsilon} \gamma^\mu \frac{-i}{(p-k)^2 + i\varepsilon} \tag{6}$$

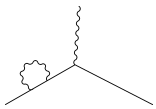
If this graph were contributing to an S-matrix element, rather than just a Green's function, we would remove the propagators from the external lines (the G_0 factors in Eq. (5)) and contract with external on-shell spinors. This $i\Sigma_2(\not{p})$ is what we would get from the normal Feynman rules without the external spinors.

Before evaluating this graph, we can observe an interesting feature that was not present in the photon case (the vacuum polarization graph). Including the self-energy graph, the *effective* electron propagator to 1-loop is

$iG(\not{p}) =$ 

$= \frac{i}{\not{p} - m} + \frac{i}{\not{p} - m} i\Sigma_2(\not{p}) \frac{i}{\not{p} - m} + \mathcal{O}(e^4)$ (7)

In an S-matrix element, this correction might appear on an external leg, like

 (8)

In that case $G(\not{p})$ is contracted with an on-shell external spinor and the result multiplied by a factor of $\not{p}-m$ from the LSZ reduction formula. Now, there is no reason to expect that $\Sigma_2(m)=0$ (and in fact it is not), so even after removing a single pole with $\not{p}-m$ we see from Eq. (7) that there will still be a pole left over. That is, the S-matrix will be singular. This problem did not come up for the photon propagator and vacuum polarization, where the corrected photon propagator had only a single pole to all orders in perturbation theory. The resolution of this apparently singular S-matrix for electron scattering is that the electron mass appearing in the LSZ formula does not necessarily have to match the electron mass appearing in the Lagrangian. In the photon case, they were equal since both were zero. Once we evaluate the self-energy graph, we will then discuss how the electron mass is renormalized and why the S-matrix remains finite.

3.1 Self-energy loop graph

Evaluating the self-energy graph with Feynman parameters (see Appendix B) gives

$$\begin{aligned} i\Sigma_2(\not{p}) =& (-ie)^2\int\frac{d^4k}{(2\pi)^4}\gamma^\mu\frac{i(\not{k}+m)}{k^2-m^2+i\varepsilon}\gamma^\mu\frac{-i}{(k-p)^2+i\varepsilon} \\ =& e^2\int\frac{d^4k}{(2\pi)^4}\int_0^1 dx\frac{2\not{k}-4m}{[(k^2-m^2)(1-x)+(p-k)^2x+i\varepsilon]^2} \end{aligned}$$

Now we complete the square in the denominator and shift $k\to k+px$ to give

$$i\Sigma_2(\not{p}) = 2e^2\int_0^1 dx\int\frac{d^4k}{(2\pi)^4}\frac{x\not{p}-2m}{[k^2-\Delta+i\varepsilon]^2} \tag{9}$$

where $\Delta=(1-x)(m^2-p^2x)$ and we have dropped the term linear in k in the numerator since it is odd under $k\to -k$ and its integral therefore vanishes. This integrand scales like $\frac{d^4k}{k^4}$ and is therefore logarithmically divergent in the UV.

To regulate the UV divergence, we have to choose a regularization scheme. For pedagogical purposes we will evaluate this loop with both Pauli-Villars (PV) and dimensional regularization (DR). Recall (from Appendix B) that Pauli-Villars introduces heavy particles, of mass Λ with negative energy for each physical particle in the theory. Pauli-Villars is nice because the scale Λ is clearly a UV deformation, with the Pauli-Villars ghosts having no effect on the low energy theory as $\Lambda\to\infty$. In dimensional regularization, which analytically continues to $4-\varepsilon$ dimensions, it is not clear that ε is a UV deformation in any sense. Dimensional regularization is much easier to use for more complicated theories than QED, so eventually we will use it exclusively. For now, it is helpful to use two regulators to see that results are regulator independent.

With a PV photon, the self-energy graph becomes

$$\Sigma_2(\not{p}) = -2ie^2\int_0^1 dx(x\not{p}-2m)\int\frac{d^4k}{(2\pi)^4}\left[\frac{1}{(k^2-\Delta)^2}-\frac{1}{(k^2-\Delta')^2}\right] \tag{10}$$

with $\Delta'=(1-x)(m^2-p^2x)+x\Lambda^2$. Since we take $\Lambda\to\infty$ we can more simply take $\Delta'=x\Lambda^2$. The regulated integral is now convergent and can be evaluated using formulas from Appendix B. The result is

$$\begin{aligned} \Sigma_2(\not{p}) =& -\frac{\alpha}{2\pi}\int_0^1 dx(2m-x\not{p})\ln\frac{x\Lambda^2}{(1-x)(m^2-p^2x)} \\ =& -\frac{\alpha}{\pi}\left(m\ln\Lambda^2-\frac{1}{4}\not{p}\ln\Lambda^2+\text{finite}\right) \end{aligned} \qquad \text{(PV)} \tag{11}$$

