Lectures on effective field theory

David B. Kaplan

February 24, 2016

Abstract

Five lectures delivered at the ICTP-SAFIR, Sao Paulo, Brasil, February 22-26, 2016.

Contents

1 Effective Quantum Mechanics

1.1 What is an effective field theory?

The uncertainty principle tells us that to probe the physics of short distances we need high momentum. On the one hand this is annoying, since creating high relative momentum in a lab costs a lot of money! On the other hand, it means that we can have predictive theories of particle physics at low energy without having to know everything about physics at short distances. For example, we can discuss precision radiative corrections in the weak interactions without having a grand unified theory or a quantum theory of gravity. The price we pay is that we have a number of parameters in the theory (such as the Higgs and fermion masses and the gauge couplings) which we cannot predict but must simply measure. But this is a lot simpler to deal with than a mess like turbulent fluid flow where the physics at many different distance scales are all entrained together.

The basic idea behind effective field theory (EFT) is the observation that the non-analytic parts of scattering amplitudes are due to intermediate process where physical particles can exist on shell (that is, kinematics are such that internal propagators $1/(p^2 - m^2 + i\epsilon)$ in Feynman diagrams can diverge with $p^2 = m^2$ so that one is sensitive to the $i\epsilon$ and sees cuts in the amplitude due to logarithms, square roots, etc). Therefore if one can construct a quantum field theory that correctly accounts for these light particles, then all the contributions to the amplitude from virtual heavy particles that cannot be physically created at these energies can be Taylor expanded p^2/M^2, where M is the energy of the heavy particle. (By "heavy" I really mean a particle whose energy is too high to create; this might be a heavy particle at rest, but it equally well applies to a pair of light particles with high relative momentum.) However, the power of of this observation is not that one can Taylor expand parts of the scattering amplitude, but that the Taylor expanded amplitude can be computed directly from a quantum field theory (the EFT) which contains *only* light particles, with local interactions between them that encode the small effects arising from virtual heavy particle exchange. Thus the standard model does not contain X gauge bosons from the GUT scale, for example, but can be easily modified to account for the very small effects such particles could have, such as causing the proton to decay, for example.

So in fact, all of our quantum field theories are EFTs; only if there is some day a Theory Of Everything (don't hold your breath) will we be able to get beyond them. So how is a set of lectures on EFT different than a quick course on quantum field theory? Traditionally a quantum field theory course is taught from the point of view that held sway from when it was originated in the late 1920s through the development of nonabelian gauge theories in the early 1970s: one starts with a ϕ^4 theory at tree level, and then computes loops and encounters renormalization; one then introduces Dirac fermions and Yukawa interactions, moves on to QED, and then nonabelian gauge theories. All of these theories have operators of dimension four or less, and it is taught that this is necessary for renormalizability. Discussion of the Fermi theory of weak interactions, with its dimension six four-fermion operator, is relegated to particle physics class. In contrast, EFT incorporates the ideas of Wilson and others that were developed in the early 1970s and completely turned on its head how we think about UV (high energy) physics and renormalization, and how we interpret the results of calculations. While the ϕ^4 interaction used to be considered one only few well-defined (renormalizable) field theories and the Fermi theory of weak interactions was viewed as useful but nonrenormalizable and sick, now the scalar theory is considered sick, while the Fermi theory is a simple prototype of all successful quantum field theories. The new view brings with it its own set of problems, such as an obsession with the fact the our universe appears to be fine-tuned. Is the modern view the last word? Probably not, and I will mention unresolved mysteries at the end of my lectures.

There are three basic uses for effective field theory I will touch on in these lectures:

- Top-down: you know the theory to high energies, but either you do not need all of its complications to arrive at the desired description of low energy physics, or else the full theory is nonperturbative and you cannot compute in it, so you construct an EFT for the light degrees of freedom, constraining their interactions from your knowledge of the symmetries of the more complete theory;
- Bottom-up: you explore small effects from high dimension operators in your low energy EFT to gain cause about what might be going on at shorter distances than you can directly probe;
- Philosophizing: you marvel at how "fine-tuned" our world appears to be, and pondering whether the way our world appears is due to some missing physics, or because we live in a special corner of the universe (the anthropic principle), or whether we live at a dynamical fixed point resulting from cosmic evolution. Such investigations are at the same time both fascinating — and possibly an incredible waste of time!

To begin with I will not discuss effective field theories, however, but effective quantum mechanics. The essential issues of approximating short range interactions with point-like interactions have nothing to do with relativity or many-body physics, and can be seen in entirety in non-relativistic quantum mechanics. I thought I would try this introduction because I feel that the way quantum mechanics and quantum field theory are traditionally taught it looks like they share nothing in common except for mysterious ladder operators, which is of course not true. What this will consist of is a discussion of scattering from delta-function potentials in different dimensions.

1.2 Scattering in 1D

1.2.1 Square well scattering in 1D

We have all solved the problem of scattering in 1D quantum mechanics, from both square barrier potentials and delta-function potentials. Consider scattering of a particle of mass m from an attractive square well potential of width Δ and depth $\frac{\alpha^2}{2m\Delta^2}$,

$$V(x) = \begin{cases} -\frac{\alpha^2}{2m\Delta^2} & 0 \le x \le \Delta \\ 0 & \text{otherwise} \end{cases} . \tag{1}$$

Here α is a dimensionless number that sets the strength of the potential. It is straight forward to compute the reflection and transmission coefficients at energy E (with $\hbar = 1$)

$$R = (1 - T) = \left[\frac{4\kappa^2 k^2 \csc^2(\kappa\Delta)}{\left(k^2 - \kappa^2\right)^2} + 1\right]^{-1} , \tag{2}$$

where

$$k = \sqrt{2mE} \ , \qquad \kappa = \sqrt{k^2 + \frac{\alpha^2}{\Delta^2}} \ . \tag{3}$$

For low k we can expand the reflection coefficient and find

$$R = 1 - \frac{4}{\alpha^2 \sin^2 \alpha} \Delta^2 k^2 + O(\Delta^4 k^4) \tag{4}$$

Note that $R \to 1$ as $k \to 0$, meaning that the potential has a huge effect at low enough energy, no matter how weak...we can say the interaction is very *relevant* at low energy.

1.2.2 Relevant δ-function scattering in 1D

Now consider scattering off a δ-function potential in 1D,

$$V(x) = -\frac{g}{2m\Delta} \delta(x) \ , \tag{5}$$

where the length scale Δ was included in order to make the coupling g dimensionless. Again one can compute the reflection coefficient and find

$$R = (1 - T) = \left[1 + \frac{4k^2\Delta^2}{g^2} \right]^{-1} = 1 - \frac{4k^2\Delta^2}{g^2} + O(k^4) \ . \tag{6}$$

By comparing the above expression to eq. (4) we see that at low momentum the δ function gives the same reflection coefficient to up to $O(k^4)$ as the square well, provided we set

$$g = \alpha \sin \alpha \ . \tag{7}$$

In the EFT business, the above equation is called a "matching condition"; this matching condition is shown in Fig. 1, and interpreting the structure in this figure – in particular the sign changes for g – is one of the problems at the end of the lecture. For small α the matching condition is simply $g \simeq \alpha^2$.

1.3 Scattering in 3D

Now let's see what happens if we try the same thing in 3D (three spatial dimensions), choosing the strength of a δ-function potential to mimic low energy scattering off a square well potential. Why this fixation with δ-function potentials? They are not particularly special in non-relativistic quantum mechanics, but in a relativistic field theory they are the *only* instantaneous potential which can be Lorentz invariant. That is why we always formulate quantum field theories as interactions between particles only when they are at the same point in spacetime. All the issues of renormalization in QFT arise from the singular nature of these δ-function interactions. So I am focussing on δ-function potentials in quantum mechanics in order to illustrate what is going on in the relativistic QFT.