In dimensional regularization, the loop is

$$\begin{aligned} \Sigma_2(\not{p}) =& -2ie^2\mu^{4-d}\int_0^1 dx(x\not{p}-2m)\int\frac{d^dk}{(2\pi)^d}\frac{1}{(k^2-\Delta+i\varepsilon)^2} \\ =& -\frac{\alpha}{2\pi}\int_0^1 dx(2m-x\not{p})\left[\frac{2}{\varepsilon}+\ln\frac{\tilde{\mu}^2}{(1-x)(m^2-p^2x)}\right] \end{aligned} \qquad \text{(DR)} \tag{12}$$

where $\tilde{\mu}^2 \equiv 4\pi e^{-\gamma_E}\mu^2$. Extracting the divergent parts, the loop can be written as

$$\Sigma_2(\not{p}) = \frac{\alpha}{\pi}\left(\frac{\not{p} - 4m}{2\varepsilon} + \text{finite}\right) \tag{13}$$

Note that in both cases $\Sigma_2(m) \neq 0$, so there will be a double-pole at 1-loop with the possibly dangerous consequences discussed below Eq. (7). Note that both results have divergences proportional to both m and to $\not{p}$. This implies that we need two quantities to renormalize, to remove both divergences.

3.2 Renormalization

As discussed in the introduction, we want the Green's function $G(\not{p})$ defined in Eq. (4) to be finite. Thus the infinities from the $\mathcal{O}(e^2)$ contribution to this Green's function must be removed through renormalization.

As with the vacuum polarization, we need to figure out what parameters in the theory can be renormalized to cancel the infinities in the self-energy graph. To begin, let us write the Lagrangian as

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + i\bar{\psi}\not{\partial}\psi - m_0\bar{\psi}\psi - e_0\bar{\psi}\not{A}\psi \tag{14}$$

In the study of vacuum polarization in Lecture II-2, we concluded that the charge in the Lagrangian, now written as e_0, called the **bare charge**, could be used to absorb an infinity. Recall that we defined a renormalized electric charge via

$$e_0^2 = e_R^2 + e_R^4\Pi_2(p_0^2) + \cdots = e_R^2\left(1 - \frac{e_R^2}{12\pi^2}\ln\frac{\Lambda^2}{-p_0^2} + \cdots\right) \tag{15}$$

where $\Pi_2(p_0^2)$ is formally infinite. Since e_0 has already been renormalized by vacuum polarization, we cannot renormalize it in a different way for the self-energy graph.

To make $G(\not{p})$ finite the obvious Lagrangian parameter which might absorb the infinity is the **bare electron mass** m_0. Indeed, from Eq. (7),

$$iG_2(\not{p}) = \frac{i}{\not{p} - m_0} + \frac{i}{\not{p} - m_0}[i\Sigma_2(\not{p})]\frac{i}{\not{p} - m_0} \tag{16}$$

we can see that an (infinite) redefinition of $m_0 = m + \Delta m$ with Δm of order e^2 could compensate for an infinity at order e^2 in $\Sigma_2(\not{p})$. Unfortunately, we saw in Eqs. (11) and (13) that $\Sigma_2(\not{p})$ has two types of infinities, one independent of $\not{p}$ and the other proportional to $\not{p}$. The mass renormalization can only remove one of these infinities. Thus, to progress further we need something else to renormalize. But what could it be? Our Lagrangian only had two parameters, m and e, and we've already defined how e gets renormalized.

In fact, there is another parameter: the normalization of the fermion wavefunction. Let us write the fermion field in terms of creation and annihilation operators that we have been using all along as the **bare free field**:

$$\psi^0(x) = \sum_s \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}(a_p^s u_p^s e^{-ipx} + b_p^{s\dagger} v_p^s e^{ipx}) \tag{17}$$

The bare free field is canonically normalized to give all the tree-level scattering results we have already calculated. We then define the **renormalized field** as

$$\psi^R(x) = \frac{1}{\sqrt{Z_2}}\sum_s \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}(a_p^s u_p^s e^{-ipx} + b_p^{s\dagger} v_p^s e^{ipx}) \equiv \frac{1}{\sqrt{Z_2}}\psi^0 \tag{18}$$

for some (formally infinite) number Z_2. This is the origin of the term *renormalization*. We index bare (infinite) fields and parameters with a 0 and physical finite renormalized fields and parameters with an R.

For the tree-level theory, $Z_2 = 1$ is required to be consistent with the normalization used in all our scattering formulae. So it is natural to account for radiative corrections by writing

$$Z_2 = 1 + \delta_2 \tag{19}$$

where δ_2 is the **counterterm**, which has a formal Taylor series expansion in e starting at order e^2. We also write

$$m_0 = Z_m m_R \tag{20}$$

and expand $Z_m = 1 + \delta_m$, with δ_m the mass counterterm.[1] Then

$$m_0 = m_R + m_R \delta_m \tag{21}$$

As we will see, particularly when we cover renormalized perturbation theory in Lecture III-6, using counterterms rather than bare and renormalized quantities directly will be extremely efficient.