First, a quick review of a few essentials of scattering theory in 3D, focussing only on s-wave scattering.

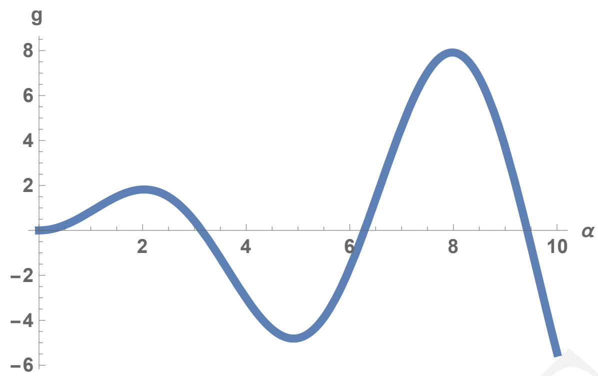

Figure 1: *The matching condition in 1D: the appropriate value of g in the effective theory for a given α in the full theory.*

A scattering solution for a particle of mass m in a finite range potential must have the asymptotic form for large $|\mathbf{r}|$

$$\psi \xrightarrow{r\to\infty} e^{ikz} + \frac{f(\theta)}{r} e^{ikr} \ . \tag{8}$$

representing an incoming plane wave in the z direction, and an outgoing scattered spherical wave. The quantity f is the scattering amplitude, depending both on scattering angle θ and incoming momentum k, and $|f|^2$ encodes the probability for scattering; in particular, the differential cross section is simply

$$\frac{d\sigma}{d\theta} = |f(\theta)|^2 \ . \tag{9}$$

For scattering off a spherically symmetric potential, both $f(\theta)$ and $e^{ikz} = e^{ikr\cos\theta}$ can be expanded in Legendre polynomials ("partial wave expansion"); I will only be interested in s-wave scattering (angle independent) and therefore will replace $f(\theta)$ simply by f — independent of angle, but still a function of k. For the plane wave we can average over θ when only considering s-wave scattering, replacing

$$e^{ikz} = e^{ikr\cos\theta} \xrightarrow[s-\text{wave}]{} \frac{1}{2}\int_0^\pi d\theta\, \sin\theta\, e^{ikr\cos\theta} = j_0(kr) \ . \tag{10}$$

Here $j_0(z)$ is a regular spherical Bessel function; we will also meet its irregular partner $n_0(z)$, where where

$$j_0(z) = \frac{\sin z}{z} \ , \qquad n_0(z) = -\frac{\cos z}{z} \ . \tag{11}$$

These functions are the s-wave solutions to the free 3D Schrödinger equation $z = kr$.

So we are interested in a solution to the Schrödinger equation with asymptotic behavior

$$\psi \xrightarrow[s-\text{wave}]{r\to\infty} j_0(kr) + \frac{f}{r}e^{ikr} = j_0(kr) + kf\left(ij_0(kr) - n_0(kr)\right) \qquad (s\text{-wave}) \tag{12}$$

Since outside the the range of the potential ψ is an exact s-wave solution to the free Schrödinger equation, and the most general solutions to the free radial Schrödinger equation are the spherical Bessel functions $j_0(kr)$, $n_0(kr)$, the asymptotic form for ψ can also be written as

$$\psi \xrightarrow{r\to\infty} A\left(\cos\delta\, j_0(kr) - \sin\delta\, n_0(kr)\right) \ . \tag{13}$$

where A and δ are real constants. The angle δ is called the phase shift, and if there was no potential, boundary conditions at $r = 0$ would require $\delta = 0$...so nonzero δ is indicative of scattering. Relating these two expressions eq. (12) and eq. (13) we find

$$f = \frac{1}{k\cot\delta - ik} \ . \tag{14}$$

So solving for the phase shift δ is equivalent to solving for the scattering amplitude f, using the formula above.

The quantity $k\cot\delta$ is interesting, since one can show that for a finite range potential it must be analytic in the energy, and so has a Taylor expansion in k involving only even powers of k , called "the effective range expansion":

$$k\cot\delta = -\frac{1}{a} + \frac{1}{2}r_0 k^2 + O(k^4) \ . \tag{15}$$

The parameters have names: a is the scattering length and r_0 is the effective range; these terms dominate low energy (low k) scattering. Proving the existence of the effective range expansion is somewhat involved and I refer you to a quantum mechanics text; there is a low-brow proof due to Bethe and a high-brow one due to Schwinger.

And the last part of this lightning review of scattering: if we have two particles of mass M scattering off each other it is often convenient to use Feynman diagrams to describe the scattering amplitude; I denote the Feynman amplitude – the sum of all diagrams – as $i\mathcal{A}$. The relation between $\mathcal{A}$ and f is

$$\mathcal{A} = \frac{4\pi}{M} f \ , \tag{16}$$

where f is the scattering amplitude for a single particle of reduced mass $m = M/2$ in the inter-particle potential. This proportionality is another result that can be priced together from quantum mechanics books, which I won't derive.

1.3.1 Square well scattering in 3D

We consider s-wave scattering off an attractive well in 3D,

$$V = \begin{cases} -\frac{\alpha^2}{m\Delta^2} & r < \Delta \\ 0 & r > \Delta \ . \end{cases} \tag{17}$$

We have for the wave functions for the two regions $r < \Delta$, $r > \Delta$ are expressed in terms of spherical Bessel functions as

$$\psi_<(r) = j_0(\kappa r) \ , \qquad \psi_>(r) = A\left[\cos\delta\, j_0(kr) - \sin\delta\, n_0(kr)\right] \tag{18}$$

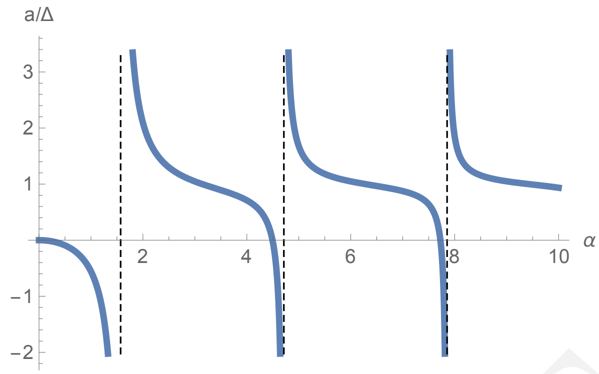

Figure 2: *a/Δ vs. the 3D potential well depth parameter α, from eq. (21).*

where $\kappa = \sqrt{k^2 + \alpha^2/\Delta^2}$ as in eq. (3) and δ is the s-wave phase shift. Equating ψ and ψ' at the edge of the potential at $r = \Delta$ allows us to solve for δ in terms of k, α, Δ, with the result

$$k \cot \delta = \frac{k(k \sin \kappa\Delta + \kappa \cot k\Delta \cos \kappa\Delta)}{k \cot k\Delta \sin \kappa\Delta - \kappa \cos \kappa\Delta} \ . \tag{19}$$

With a little help from Mathematica we can expand this in powers of k^2 and find

$$k \cot \delta = \frac{1}{\Delta}\left(\frac{\tan \alpha}{\alpha} - 1\right)^{-1} + O(k^2) \tag{20}$$

where on comparing with eq. (15) we can read off the scattering length from the k^2 expansion,

$$a = -\Delta\left(\frac{\tan \alpha}{\alpha} - 1\right) \ , \tag{21}$$

a relation shown in Fig. 2. The singularities one finds for the scattering length as the strength of the potential α increases correspond to the critical values $\alpha_c = (2n+1)\pi/2$, $n = 0, 1, 2 \ldots$ where a new bound state appears.

1.3.2 Irrelevant δ-function scattering in 3D

Now we look at reproducing the above scattering length from scattering in 3D off a delta function potential. At first look this seems hopeless: note that the result for a square well of width Δ and coupling $\alpha = O(1)$ gives a scattering length that is $a = O(\Delta)$; this is to be expected since a is a length, and Δ is the natural length scale in the problem. Therefore if you extrapolate to a potential of zero width (a δ function) you would conclude that the scattering length would go to zero, and the scattering amplitude would vanish for low k. This is an example of an *irrelevant* interaction.

On second look the situation is even worse: since $-\delta^3(\mathbf{r})$ scales as $-1/r^3$ while the kinetic $-\nabla^2$ term in the Schrödinger equation only scales as $1/r^2$ you can see that the system does not have a finite energy ground state. For example if you performed a variational calculation, you could lower the energy without bound by scaling the wave function to smaller and smaller extent. Therefore the definition of a δ-function has to be modified in 3D – this is the essence of renormalization.

These two features go hand in hand: typically singular interactions are "irrelevant" and at the same time require renormalization. We can sometimes turn an irrelevant interaction into a relevant one by fixing a certain renormalization condition which forces a fine tuning of the coupling to a critical value, and that is the case here. For example, consider defining the δ-function as the $\rho \to 0$ limit of a square well of width ρ and depth $V_0 = \bar{\alpha}^2(\rho)/(m\rho^2)$, while adjusting the coupling strength $\bar{\alpha}(\rho)$ to keep the scattering length fixed to the desired value of a given in eq. (21). We find

$$a = \rho\left(1 - \frac{\tan\bar{\alpha}(\rho)}{\bar{\alpha}(\rho)}\right) \tag{22}$$

as $\rho \to 0$. There are an infinite number of solutions, corresponding to $\alpha \simeq \alpha_c = (2n+1)\pi/2$, $n = 0, 1, 2\ldots$, and the $n = 0$ possibility is

$$\bar{\alpha}(\rho) \xrightarrow{\rho\to 0} \frac{\pi}{2} + \frac{2\rho}{\pi a} + O(\rho^2)\ . \tag{23}$$

in other words, we have to tune this vanishingly thin square well to have a single bound state right near threshold ($\alpha \simeq \pi/2$). However, note that while naively you might think a potential $-g\delta^3(\mathbf{r})$ would be approximated by a square well of depth $V_0 \propto 1/\rho^3$ as $\rho \to 0$, but we see that instead we get $V_0 \propto 1/\rho^2$. This is sort of like using a potential $-r\delta^3(\mathbf{r})$ instead of $-\delta^3(\mathbf{r})$.

We have struck a delicate balance: A naive δ function potential is too strong and singular to have a ground state; a typical square well of depth $\alpha^2/m\rho^2$ becomes irrelevant for fixed α in the $\rho \to 0$ limit; but a strongly coupled potential of form $V \simeq -(\pi/2)^2/m\rho^2$ can lead to a relevant interaction so long as we tune α its critical value $\alpha_c = \pi/2$ in precisely the right way as we take $\rho \to 0$.

This may all seem more familiar to you if I to use field theory methods and renormalization. Consider two colliding particles of mass M in three spatial dimensions with a δ-function interaction; this is identical to the problem of potential scattering when we identify m with the reduced mass of the two particle system,

$$m = \frac{M}{2}\ . \tag{24}$$

We introduce the field ψ for the scattering particles (assuming they are spinless bosons) and the Lagrange density

$$\mathcal{L} = \psi^\dagger\left(i\partial_t + \frac{\nabla^2}{2M}\right)\psi - \frac{C_0}{4}\left(\psi^\dagger\psi\right)^2\ . \tag{25}$$

Here $C_0 > 0$ implies a repulsive interaction. As in a relativistic field theory, ψ annihilates particles and $\psi^\dagger$ creates them; unlike in a relativistic field theory, however, there are no anti-particles.

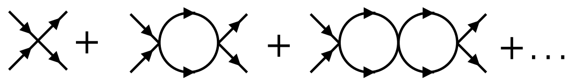

Figure 3: *The sum of Feynman diagrams giving the exact scattering amplitude for two particles interaction via a δ-function potential.*

The kinetic term gives rise to the free propagator

$$G(E, \mathbf{p}) = \frac{i}{E - \mathbf{p}^2/(2M) + i\epsilon} \ , \tag{26}$$

while the interaction term gives the vertex $-iC_0$. The total Feynman amplitude for two particles then is the sum of diagrams in Fig. 3, which is the geometric series

$$i\mathcal{A} = -iC_0 \left[1 + (C_0 B(E)) + (C_0 B(E))^2 + \ldots\right] = \frac{i}{-\frac{1}{C_0} + B(E)} \ , \tag{27}$$

where B is the 1-loop diagram, which in the center of momentum frame (where the incoming particles have momenta $\pm\mathbf{p}$ and energy $E/2 = \mathbf{k}^2/(2M)$) is given by

$$B(E) = -i \int \frac{d^4q}{(2\pi)^4} \frac{i}{\left(\frac{E}{2} + q_0 - \frac{\mathbf{q}^2}{2M} + i\epsilon\right)} \frac{i}{\left(\frac{E}{2} - q_0 - \frac{\mathbf{q}^2}{2M} + i\epsilon\right)} = \int \frac{d^3q}{(2\pi)^3} \frac{1}{E - \frac{\mathbf{q}^2}{M} + i\epsilon} \ . \tag{28}$$

The B integral is linearly divergent and so I will regulate it with a momentum cutoff and renormalize the coupling C_0:

$$\begin{aligned} B(E, \Lambda) &= \int^{\Lambda} \frac{d^3q}{(2\pi)^3} \frac{1}{E - \frac{\mathbf{q}^2}{M} + i\epsilon} \\ &= -\frac{M\left(\Lambda - \sqrt{-EM - i\epsilon}\tan^{-1}\left(\frac{\Lambda}{\sqrt{-EM - i\epsilon}}\right)\right)}{2\pi^2} \\ &= -\frac{M\Lambda}{2\pi^2} + \frac{M}{4\pi}\sqrt{-ME - i\epsilon} + O\left(\frac{1}{\Lambda}\right) \\ &= -\frac{M\Lambda}{2\pi^2} - i\frac{Mk}{4\pi} + O\left(\frac{1}{\Lambda}\right) \ . \end{aligned} \tag{29}$$

Thus from eq. (27) we get the Feynman amplitude

$$\mathcal{A} = \frac{1}{-\frac{1}{C_0} + B(E)} = \frac{1}{-\frac{1}{C_0} - \frac{M\Lambda}{2\pi^2} - i\frac{Mk}{4\pi}} = \frac{4\pi}{M} \frac{1}{\left(-\frac{4\pi}{M\overline{C}_0} - ik\right)} \tag{30}$$

where

$$\frac{1}{\overline{C}_0} = \frac{1}{C_0} - \frac{M\Lambda}{2\pi^2} \ . \tag{31}$$

$\overline{C}_0$ is our renormalized coupling, C_0 is our bare coupling, and Λ is our UV cutoff. Since in in 3D we have (eq. (14), eq. (16))

$$\mathcal{A} = \frac{4\pi}{M} \frac{1}{k\cot\delta - ik} \ , \qquad k\cot\delta = -\frac{1}{a} + \frac{1}{2} r_0 k^2 + \ldots \tag{32}$$

we see that this theory relates $\overline{C}_0$ to the scattering length as

$$\overline{C}_0 = \frac{4\pi a}{M} \ . \tag{33}$$

Therefore we can reproduce square well scattering length eq. (21) by taking

$$\overline{C}_0 = -\frac{4\pi\Delta}{M}\left(\frac{\tan\alpha}{\alpha} - 1\right) \ . \tag{34}$$

With our simple EFT we can reproduce the scattering length of the square well problem, but not the next term in the low k expansion, the effective range. There is a one-to-one correspondence between the number of terms we can fit in the effective range expansion and the number of operators we include in the EFT; to account for the effective range we would have to include a new contact interaction involving two derivatives and match its coefficient.