All the calculations we have done so far have been with fields with the conventional (bare) normalization. However, it is the Green's function of *renormalized* fields which should have finite physical values. So we define

$$\langle \psi^0(x)\bar{\psi}^0(y)\rangle = i\int \frac{d^4p}{(2\pi)^4} e^{-ip(x-y)} G^{\text{bare}}(\not{p}) \tag{22}$$

and

$$\langle \psi^R(x)\bar{\psi}^R(y)\rangle = i\int \frac{d^4p}{(2\pi)^4} e^{-ip(x-y)} G^{R}(\not{p}) \tag{23}$$

and expect $G^R(\not{p})$ to be finite. By definition

$$G^R(\not{p}) = \frac{1}{Z_2} G^{\text{bare}}(\not{p}) \tag{24}$$

Now since Z_2 is just a number, the tree-level propagator for the renormalized fields can be expressed in terms of the propagator of the bare fields as

$$\begin{aligned} iG^R(\not{p}) &= \frac{1}{Z_2}\frac{i}{\not{p} - m_0} + \text{loops} \\ &= \left(\frac{1}{1+\delta_2}\right)\left(\frac{i}{\not{p} - m_R - \delta_m m_R}\right) + \text{loops} \\ &= \frac{i}{\not{p} - m_R} + \frac{i}{\not{p} - m_R}[i(\delta_2 \not{p} - (\delta_2 + \delta_m)m_R)]\frac{i}{\not{p} - m_R} + \text{loops} + \mathcal{O}(e^4) \end{aligned} \tag{25}$$

Adding the 1-loop contribution as in Eqs. (7) or (16) gives

$$iG^R(\not{p}) = \frac{i}{\not{p} - m_R} + \frac{i}{\not{p} - m_R}[i(\delta_2 \not{p} - (\delta_2 + \delta_m)m_R + \Sigma_2(\not{p}))]\frac{i}{\not{p} - m_R} + \mathcal{O}(e^4) \tag{26}$$

So now we can choose δ_2 and δ_m to remove all the infinities in the electron self-energy.

To be explicit, from Eq. (11) we see that choosing

$$\delta_2 = -\frac{\alpha}{4\pi}\ln\Lambda^2, \quad \delta_m = -\frac{3\alpha}{4\pi}\ln\Lambda^2 \quad \text{(PV)} \tag{27}$$

for Pauli-Villars or

$$\delta_2 = -\frac{\alpha}{4\pi}\frac{2}{\varepsilon}, \quad \delta_m = -\frac{3\alpha}{4\pi}\frac{2}{\varepsilon} \quad \text{(DR)} \tag{28}$$

for dimensional regularization will remove the infinities. With these choices, we will get finite answers for the two point function $G^R(\not{p})$ at any scale p.

We can choose different values for the counterterms which differ from these by finite numbers and $G^R(\not{p})$ will still be finite. Any prescription for choosing the finite parts of the counterterms is known as a **subtraction scheme**. Not only must observables in a renormalized theory be finite, but they also must be independent of the subtraction scheme, as we will see. Nevertheless, there are some smart choices for subtraction schemes and some not-so-smart choices.

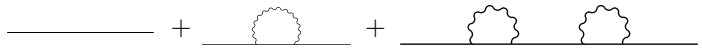
1. Another common convention is $Z_2 m_0 \equiv m_R + \delta_m$. Our convention is more commonly used in modern field theory calculations.

The two subtraction schemes most often used in quantum field theory are the *on-shell* subtraction scheme and the *minimal subtraction* (MS) scheme. Minimal subtraction is by far the simplest scheme and the one used in almost all modern quantum field theory calculations. In minimal subtraction the counterterms are defined to have no finite parts at all, so that δ_2 and δ_m are given by Eqs. (27) and (28). More commonly, a slightly modified version of this prescription known as modified minimal subtraction ($\overline{\text{MS}}$) is used, in which $\ln(4\pi)$ and γ_E finite parts in dimensionally regulated results are also subtracted off. $\overline{\text{MS}}$ just turns $\tilde{\mu}$ back into μ in dimensionally regularized amplitudes.

In on-shell subtraction the renormalized mass m_R appearing in Green's functions is identified with the observed electron mass m_P which can be defined to all orders as the position of the pole in the S-matrix.[2] To see how this identification works in practice, it is helpful to look at the possible form of the higher order corrections.

4 The pole mass

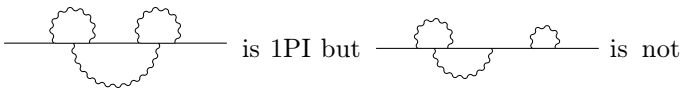
So far, we have only included one particular self-energy correction. The 2-point function $G(\not{p})$ in fact gets corrections from an infinite number of graphs. One particular series of corrections, of the form

$iG^{\text{bare}}(\not{p}) =$  (29)

just produces a geometric series

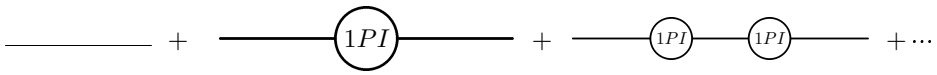
$$iG^{\text{bare}}(\not{p}) = \frac{i}{\not{p}-m_0} + \frac{i}{\not{p}-m_0}\Big(i\Sigma_2(\not{p})\Big)\frac{i}{\not{p}-m_0} + \frac{i}{\not{p}-m_0}\Big(i\Sigma_2(\not{p})\Big)\frac{i}{\not{p}-m_0}\Big(i\Sigma_2(\not{p})\Big)\frac{i}{\not{p}-m_0} + \cdots \tag{30}$$

which is easy to sum. More generally, any possible graph contributing to this Green's function is part of *some* geometric series. Conversely, the entire Green's function can be written as the sum of a single geometric series constructed by sewing together graphs which cannot be cut in two by slicing a single propagator. We call such graphs **one-particle irreducible (1PI)**. For example,