What have we accomplished? We have shown that one can reproduce low energy scattering from a finite range potential in 3D with a δ-function interaction, with errors of $O(k^2\Delta^2)$ with the caveat that renormalization is necessary if we want to make sense of the theory.

However there is second important and subtle lesson: We can view eq. (31) plus eq. (33) to imply a fine tuning of the inverse bare coupling $1/C_0$ coupling as $\Lambda \to \infty$: $M\Lambda C_0/(2\pi^2)$ must be tuned to $1 + O(1/a\Lambda)$ as $\Lambda \to \infty$. This is the same lesson we learned looking at square wells: if C_0 didn't vanish at least linearly with the cutoff, the interaction would be too strong to makes sense; while if ΛC_0 went to zero or a small constant, the interaction would be irrelevant. Only if ΛC_0 is fine-tuned to a critical value can we obtain nontrivial scattering at low k.

1.4 Scattering in 2D

1.4.1 Square well scattering in 2D

Finally, let's look at the intermediary case of scattering in two spatial dimensions, where we take the same potential as in eq. (17). This is not just a tour of special functions — something interesting happens! The analogue of eq. (35) for the two dimensional square well problem is

$$\psi_<(r) = J_0(\kappa r) \ , \qquad \psi_>(r) = A\left[\cos\delta\, J_0(kr) - \sin\delta\, Y_0(kr)\right] \tag{35}$$

where κ is given in eq. (3) and J, Y are the regular and irregular Bessel functions. Equating ψ and ψ' at the boundary $r = \Delta$ gives[1]

$$\begin{aligned}\cot\delta &= \frac{kJ_0(\Delta\kappa)Y_1(\Delta k) - \kappa J_1(\Delta\kappa)Y_0(\Delta k)}{kJ_0(\Delta\kappa)J_1(\Delta k) - \kappa J_1(\Delta\kappa)J_0(k\Delta)} \\ &= \frac{2\left(\frac{J_0(\alpha)}{\alpha J_1(\alpha)} + \log\left(\frac{\Delta k}{2}\right) + \gamma_E\right)}{\pi} + O\left(k^2\right)\end{aligned} \tag{36}$$

[1]In the following expressions $\gamma_E = 0.577\ldots$ is the Euler constant.

This result looks very odd because of the logarithm that depends on k! The interesting feature of this expression is not that $\cot\delta(k) \to -\infty$ for $k \to 0$: that just means that the phase shift vanishes at low k. What is curious is that for our attractive potential, the function $J_0(\alpha)/(\alpha J_1(\alpha))$ is strictly positive, and therefore $\cot\delta$ changes sign at a special value for k,

$$k = \Lambda \simeq \frac{2e^{-\frac{J_0(\alpha)}{\alpha J_1(\alpha)} - \gamma_E}}{\Delta} \ , \tag{37}$$

where the scale Λ is *exponentially* lower than our fundamental scale Δ for weak coupling, since then $J_0(\alpha)/\alpha J_1(\alpha) \sim 2/\alpha^2 \gg 1$. This is evidence in the scattering amplitude for a bound state of size $\sim 1/\Lambda$...exponentially larger than the size of the potential!

On the other hand, if the interaction is repulsive, the $J_0(\alpha)/\alpha J_1(\alpha)$ factor is replaced by $-I_0(\alpha)/\alpha I_1(\alpha) < 0$, I_n being one of the other Bessel functions, and the numerator in eq. (36) is always negative, and there is no bound state.

1.4.2 Marginal δ-function scattering in 2D & asymptotic freedom

If we now look at the Schrödinger equation with a δ-function to mock up the effects of the square well for low k we find something funny: the equation is scale invariant. What that means is that the existence of any solution $\psi(r)$ to the equation

$$\left[-\frac{1}{2m}\nabla^2 + \frac{g}{m}\delta^2(\mathbf{r})\right]\psi(r) = E\psi(r) \tag{38}$$

implies a continuous family of solutions $\psi_\lambda(r) = \psi(\lambda r)$ – the same functional form except scaled smaller by a factor of λ – with energy $E_\lambda = \lambda^2 E$. Thus it seems that all possible energy eigenvalues with the same sign as E exist and there are no discrete eigenstates...which is OK if only positive energy scattering solutions exist, the case for a repulsive interaction — but not OK if there are bound states: it appears that if there is any one negative energy state, then there is an unbounded continuum of negative energy states and no ground state. The problem is that ∇^2 and $\delta^2(\mathbf{r})$ have the same dimension, 1/length^2, and so there is no inherent scale to the left side of the equation. Since the scaling property of $\delta^D(\mathbf{r})$ changes with dimension D, while the scaling property of ∇^2 does not, $D = 2$ is special.

Since the δ-function interaction seems to be scale invariant, we say that it is neither relevant (dominating IR physics, as in 1D) nor irrelevant (unimportant to IR physics, as in 3D) but apparently of equal important at all scales, which we call *marginal*. However, we know that (i) the δ function description appears to be sick, and (ii) from our exact analysis of the square well that the IR description of the full theory is not really scale invariant, due to the logarithm. Therefore it is a reasonable guess that our analysis of the δ-function is incorrect due to its singularity, and that we are going to have to be more careful, and renormalize.

We can repeat the Feynman diagram approach we used in 3D, only now in 2D. Now the loop integral in eq. (29) is required in $d = 3$ spacetime dimensions instead of $d = 4$. It is still divergent, but now only log divergent, not linearly divergent. It still needs regularization, but this time instead of using a momentum cutoff I will use dimensional regularization, to make it look even more like conventional QFT calculations. Therefore we keep the number of spacetime dimensions d arbitrary in computing the integral, and subsequently expand about $d = 3$ (for scattering in $D = 2$ spatial dimensions)[2]. We take for our action

$$S = \int dt \int d^{d-1}x \left[\psi^\dagger \left(i\partial_t + \frac{\nabla^2}{2M} \right) \psi - \mu^{d-3} \frac{C_0}{4} \left(\psi^\dagger \psi \right)^2 \right] . \tag{39}$$

where the renormalization scale μ was introduced to keep C_0 dimensionless (see problem). Then the Feynman rules are the same as in the previous case, except for the factor of μ^{d-3} at the vertices, and we find

$$\begin{aligned} B(E) &= \mu^{3-d} \int \frac{d^{d-1}q}{(2\pi)^{d-1}} \frac{1}{E - \frac{\mathbf{q}^2}{M} + i\epsilon} \\ &= -M \left(-ME - i\epsilon\right)^{\frac{d-3}{2}} \Gamma\left(\frac{3-d}{2}\right) \frac{\mu^{3-d}}{(4\pi)^{(d-1)/2}} \\ &\xrightarrow{d\to 3} \frac{M}{2\pi} \frac{1}{(d-3)} + \frac{M}{4\pi} \left(\gamma_E - \ln 4\pi + \ln \frac{k^2}{\mu^2} - i\pi \right) + O(d-3) \end{aligned} \tag{40}$$

where $k = \sqrt{ME}$ is the magnitude of the momentum of each incoming particle in the center of momentum frame, and the scattering amplitude is therefore

$$\mathcal{A} = \frac{1}{-\frac{1}{C_0} + B(E)} = \left[-\frac{1}{C_0} + \frac{M}{2\pi} \frac{1}{(d-3)} + \frac{M}{4\pi} \left(\gamma_E - \ln 4\pi + \ln \frac{k^2}{\mu^2} - i\pi \right) \right]^{-1} \tag{41}$$

At this point it is convenient to define the dimensionless coupling constant g:

$$C_0 \equiv g \frac{4\pi}{M} . \tag{42}$$

Given the definition of our Lagrangian, $g > 0$ corresponds to a repulsive potential, and $g < 0$ is attractive. so that the amplitude is

$$\mathcal{A} = \frac{4\pi}{M} \left[-\frac{1}{g} - \frac{2}{(d-3)} + \gamma_E - \ln 4\pi + \ln \frac{k^2}{\mu^2} - i\pi \right]^{-1} \tag{43}$$

To make sense of this at $d = 3$ we have to renormalize g with the definition:

$$\frac{1}{g} = \frac{1}{\bar{g}(\mu)} + \frac{2}{(d-3)} + \gamma_E - \ln 4\pi , \tag{44}$$

where $\bar{g}(\mu)$ is the renormalized running coupling constant, and so the amplitude is given by

$$\mathcal{A} = \frac{4\pi}{M} \left[-\frac{1}{\bar{g}(\mu)} + \ln \frac{k^2}{\mu^2} - i\pi \right]^{-1} \tag{45}$$

[2] If you are curious why I did not use dimensional regularization for the $D = 3$ case: dim reg ignores power divergences, and so when computing graphs with power law divergences using dim reg you do not explicitly notice that you are fine-tuning the theory. This happens in the standard model with the quadratic divergence of the Higgs mass[2]...every few years someone publishes a preprint saying there is no fine-tuning problem since one can compute diagrams using dim reg, where there is no quadratic divergence, which is silly. I used a momentum cutoff in the previous section so we could see the fine-tuning of C_0.

Since this must be independent of μ it follows that

$$\mu\frac{d}{d\mu}\left(-\frac{1}{\bar{g}(\mu)}+\ln\frac{k^2}{\mu^2}\right)=0 \tag{46}$$

or equivalently,

$$\mu\frac{d\bar{g}(\mu)}{d\mu}=\beta(\bar{g})\ ,\qquad \beta(\bar{g})=2\bar{g}(\mu)^2\ . \tag{47}$$

If we specify the renormalization condition $\bar{g}(\mu_0)\equiv\bar{g}_0$, then the solution to this renormalization group equation is

$$\bar{g}(\mu)=\frac{1}{\frac{1}{\bar{g}_0}+2\ln\frac{\mu_0}{\mu}}\ . \tag{48}$$

Note that this solution $\bar{g}(\mu)$ blows up at

$$\mu=\mu_0 e^{1/(2\bar{g}_0)}\equiv\Lambda\ . \tag{49}$$

For $0<g_0\ll 1$ (weak repulsive interaction) we have $\Lambda\gg\mu_0$, while for $-1\ll g_0<0$ (weak attractive interaction) Λ is an infrared scale, $\Lambda\ll\mu_0$. If we set $\mu_0=\Lambda$ in eq. (48) and $g_0=\infty$, we find

$$\bar{g}(\mu)=\frac{1}{\ln\frac{\Lambda^2}{\mu^2}}\ , \tag{50}$$

and the amplitude as

$$\mathcal{A}=\frac{4\pi}{M}\frac{1}{\ln\frac{k^2}{\Lambda^2}+i\pi}\ , \tag{51}$$

or equivalently,

$$\cot\delta=-\frac{1}{\pi}\ln\frac{k^2}{\Lambda^2}\ . \tag{52}$$

Now just have to specify Λ instead of $\bar{g}_0$ to define the theory ("dimensional transmutation").

Finally, we can match this δ-function scattering amplitude to the square well scattering amplitude at low k by equating eq. (52) with our expression eq. (36), yielding the matching condition

$$\ln\frac{k^2}{\Lambda^2}=2\left(\frac{J_0(\alpha)}{\alpha J_1(\alpha)}+\log\left(\frac{\Delta k}{2}\right)+\gamma_E\right) \tag{53}$$

from which the k dependence drops out and we arrive at and expression for Λ in terms of the coupling constant α of the square well:

$$\Lambda=\frac{2e^{-\frac{J_0(\alpha)}{\alpha J_1(\alpha)}-\gamma}}{\Delta} \tag{54}$$

If $g_0 < 0$ (attractive interaction) the scale Λ is in the IR ($\mu \ll \mu_0$ if g_0 is moderately small) and we say that the interaction is asymptotically free, with Λ playing the same role as Λ_{QCD} in the Standard Model – except that here we are not using perturbation theory, the β-function is exact, and we can take $\mu < \Lambda$ and watch $\bar{g}(\mu)$ change from $+\infty$ to $-\infty$ as we scale through a bound state. If instead $g_0 > 0$ (repulsive interaction) then Λ is in the UV, we say the theory is asymptotically unfree, and Λ is similar to the Landau pole in QED. So we see that while the Schrödinger equation *appeared* to have a scale invariance and therefore no discrete states, in reality when one makes sense of the singular interaction, a scale Λ seeks into the theory, and it is no longer scale invariant.

1.5 Lessons learned

We have learned the following by studying scattering from a finite range potential at low k in various dimensions:

- A contact interaction (δ-function) is more irrelevant in higher dimensions;
- marginal interactions are characterized naive scale invariance, and by logarithms of the energy and running couplings when renormalization is accounted for; they can either look like relevant or irrelevant interactions depending on whether the running is asymptotically free or not; and in either case they are characterized by a mass scale Λ exponentially far away from the fundamental length scale of the interaction, Δ.
- Irrelevant interactions and marginal interactions typically require renormalization; an irrelevant interaction can sometimes be made relevant if its coefficient is tuned to a critical value.

All of these lessons will be pertinent in relativistic quantum field theory as well.

1.6 Problems for lecture I

I.1) Explain Fig. 1: how do you interpret those oscillations? Similarly, what about the cycles in Fig. 2?

I.2) Consider dimensional analysis for the non-relativistic action eq. (39). Take momenta p to have dimension 1 by definition in any spacetime dimension d; with the uncertainty principle $[x, p] = i\hbar$ and $\hbar = 1$ we then must assign dimension -1 to spatial coordinate x. Write this as

$$[p] = [\partial_x] = 1 \ , \qquad [x] = -1 \ . \tag{55}$$

Unlike in the relativistic theory we can treat M as a dimensionless parameter under this scaling law. If we do that, use eq. (39) with the factor of μ omitted to figure out the scaling dimensions

$$[t] \ , \quad [\partial_t] \ , \quad [\psi] \ , \quad [C_0] \tag{56}$$

for arbitrary d, using that fact that the action S must be dimensionless (after all, in a path integral we exponentiate $S/\hbar$, which would make no sense if that was a dimensional quantity). What is special about $[C_0]$ at $d = 3$? Confirm that including the factor of μ^{d-3}, where μ has scaling dimension 1 ($[\mu] = 1$) allows C_0 to maintain its $d = 3$ scaling dimension for any d.

I.3) In eq. (52) the distinction between attractive and repulsive interactions seems to have been completely lost since that equation holds for both cases! By looking at how the 2D matching works in describing the square well by a δ-function, explain how the low energy theory described by eq. (52) behaves differently when the square well scattering is attractive versus repulsive. Is there physical significance to the scale Λ in the effective theory for an attractive interaction? What about for a repulsive interaction?

2 EFT at tree level

To construct a relativistic effective field theory valid up to some scale Λ, we will take for our action made out of all light fields (those corresponding to particles with masses or energies much less that Λ) including all possible local operators consistent with the underlying symmetries that we think govern the world. All UV physics that we are not including explicitly is encoded in the coefficients of these operators, in the same way we saw in the previous section that a contact interaction (δ-function potential) was able to reproduce the scattering length for scattering off a square well if its coefficient was chosen appropriately (we "matched" it to the UV physics). However, in the previous examples we just tried matching the scattering lengths; we could have tried to also reproduce $O(k^2\Delta^2)$ effects, and so on, but to do so would have required introducing more and more singular contributions to the potential in the effective theory, such as $\nabla^2\delta(\mathbf{r})$, $\nabla^4\delta(\mathbf{r})$, and so on. Going to all orders in k^2 would require an infinite number of such terms, and the same is true for a relativistic EFT. Such a theory is not "renormalizable" in the historical sense: there is typically no finite set of coupling constants that can be renormalized with a finite pieces of experimental data to render the theory finite. Instead there are an infinite number of counterterms need to make the theory finite, and therefore an infinite number of experimental data needed to fix the finite parts of the counterterms. Such a theory would be unless there existed some sort of expansion that let us deal with only a finite set of operators at each order in that expansion.