(31)

Thus

$iG(\not{p}) =$ 

Defining $i\Sigma(\not{p})$ as the sum of all of the 1PI graphs, we find

$$\begin{aligned} iG(\not{p}) &= \frac{i}{\not{p}-m} + \frac{i}{\not{p}-m}\Big(i\Sigma(\not{p})\Big)\frac{i}{\not{p}-m} + \frac{i}{\not{p}-m}\Big(i\Sigma(\not{p})\Big)\frac{i}{\not{p}-m}\Big(i\Sigma(\not{p})\Big)\frac{i}{\not{p}-m} + \\ &= \frac{i}{\not{p}-m}\left(1 + \frac{-\Sigma(\not{p})}{\not{p}-m} + \left(\frac{-\Sigma(\not{p})}{\not{p}-m}\right)^2 + \cdots\right) \\ &= \frac{i}{\not{p}-m}\frac{1}{1+\frac{\Sigma(\not{p})}{\not{p}-m}} \\ &= \frac{i}{\not{p}-m+\Sigma(\not{p})} \end{aligned} \tag{32}$$

2. Actually, there isn't an isolated pole in the S-matrix associated with the electron. Rather, the electron mass is the beginning of a cut in the complex plane. This will be discussed more in Lecture III-10.

This is just a general expression for a sum of Feynman diagrams, applying whether $m = m_0$ or $m = m_R$. For the bare Green's function, there was just a single 1PI diagram at order e^2 and so $\Sigma(\not p) = \Sigma_2(\not p) + \mathcal{O}(e^4)$. Then we have

$$iG^{\rm bare}(\not p) = \frac{i}{\not p - m_0 + \Sigma_2(\not p) + \cdots} \tag{33}$$

This expression is the sum of the series in Eq. (29).

From the bare Green's function we can compute the renormalized Green's function as

$$iG^R(\not p) = \frac{1}{1+\delta_2} iG^{\rm bare}(\not p) = \left(\frac{1}{1+\delta_2}\right)\frac{i}{\not p - m_0 + \Sigma_2(\not p) + \cdots} = \frac{i}{\not p - m_0 + \delta_2 \not p - m_0\delta_2 + \Sigma_2(\not p) + \cdots}$$

where the $\cdots$ are formally $\mathcal{O}(e^4)$ or higher. Then using Eq.(21), $m_0 = m_R + m_R\delta_m$, this becomes

$$iG^R(\not p) = \frac{i}{\not p - m_R + \delta_2 \not p - (\delta_2 + \delta_m) m_R + \Sigma_2(\not p) + \cdots} \tag{34}$$

We will write this more conveniently as

$$iG^R(\not p) = \frac{i}{\not p - m_R + \Sigma_R(\not p)} \tag{35}$$

with $\Sigma_R(\not p) = \Sigma_2(\not p) + \delta_2 \not p - (\delta_m + \delta_2) m_R + \mathcal{O}(e^4)$.

You may have noted that this result would follow easily from Eq. (26) if we could treat the counterterms as contributions to 1P1 graphs. To justify such treatment, all we have to do is rewrite the bare free Lagrangian in terms of renormalized fields

$$\mathcal{L} = i\bar\psi^{\,0}\not\partial\psi^0 - m_0\bar\psi^{\,0}\psi^0 = i\,Z_2\bar\psi^{\,R}\not\partial\psi^R - Z_2Z_m m_R\bar\psi^{\,R}\psi^R \tag{36}$$

using Eqs. (19) and (20) this becomes

$$\mathcal{L} = i\,\bar\psi^{\,R}\not\partial\psi^R - m^R\bar\psi^{\,R}\psi^R + i\delta_2\bar\psi^{\,R}\not\partial\psi^R - m_R(\delta_2+\delta_m)\bar\psi^{\,R}\psi^R \tag{37}$$

Thus we can treat the counterterms, which start at order e^2, as interactions whose Feynman rules give contributions $\delta_2 \not p$ and $-(\delta_2 + \delta_m)m_R$ to the 1PI graphs. Then Eq. (35) follows from the general form Eq. (32) with $m = m_R$ and $\Sigma = \Sigma_R$. Expanding the Lagrangian in terms of renormalized quantities leads to so-called **renormalized perturbation theory**. Renormalized perturbation theory will be discussed more completely, including interactions and the photon field, in the next lecture.

4.1 On-shell subtraction

Having summed all of the 1PI diagrams into the renormalized propagator, we can now identify the physical electron mass m_P as the location of its pole. More precisely, the renormalized propagator should have a single pole at $\not p = m_P$ with residue i. The location of the pole is a *definition* of mass, known as the **pole mass**. It is important to keep in mind that the pole mass is physical and independent of any subtraction scheme used to set the finite parts of the counterterms. In the on-shell subtraction scheme, the finite parts of the counterterms are chosen so that $m_R = m_P$. In minimal subtraction, $m_R \neq m_P$. In either case the 2-point Green's function still has a pole at m_P.