Wilson provided such an expansion. The first thing to accept is that the EFT has an intrinsic, finite UV cutoff Λ. This scale is typically the mass of the lightest particles *omitted* from the theory. For example, in the Fermi theory of the weak interactions, $\Lambda = M_W$. With a cutoff in place, all radiative corrections in the theory are finite, even if they are proportional to powers or logarithms of Λ. The useful expansion then is a momentum expansion, in powers of k/Λ, where k is the external momentum in some physical process of interest, such as a particle decay, two particle scattering, two particle annihilation, etc. This momentum expansion is the key tool that makes EFTs useful. To understand how this works, we need to develop the concept of operator dimension. In this lecture we will only consider the EFT at tree level.

2.1 Scaling in a relativistic EFT

As a prototypical example of an EFT, consider the Lagrangian (in four dimensional Euclidean spacetime, after a Wick rotation to imaginary time) for relativistic scalar field with a $\phi \to -\phi$ symmetry:

$$\mathcal{L}_E = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}\phi^4 + \sum_{n=1}^{\infty}\left(\frac{c_n}{\Lambda^{2n}}\phi^{4+2n} + \frac{d_n}{\Lambda^{2n}}(\partial\phi)^2\phi^{2n} + \dots\right) \tag{57}$$

We are setting $\hbar = c = 1$ so that momenta have dimension of mass, and spacetime coordinates have dimension of inverse mass. I indicate this as

$$[p] = 1\ , \qquad [x] = [t] = -1\ , \qquad [\partial_x] = [\partial_t] = 1\ . \tag{58}$$

Since the action is dimensionless, then — in $d = 4$ spacetime dimensions — from the kinetic term for ϕ we see that ϕ has dimension of mass:

$$[\phi] = 1 \ . \tag{59}$$

That means that the operator ϕ^6 is dimension 6, and the contribution to the action $\int d^4x\, \phi^6$ has dimension 2, and so its coupling constant must have dimension -2, or $1/\text{mass}^2$. The operator $\phi^2(\partial^2\phi)^2$ is dimension 8 and must have a coefficient which is dimension -4, or $1/\text{mass}^4$. In eq. (57) I have introduced the cutoff scale Λ explicitly into the Lagrangian in such a way as to make the the couplings λ, c_n and d_n all dimensionless, with no loss of generality. I will assume here that $\lambda \ll 1$, $c_n \ll 1$ and $d_n \ll 1$ so that a perturbative expansions in these couplings is reasonable.

You might ask why we do things this way — why not rescale the ϕ^6 operator to have coefficient 1 instead of the kinetic term, and declare ϕ to have dimension $2/3$? The reason why is because the kinetic term is more important and determines the size of quantum fluctuations for a relativistic excitation. To see this, consider the path integral

$$\int D\phi\, e^{-S_E} \ , \qquad S_E = \int d^4x\, \mathcal{L}_E \ . \tag{60}$$

Now consider a particular field configuration contributing to this path integral that looks like the "wavelet" pictured in Fig. 4, with wavenumber $|k_\mu| \sim k$, localized to a spacetime volume of size L^4, where $L \simeq 2\pi/k$, and with amplitude ϕ_k. Derivatives acting on such a configuration give powers of k, while spacetime integration gives a factor of $L^4 \simeq (4\pi/k)^4$. With this configuration, the Euclidean action is given by

$$\begin{aligned} S_E &\simeq \left(\frac{2\pi}{k}\right)^4 \left[\frac{k^2\phi_k^2}{2} + \frac{1}{2}m^2\phi_k^2 + \frac{\lambda}{4!}{\phi_k}^4 + \sum_{n=1}^{\infty}\left(\frac{c_n}{\Lambda^{2n}}{\phi_k}^{4+2n} + \frac{d_n k^2}{\Lambda^{2n}}{\phi_k}^{2+2n} + \ldots\right)\right] \\ &= (2\pi)^4 \left[\frac{\widehat{\phi}_k^2}{2} + \frac{1}{2}\frac{m^2}{k^2}\widehat{\phi}_k^2 + \frac{\lambda}{4!}\widehat{\phi}_k^4 + \sum_n \left(c_n\left(\frac{k^2}{\Lambda^2}\right)^n \widehat{\phi}_k^{4+2n} + d_n\left(\frac{k^2}{\Lambda^2}\right)^n \widehat{\phi}_k^{2+2n} + \ldots\right)\right] , \end{aligned} \tag{61}$$

where in the second line I have rescaled the amplitude by k,

$$\widehat{\phi}_k \equiv \phi_k/k \ . \tag{62}$$

In the above expression, the factor of $\left(\frac{2\pi}{k}\right)^4$ in front is the spacetime volume occupied by our wavelet and comes from $\int d^4x$, while for every operator I have substituted $\partial \to k$ and $\phi \to k\widehat{\phi}_k$. Now for the path integral, consider ordinary integration over the amplitude $\widehat{\phi}_k$ for this particular mode:

$$\int d\widehat{\phi}_k\, e^{-S_E} \ . \tag{63}$$

The integral is dominated by those values of $\widehat{\phi}_k$ for which $S_E \lesssim 1$, because otherwise $\exp(-S_E)$ is very small. Which are the important terms in S_E in this region? First, assume that the particle is relativistic, $m \ll k \ll \Lambda$ so that both m^2/k^2 and k^2/Λ^2 are very small,

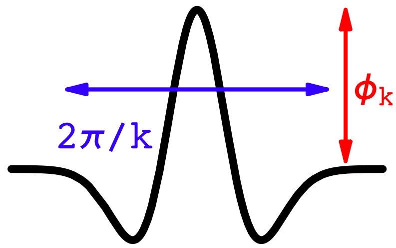

Figure 4: *sample configuration contributing to the path integral for the scalar field theory in eq. (57). Its amplitude is ϕ_k and has wave number $\sim k$ and spatial extent $\sim 2\pi/k$.*

and assume the dimensionless couplings λ, c_n, d_n are $\lesssim O(1)$. Then as one increases the amplitude $\widehat{\phi}_k$ from zero, the first term in S_E to become become $O(1)$ is the kinetic term, $(2\pi)^4\widehat{\phi}_k^2$, which occurs for $\phi_k = k\widehat{\phi}_k \sim k/(2\pi)^2$. It is because the kinetic term controls the fluctuations of the scalar field that we "canonically normalize" the field such that the kinetic term is $\frac{1}{2}(\partial\phi)^2$, and perturb in the coefficients of the other operators in the theory. Is this conclusion always true? No. For low enough k, for example, the mass term with its factor of m^2/k^2 will eventually dominate and a different scaling regime takes over. Also, if some of the dimensionless couplings λ, c_n, d_n are large, those terms may dominate and the theory will change its nature dramatically. In the first lecture we looked at scattering off a δ-function in $D = 3$ and saw that if the coupling was tuned to a particular strong value its effects could dominate low energy scattering, even though it is naively an irrelevant interaction.

What happens as we consider different momenta k? We see from eq. (61) that as k is reduced, the c_n and d_n terms, proportional to $(k^2/\Lambda^2)^n$, get smaller. Such operators are "irrelevant" operators in Wilson's language, because they become unimportant in the infrared (low k). In contrast, the mass term becomes more important; it is called a "relevant" operator. The kinetic term and the $\lambda\phi^4$ interaction do not change; such operators are called "marginal". It used to be thought that the irrelevant operators were dangerous, making the theory nonrenormalizable, while the relevant operators were safe – "superrenormalizable". As we consider radiative corrections later we will see that Wilson flipped this entirely on its head, so that irrelevant operators are now considered safe, while the existence of relevant operators is thought to be a serious problem to be solved.

In practice, when working with a relativistic theory in d spacetime dimensions with small dimensionless coupling constants, the operators with dimension d are the marginal ones, those with higher dimension are irrelevant, and those with lower dimension are relevant. The bottom line is that we can analyze the theory in a momentum expansion, working to a particular order and ignoring irrelevant operators above a certain dimension. The ability to do so will persist even when we include radiative corrections.

2.1.1 Dimensional analysis: Fermi's theory of the weak interactions

To see why dimensional analysis has practical consequences, first consider Fermi's theory of the weak interactions. Originally this was a "bottom-up" sort of EFT — Fermi did not have a complete UV description of the weak interactions, and so constructed the theory as a phenomenological modification of QED to account for neutron decay. Now we have the SM, and so we think of the Fermi theory as a "top-down" EFT: not necessary for doing calculations since we have the SM, but very practical for processes at energies far below the W mass, such as β-decay (Fermi's original application of his theory) and low energy neutral current scattering due to Z exchange (about which Fermi knew nothing).

The weak interactions refer to processes mediated by the $W^{\pm}$ or Z^0 bosons, whose masses are approximately 80 GeV and 91 GeV respectively. The couplings of these gauge bosons to quarks and leptons can be written in terms of the electromagnetic current

$$j^{\mu}_{\text{em}} = \frac{2}{3}\bar{u}_i\gamma^{\mu}u_i - \frac{2}{3}\bar{d}_i\gamma^{\mu}d_i - \bar{e}_i\gamma^{\mu}e_i \tag{64}$$

where $i = 1, 2, 3$ runs over families, and the left-handed $SU(2)$ currents

$$j^{\mu}_a = \sum_{\psi}\overline{\psi}\gamma^{\mu}\left(\frac{1-\gamma_5}{2}\right)\frac{\tau_a}{2}\psi\ , \qquad a = 1, 2, 3\ , \tag{65}$$

where the $\overline{\psi}, \psi$ fields in the currents are either the lepton doublets

$$\begin{pmatrix}\nu_e\\ e\end{pmatrix}\ ,\quad \begin{pmatrix}\nu_{\mu}\\ \mu\end{pmatrix}\ ,\quad \begin{pmatrix}\nu_{\tau}\\ \tau\end{pmatrix}\ , \tag{66}$$

or the quark doublets

$$\psi = \begin{pmatrix}u\\ d'\end{pmatrix}\ ,\quad \begin{pmatrix}c\\ s'\end{pmatrix}\ ,\quad \begin{pmatrix}t\\ b'\end{pmatrix}\ , \tag{67}$$

with the "flavor eigenstates" d', s' and b' being related to the mass eigenstates d, s and b by the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix[3]:

$$q'_i = V_{ij}q_j\ . \tag{68}$$

The SM coupling of the heavy gauge bosons to these currents is

$$\mathcal{L}_{\text{J}} = \frac{e}{\sin\theta_w}\left(W^+_{\mu}J^{\mu}_{-} + W^-_{\mu}J^{\mu}_{+}\right) + \frac{e}{\sin\theta_w\,\cos\theta_w}Z_{\mu}\left(j^{\mu}_3 - \sin^2\theta_w j^{\mu}_{\text{em}}\right) \tag{69}$$

where

$$J^{\mu}_{\pm} = \frac{j^{\mu}_1 \mp i\, j^{\mu}_2}{\sqrt{2}}\ . \tag{70}$$

Tree level exchange of a W boson then gives the amplitude at low momentum exchange

$$i\mathcal{A} = \left(-i\frac{e}{\sin\theta_w}\right)^2 J^{\mu}_{-}J^{\nu}_{+}\frac{-ig_{\mu\nu}}{q^2 - M_W^2} = -i\frac{e^2}{\sin^2\theta_w M_W^2}J^{\mu}_{-}J_{\mu+} + O\left(\frac{q^2}{M_W^2}\right)\ . \tag{71}$$

[3]The elements of the CKM matrix are named after which quarks they couple through the charged current, namely $V_{11} \equiv V_{ud}$, $V_{12} \equiv V_{us}$, $V_{21} \equiv V_{cd}$, etc.

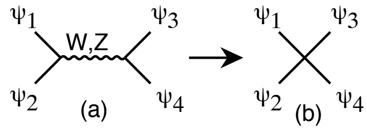

Figure 5: *(a) Tree level W and Z exchange between four fermions. (b) The effective vertex in the low energy effective theory (Fermi interaction).*

This amplitude can be reproduced to lowest order in q^2/M_W^2 by a low energy EFT with a contact interaction, Fig. 5,

$$\mathcal{L}_{\rm F} = -\frac{e^2}{\sin^2\theta_w M_W^2} J_-^\mu J_{\mu+} = \frac{8}{\sqrt{2}} G_F J_-^\mu J_{\mu+} \ , \tag{72}$$

$$G_F \equiv \frac{\sqrt{2}}{8}\frac{e^2}{\sin^2\theta_w M_W^2} = 1.166\times 10^{-5}\ \ \mathrm{GeV}^2\ . \tag{73}$$

This is our matching condition, analogous to the matching we did for δ-function scattering in order to reproduce the low energy behavior of square well scattering. This charged current interaction, written in terms of leptons and nucleons instead of leptons and quarks, was postulated by Fermi to explain neutron decay; the $8/\sqrt{2}$ numerical factor looks funny here because I am normalizing the currents in the way they appear in the SM, while weak currents are historically (pre-SM) normalized differently. Neutral currents were proposed in the 60's and discovered in the 70's.

Since the four-fermion Fermi interaction has dimension 6, it is an irrelevant interaction, according to our previous discussion, explaining why we say the interactions are "weak" and neutrinos are "weakly interacting". Consider, for example, some low energy neutrino scattering cross section σ. Since neutrinos only interact via W and Z exchange, the cross-section σ must be proportional to G_F^2 which has dimension -4. But a cross section has dimensions of area, or mass dimension -2. Since the only other scale around is the center of mass energy $\sqrt{s}$, on purely dimensional grounds σ must scale with energy as

$$\sigma_\nu \simeq G_F^2 s\ , \tag{74}$$

This explains why low energy neutrinos are so hard to detect, and the weak interactions are weak; at LHC energies, however, where the effective field theory has broken down, the weak interactions are marginal and characterized by the $SU(2)$ coupling constant $g \simeq 0.6$, about twice as strong as the electromagnetic coupling. It is a simple result for which one does not need the full machinery of the SM to derive.

It looks like the neutrino cross section grows with s without bound, but remember that this EFT is only valid up to $s \simeq M_W$.

2.1.2 Dimensional analysis: the blue sky

Another top-down application of EFT is to answer the question of why the sky is blue. More precisely, why low energy light scattering from neutral atoms in their ground state (Rayleigh scattering) is much stronger for blue light than red[4]. The physics of the scattering process could be analyzed using exact or approximate atomic wave functions and matrix elements, but that is overkill for low energy scattering. Let's construct an "effective Lagrangian" to describe this process. This means that we are going to write down a Lagrangian with all interactions describing elastic photon-atom scattering that are allowed by the symmetries of the world — namely Lorentz invariance and gauge invariance. Photons are described by a field A_μ which creates and destroys photons; a gauge invariant object constructed from A_μ is the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The atomic field is defined as ϕ_v, where ϕ_v destroys an atom with four-velocity v_μ (satisfying $v_\mu v^\mu = 1$, with $v_\mu = (1,0,0,0)$ in the rest-frame of the atom), while $\phi_v^\dagger$ creates an atom with four-velocity v_μ. In this case we should use relativistic scaling, since we are interested in on-shell photons, and are uninterested in recoil effects (the kinetic energy of the atom):

$$[x] = [t] = -1, \quad [p] = [E] = [A_\mu] = 1 \, , \qquad [\phi] = \frac{3}{2} \, , \tag{75}$$

where the atomic field ϕ destroys an atom with four-velocity v_μ (satisfying $v_\mu v^\mu = 1$, with $v_\mu = (1,0,0,0)$ in the rest-frame of the atom), while $\phi^\dagger$ creates an atom with four-velocity v_μ.