From Eq. (35), for $G^R(\not p)$ to have a pole at $\not p = m_P$ the 1PI graphs must satisfy

$$\boxed{\Sigma_R(m_P) = m_R - m_P} \tag{38}$$

Having residue i implies

$$i = \lim_{\not p \to m_P} (\not p - m_P)\frac{i}{\not p - m_R + \Sigma_R(\not p)} = \lim_{\not p \to m_P} \frac{i}{1 + \frac{d}{d\not p}\Sigma_R(\not p)} \tag{39}$$

where we have used L'Hôpital's rule. This implies

$$\boxed{\frac{d}{d\not p}\Sigma_R(\not p)\bigg|_{\not p = m_P} = 0} \tag{40}$$

These conditions define the pole mass, independent of subtraction scheme.

In the **on-shell subtraction scheme**, the renormalized mass m_R is set equal to the pole mass m_P. Then, recalling $\Sigma_R(\not p) = \Sigma_2(\not p) + \delta_2 \not p - (\delta_m + \delta_2) m_R + \cdots$, these conditions imply to order e^2

$$\delta_2 = -\frac{d}{d\not p}\Sigma_2(\not p)\bigg|_{\not p = m_P} \tag{41}$$

and

$$\delta_m m_P = \Sigma_2(m_P) \tag{42}$$

which we can now evaluate in our different regulators.

With Pauli-Villars, Eq. (12) implies

$$\Sigma_2(m_P) = -\frac{\alpha}{2\pi} m_P \left(\frac{3}{2} \ln \frac{\Lambda^2}{m_P^2} + \frac{3}{4} \right) \qquad \text{(PV)} \tag{43}$$

which is one of our conditions. Unfortunately, when we try to evaluate the derivative, we find

$$\frac{d}{d\not p}\Sigma(\not p)\bigg|_{\not p = m_P} = \frac{\alpha}{2\pi}\left(\frac{1}{2}\ln\frac{\Lambda^2}{m_P^2} + \frac{5}{4} - \int_0^1 dx \frac{2x(2-x)}{1-x} \right) \qquad \text{(PV)}$$

This last integral is divergent. This divergence is an **infrared divergence**, due to the integration region near $k^2 = 0$. In this case, the divergence does not come from the loop integral itself, but from our choice of subtraction scheme which involved $\Sigma'(m_P)$. Nevertheless, infrared divergences in renormalized Green's functions and S-matrix elements are unavoidable. We will see how they drop out of physical observables in Lecture III-6.

For now, a quick way to sequester the infrared divergence is to pretend that the photon has a tiny mass m_γ. As with ultraviolet divergences, infrared divergences will cancel in physical processes, so we will eventually be able to take $m_\gamma \to 0$. If you are skeptical about how this could happen, recall that in the vacuum polarization calculation at momentum transfers $-p^2 \gg m^2$, the corrections to the Coulomb potential were independent of m. In fact, the vacuum polarization graph would be infrared divergent if we set $m = 0$ before evaluating the loop. Thus at very short distances, the electron mass acts only as a regulator, just as m_γ does here.

The effect of a photon mass is to change Δ to $\Delta = (1-x)(m_P^2 - p^2 x) + x m_\gamma^2$ so that

$$\Sigma_2(\not p) = \frac{\alpha}{2\pi}\int_0^1 dx (x\not p - 2m_P) \ln \frac{x\Lambda^2}{(1-x)(m_P^2 - p^2 x) + x m_\gamma^2} \qquad \text{(PV)} \tag{44}$$

Then, keeping only the leading terms in m_γ

$$\delta_2 = -\Sigma'(m_P) = \frac{\alpha}{2\pi}\left(-\frac{1}{2}\ln\frac{\Lambda^2}{m_P^2} - \frac{9}{4} - \ln\frac{m_\gamma^2}{m_P^2} \right) \qquad \text{(PV)}$$

which is now finite. Then

$$\delta_m = \frac{1}{m_P}\Sigma_2(m_P) = \frac{\alpha}{2\pi}\left(-\frac{3}{2}\ln\frac{\Lambda^2}{m_P^2} - \frac{3}{4} \right) \qquad \text{(PV)} \tag{45}$$

In dimensional regularization, with the photon mass added, the loop gives

$$\Sigma_2(\not p) = \frac{\alpha}{2\pi}\int_0^1 dx (x\not p - 2m)\left(\frac{2}{\varepsilon} + \ln \frac{\tilde\mu^2}{(1-x)(m_P^2 - p^2 x) + x m_\gamma^2} \right) \qquad \text{(DR)} \tag{46}$$

leading to

$$\delta_2 = -\Sigma_2'(m_P) = -\frac{\alpha}{2\pi}\left(\frac{1}{\varepsilon} + \frac{1}{2}\ln\frac{\tilde\mu^2}{m_P^2} + \frac{5}{2} + \ln\frac{m_\gamma^2}{m_P^2} \right) \qquad \text{(DR)} \tag{47}$$

and

$$\delta_m = \frac{1}{m_P}\Sigma_2(m_P) = \frac{\alpha}{2\pi}\left(-\frac{3}{\varepsilon} - \frac{3}{2}\ln\frac{\tilde\mu^2}{m_P^2} - \frac{5}{2} \right) \qquad \text{(DR)} \tag{48}$$

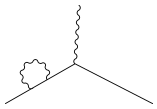
4.2 Amputation

Recall that the LSZ theorem converts Green's functions to S-matrix elements by adding external polarizations and factors of $\not{p} - m_0$ to project on to physical one-particle states. However, we have now seen that the location of the pole in the electron propagator is not the value of the mass m_0 appearing in the Lagrangian, but rather at some other location m_P. Moreover, we have found that only Green's functions of renormalized fields like $G^R \sim \langle \bar{\psi}^R \psi^R \rangle$ should be finite. Thus it would be natural to modify the LSZ theorem to

$$\langle f|S|i\rangle \sim (\not{p}_f - m_P)\cdots(\not{p}_i - m_P)\langle \psi^R \cdots \psi^R \rangle \tag{49}$$

This is almost correct.

The subtlety is that in the derivation of LSZ, we had to assume that all the interactions happened during some finite time interval and that as $t \to \pm\infty$ we could treat the theory as free. In the free theory, the pole would be at m_0. Thus we really want the theory not to be entirely free at asymptotic times, but to include all of the corrections which move the pole from m_0 to m_P. Those corrections are precisely the series of 1PI insertions onto the electron propagator. Thus in projecting onto the pole mass, with the $(\not{p} - m_P)$ factors, we must assume that all of the corrections to the on-shell external electron propagator have been included. For example, diagrams like



(50)

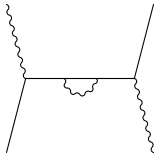
would only contribute to correcting the external electron propagator, which would then be removed by LSZ.

Thus the LSZ theorem in renormalized perturbation theory is

$$\langle f|S|i\rangle = (\not{p}_f - m_P)\cdots(\not{p}_i - m_P)\langle \psi^R \cdots \psi^R \rangle_{\text{amputated}} \tag{51}$$

where **amputated** means chop off the external lines until they begin interacting with the other fields. Only amputated diagrams contribute to S-matrix elements.

Note that amputating diagrams does not mean that self-energy graphs are never important. When a self-energy bubble occurs on an internal line, like in this correction to Compton scattering



(52)

it will have an important physical effect. All the renormalized LSZ theorem says is that you should not correct external lines for S-matrix elements since those corrections are already accounted for in the updated definition of asymptotic states.

5 Minimal subtraction

In minimal subtraction, the counterterms are fixed with no reference to the pole mass. The prescription is simply that the counterterms should have no finite parts. Thus, with Pauli-Villars, we get Eq. (27)

$$\delta_2 = -\frac{\alpha}{4\pi}\ln\Lambda^2, \quad \delta_m = -\frac{3\alpha}{4\pi}\ln\Lambda^2 \quad \text{(PV)} \tag{53}$$

and then $\Sigma_R(\not{p}) = \Sigma_2(\not{p}) + \delta_2\not{p} - (\delta_m + \delta_2)m_R$ is

$$\Sigma_R(\not{p}) = \frac{\alpha}{2\pi}\int_0^1 dx(x\not{p} - 2m_R)\ln\frac{x}{(1-x)(m_R^2 - p^2x)} \tag{54}$$

which is finite, but has nonsensical dimensions. Instead, we can modify the minimal subtraction for use with Pauli-Villars so that

$$\delta_2 = -\frac{\alpha}{4\pi}\ln\frac{\Lambda^2}{\mu^2}, \quad \delta_m = -\frac{3\alpha}{4\pi}\ln\frac{\Lambda^2}{\mu^2} \quad \text{(PV)} \tag{55}$$

with μ some arbitrary scale with dimensions of mass. μ should be thought of as a low energy scale, say 1 GeV, which is not taken to infinity. Then,

$$\Sigma_R(\not p) = \frac{\alpha}{2\pi}\int_0^1 dx(x\not p - 2m_R)\ln\frac{x\mu^2}{(1-x)(m_R^2 - p^2x)} \tag{56}$$

By introducing μ we have established a one-parameter family of subtraction schemes. Any physical observable must be independent of μ, but μ is *not* taken to infinity. μ is sometimes called the **subtraction point**.

The subtraction point already appeared in Lecture III-2 on vacuum polarization, where it was set equal to the long distance scale where the renormalized electric charge e_R was defined. As in that case, when one compares observables, such as combinations of the Coulomb potential $r_1V(r_1) - r_2V(r_2)$ measured at different scales, the subtraction point will drop out.

The subtraction point also appears as the parameter μ in dimensional regularization. Recall that in dimensional regularization μ is introduced by the rescaling $e^2 \to \mu^{4-d}e^2$ which lets the electric charge remain dimensionless in d dimensions. In dimensional regularization, minimal subtraction gives Eq. (28)

$$\delta_2 = -\frac{\alpha}{4\pi}\frac{2}{\varepsilon}, \quad \delta_m = -\frac{3\alpha}{4\pi}\frac{2}{\varepsilon} \quad \text{(DR, MS)} \tag{57}$$

In dimensional regularization, minimal subtraction is almost always upgraded to modified minimal subtraction ($\overline{\text{MS}}$), where the $\ln(4\pi)$ and γ_E factors are also removed. Expanding $\tilde\mu^2$ in Eq. (12)

$$\Sigma_2(\not p) = \frac{\alpha}{2\pi}\int_0^1 dx(x\not p - 2m_R)\left[\frac{2}{\varepsilon} + \ln\frac{4\pi e^{-\gamma_E}\mu^2}{(1-x)(m_R^2 - p^2x)}\right] \tag{58}$$

$$= \frac{\alpha}{2\pi}\left[\frac{1}{2}\not p\left(\frac{2}{\varepsilon} + \ln(4\pi e^{-\gamma_E})\right) - 2m_R\left(\frac{2}{\varepsilon} + \ln(4\pi e^{-\gamma_E})\right) + \text{finite}\right] \tag{59}$$

So in $\overline{\text{MS}}$

$$\delta_2 = -\frac{\alpha}{4\pi}\left(\frac{2}{\varepsilon} + \ln(4\pi e^{-\gamma_E})\right), \quad \delta_m = -\frac{3\alpha}{4\pi}\left(\frac{2}{\varepsilon} + \ln(4\pi e^{-\gamma_E})\right) \quad \text{(DR, } \overline{\text{MS}}\text{)} \tag{60}$$

and then

$$\Sigma_R(\not p) = \frac{\alpha}{2\pi}\int_0^1 dx(x\not p - 2m_R)\left[\ln\frac{\mu^2}{(1-x)(m_R^2 - p^2x)}\right] \tag{61}$$

which is UV finite and has μ in it not $\tilde\mu$. As with Pauli-Villars, there is a 1-parameter family of renormalized 1PI corrections. In both cases, the subtraction point μ is an arbitrary scale which is *not* taken to infinity but will drop out of physical calculations.

The value of m_R is finite in $\overline{\text{MS}}$ and known as the **$\overline{\text{MS}}$ mass**. The renormalized electron propagator will in general not have a pole at $\not p = m_R$. There is still a pole at $\not p = m_P$ with residue i, but $m_P \neq m_R$. Recalling the renormalized electron propagator from Eq. (35)

$$iG^R(\not p) = \frac{i}{\not p - m_R + \Sigma_R(\not p)} \tag{62}$$

we can now easily relate the pole mass and the $\overline{\text{MS}}$ mass. Requiring a pole in this propagator at $\not p = m_P$ gives

$$m_P - m_R + \Sigma_R(m_P) = 0 \tag{63}$$

Using $m_P = m_R$ at leading order, we then have

$$m_R = m_P + \Sigma_R(m_P) = m_P\left[1 - \frac{\alpha}{4\pi}\left(5 + 3\ln\frac{\mu^2}{m_P^2}\right) + \mathcal{O}(\alpha^2)\right] \quad \text{(DR)} \tag{64}$$

In particular, the $\overline{\text{MS}}$ mass depends on the arbitrary scale μ.

While your first instinct might be that this extra parameter μ in minimal subtraction adds an unnecessary complication, it is actually extremely useful. The fact that physical observables are independent of μ gives a powerful constraint. Indeed, demanding $\frac{d}{d\mu}\mathcal{O} = 0$ where $\mathcal{O}$ is some observable is the *renormalization group equation*, to be discussed in Lecture III-9.

6 Summary and Discussion

In this lecture we saw that the electron self-energy graph contributes loop corrections to the electron propagator. This loop was divergent, but the divergence could be removed by renormalizing the electron's quantum field $\psi^0 = \sqrt{Z_2}\psi^R$ and redefining the electron mass $m_0 = Z_m m_R$. In these equations, ψ^0 and m^0 refer to bare quantities which are formally infinite, while ψ^R and m_R are finite renormalized quantities. The quantities δ_m and δ_2 defined by expanding the renormalization factors around the classical values, e.g. $Z_2 = 1 + \delta_2$, are known as *counterterms*. These counterterms can be chosen to cancel the infinite contribution of the electron self-energy graph to the renormalized electron propagator. While the cancellation fixes the infinite parts of the counterterms, the finite parts are arbitrary. Conventions for fixing the finite parts are known as *subtraction schemes*.

We saw that the general geometric series of loops correcting the propagator can be summed to all orders in α leading to a renormalized propagator of the form

$$iG^R(\not{p}) = \frac{i}{\not{p} - m_R + \Sigma_R(\not{p})} \tag{65}$$

Here $\Sigma_R(\not{p})$ represents *one-particle irreducible Feynman diagrams* plus counterterm contributions. Up to order e^2, we found $\Sigma_R(\not{p}) = \Sigma_2(\not{p}) + \delta_2\not{p} - (\delta_m + \delta_2)m_R$. This renormalized propagator should have a pole at the physical electron mass, the **pole mass**, with residue i

$$iG^R(\not{p}) = \frac{i}{\not{p} - m_P} + \text{terms regular at } \not{p} = m_P \tag{66}$$

In terms of the bare propagator, $G^{\text{bare}}(\not{p}) = Z_2 G^R(\not{p})$, we can write

$$iG^{\text{bare}}(\not{p}) = \frac{iZ_2}{\not{p} - m_P} + \text{terms regular at } \not{p} = m_P \tag{67}$$

Sometimes people use this to interpret Z_2 as the residue of the pole. However, since both Z_2 and the bare propagator are formally infinite, this interpretation must be made with care.

Two subtraction schemes were discussed. The first, the *on-shell scheme*, was defined by identifying the location of the pole of the propagator, m_P, with the renormalized mass $m_R \equiv m_P$. This, along with a constraint on the residue of the pole, generated two equations

$$\Sigma_R(m_P) = 0, \quad \frac{d}{d\not{p}}\Sigma_R(\not{p})\bigg|_{\not{p}=m_P} = 0 \tag{68}$$

These equations, which apply to all orders in perturbation theory, fix the counterterms δ_2 and δ_m. They are known as the **on-shell renormalization conditions**.

The second scheme, known as *minimal subtraction*, simply sets the finite parts of δ_2 and δ_m to zero. *Modified minimal subtraction* also subtracts off $\ln(4\pi)$ and γ_E factors which effectively replaces $\tilde{\mu}$ by μ in dimensionally regulated amplitudes. In minimal subtraction, the renormalized mass (written as m_R or often just m) is known as the $\overline{\text{MS}}$ *mass*. It is in general different from the pole mass. At 1-loop, we found

$$m_R = m_P + \Sigma_R(m_P) = m_P\left[1 - \frac{\alpha}{4\pi}\left(5 + 3\ln\frac{\mu^2}{m_P^2}\right)\right] \tag{69}$$

This expression depends on an arbitrary scale μ known as the *subtraction point* which is not taken to ∞. While the extra parameter μ may seem superfluous, we will see in Lecture III-9 that physical observables being independent of μ leads to an important constraint, the renormalization group equations. Even without using the renormalization group, μ-independence order-by-order in perturbation theory gives an important check that an observable has been calculated correctly. We will provide a number of examples in the next two lectures.

You might wonder why on earth anyone would use an unphysical and arbitrary $\overline{\rm MS}$ mass rather than the physical pole mass. The basic answer is that $\overline{\rm MS}$ is a much simpler subtraction scheme than the on-shell scheme. It is often easier to compute loops in $\overline{\rm MS}$ and then convert the masses back to the pole mass at the end rather than to do the computations in terms of the pole mass from the beginning. Numerically, the differences between pole masses and $\overline{\rm MS}$ are often quite small for μ chosen of order m_P. One important exception is the top quark mass, where $m_P \sim 175$ GeV but $m_R \sim 163$ GeV. This 5% difference is important for precision physics, to be discussed in Lecture IV-5. A more sophisticated answer is that the $\overline{\rm MS}$ mass has an appealing property that it is free of ambiguities related to non-perturbative effects in quantum chromodynamics (so-called *renormalon ambiguities*). Indeed for particles like quarks, which can never be seen as asymptotic states, there is not actually a pole in the S-matrix, so the pole mass is not always a useful mass definition.

It is important to keep in mind that the physical electron mass m_P is the location of the pole in the electron propagator whether or not we identified this mass with m_R. In the on-shell scheme, we cannot ask about radiative corrections to the electron mass m_P since by definition, it does not receive any. In minimal subtraction, the electron mass m_R *does* get radiative corrections, as Eq. (69) shows. A physical effect of these radiative corrections can be seen, in principle, in logarithmic corrections to the Yukawa potential, which is easiest to understand using renormalization group methods (cf. Lecture III-10).

It is not always easy to determine what scheme experimental mass measurements correspond to. For example, the top mass has been measured at the Tevatron and the Large Hadron Collider by fitting a line shape to the output of a particular Monte Carlo event generator called Pythia. Thus one can say the top mass is measured in the **Pythia scheme**. Although the Pythia scheme is close to the on-shell scheme, for a precision top mass measurement it is necessary to have a systematic way to convert between the two.

Finally, we discussed that for S-matrix elements, the LSZ reduction theorem should be modified to

$$\langle f|S|i\rangle = (\not{p}_f - m_P)\cdots(\not{p}_i - m_P)\langle\psi^R\cdots\psi^R\rangle_{\rm amputated} \tag{70}$$

where *amputated* refers to not including diagrams with 1PI corrections to external legs. This was necessary because those corrections are already included in what we call external states, with poles at m_P.

Despite the amputation of corrections to external legs, there are physical implications of the electron self-energy when the graph corrects internal lines. Historically, the most important such correction was the Lamb shift (the splitting between the $2S_{1/2}$ and $2P_{1/2}$ levels of the Hydrogen atom). Radiative corrections to the electron propagator were what Oppenheimer was missing when he calculated this shift in Old-Fashioned Perturbation Theory in 1932. Hans Bethe's famous estimate

$$\Delta E(2S_{1/2}) = m\frac{4Z^4\alpha^5}{3\pi n^3}\ln\frac{m}{E_0} \sim 1000\,{\rm MHz} \tag{71}$$

for the Lamb shift from 1947 came from cutting off the infrared divergence in the self-energy graph at the energy E_0 of the Hydrogen atom ground state. The self-energy graph was one of the ingredients Feynman, Schwinger and Tomonaga needed to produce the correct, infrared- and ultraviolet-finite Lamb shift in 1949. More generally, the self-energy graph contributes in some way to almost every precision process that has been calculated in QED.