So what is the most general form for $\mathcal{L}_{\rm eft}$? Since the atom is electrically neutral, gauge invariance implies that ϕ can only be coupled to $F_{\mu\nu}$ and not directly to A_μ. So $\mathcal{L}_{\rm eft}$ is comprised of all local, Hermitian monomials in $\phi^\dagger\phi$, $F_{\mu\nu}$, v_μ, and ∂_μ. Certain combinations we needn't consider for the problem at hand — for example $\partial_\mu F^{\mu\nu} = 0$ for radiation (by Maxwell's equations); also, if we define the energy of the atom at rest in it's ground state to be zero, then $v^\mu \partial_\mu \phi = 0$, since $v_\mu = (1,0,0,0)$ in the rest frame, where $\partial_t \phi = 0$. Similarly, $\partial_\mu \partial^\mu \phi = 0$. Thus we are led to consider the interaction Lagrangian

$$\begin{aligned}\mathcal{L}_{\rm eft} \;=\; & c_1 \phi^\dagger \phi F_{\mu\nu} F^{\mu\nu} + c_2 \phi^\dagger \phi v^\alpha F_{\alpha\mu} v_\beta F^{\beta\mu} \\ & + c_3 \phi^\dagger \phi (v^\alpha \partial_\alpha) F_{\mu\nu} F^{\mu\nu} + \ldots \end{aligned} \tag{76}$$

The above expression involves an infinite number of operators and an infinite number of unknown coefficients! Nevertheless, dimensional analysis allows us to identify the leading contribution to low energy scattering of light by neutral atoms.

With the scaling behavior eq. (75), and the need for $\mathcal{L}$ to have dimension 4, we find the dimensions of our couplings to be

$$[c_1] = [c_2] = -3 \, , \qquad [c_3] = -4 \, . \tag{77}$$

Since the c_3 operator has higher dimension, we will ignore it. What are the sizes of the coefficients $c_{1,2}$? To do a careful analysis one needs to go back to the full Hamiltonian for the

[4]By "low energy" I mean that the photon energy E_γ is much smaller than the excitation energy ΔE of the atom, which is of course much smaller than its inverse size or mass:

$$E_\gamma \ll \Delta E \ll r_0^{-1} \ll M_{atom}$$

where r_0 is the atomic size, roughly the Bohr radius. Thus the process is necessarily elastic scattering, and to a good approximation we can ignore that the atom recoils, treating it as infinitely heavy.

atom in question interacting with light, and "match" the full theory to the effective theory. Here I will just estimate the sizes of the c_i coefficients, rather than doing some atomic physics calculations. Note that extremely low energy photons cannot probe the internal structure of the atom, and so the cross-section ought to be classical, only depending on the size of the scatterer, which I will denote as r_0, roughly the Bohr radius. Since such low energy scattering can be described entirely in terms of the coefficients c_1 and c_2, we conclude that

$$c_1 \simeq c_2 \simeq r_0^3 \ .$$

The effective Lagrangian for low energy scattering of light is therefore

$$\mathcal{L}_{\rm eft} = r_0^3 \left(a_1 \phi^\dagger \phi F_{\mu\nu} F^{\mu\nu} + a_2 \phi^\dagger \phi v^\alpha F_{\alpha\mu} v_\beta F^{\beta\mu} \right) \tag{78}$$

where a_1 and a_2 are dimensionless, and expected to be $\mathcal{O}(1)$. The cross-section (which goes as the amplitude squared) must therefore be proportional to r_0^6. But a cross section σ has dimensions of area, or $[\sigma] = -2$, while $[r_0^6] = -6$. Therefore the cross section must be proportional to

$$\sigma \propto E_\gamma^4 r_0^6 \ , \tag{79}$$

growing like the fourth power of the photon energy. Thus blue light is scattered more strongly than red, and the sky far from the sun looks blue. The two independent coefficients in this calculation must correspond to the electric and magnetic polarizabilities of the atom.

Is the expression eq. (79) valid for arbitrarily high energy? No, because we ignored higher dimension terms in the effective Lagrangian we used, terms which become more important at higher energies — and at sufficiently high energy these terms are all in principle equally important and the EFT breaks down. To understand the size of corrections to eq. (79) we need to know the size of the c_3 operator (and the rest we ignored). Since $[c_3] = -4$, we expect the effect of the c_3 operator on the scattering amplitude to be smaller than the leading effects by a factor of E_γ/Λ, where Λ is some energy scale. But does Λ equal M_{atom}, $r_0^{-1} \sim \alpha m_e$ or $\Delta E \sim \alpha^2 m_e$? The latter — the energy required to excite the atom — is the smallest energy scale and hence the most important. We expect our approximations to break down as $E_\gamma \to \Delta E$ since for such energies the photon can excite the atom. Hence we predict

$$\sigma \propto E_\gamma^4 r_0^6 \left(1 + \mathcal{O}(E_\gamma/\Delta E)\right) . \tag{80}$$

The Rayleigh scattering formula ought to work pretty well for blue light, but not very far into the ultraviolet. Note that eq. eq. (80) contains a lot of physics even though we did very little work. More work is needed to compute the constant of proportionality.

2.2 Accidental symmetry and BSM physics

Now let's switch tactics and use dimensional analysis to talk about bottom-up applications of EFT. We would like to have clues of physics beyond the SM (BSM). Evidence we currently have for BSM physics are the existence of gravity, neutrino masses and dark matter. Hints for additional BSM physics include circumstantial evidence for Grand Unification and for inflation, the absence of a neutron electric dipole moment, and the baryon number asymmetry of the universe. Great puzzles include the origin of flavor and family structure, why the electroweak scale is so low compared to the Planck scale (but not so far from the QCD scale), and why we live in an epoch where matter, dark matter, and dark energy all have have rather similar densities.

In order to make progress we would like to have more data, and looking for subtle effects due to irrelevant operators can in some cases give us a much farther experimental reach than can collider physics, exactly in the same way the Fermi interaction provided a critical clue which eventually led to the SM. Those cases are necessarily ones where the irrelevant operators violate symmetries that are preserved by the marginal and irrelevant operators in the SM, and are therefore the leading contribution to certain processes. We call these symmetries "accidental symmetries": they are not symmetries of the UV theory, but they are approximate symmetries of the IR theory.

A simple and practical example of an accidental symmetry is $SO(4)$ symmetry in lattice QCD — the Euclidian version of the Lorentz group. Lattice QCD formulates QCD on a 4d hypercubic lattice, and then looks in the IR on this lattice, focusing on modes whose wavelengths are so long that they are insensitive to the discretization of spacetime. But why is it obvious that a hypercubic lattice will yield a continuum Lorentz invariant theory?

The reason lattice field theory works is because of accidental symmetry: Operators on the lattice are constrained by gauge invariance and the hypercubic symmetry of the lattice. While it is possible to write down operators which are invariant under these symmetries while violating the $SO(4)$ Lorentz symmetry, such operators all have high dimension and are not relevant. For example, if A_μ is a vector field, the $SO(4)$-violating operator $A_1A_2A_3A_4$ is hypercubic invariant and marginal (dimension 4) and so could spoil the continuum limit we desire; however, the only vector field in lattice QCD is the gauge potential, and such an operator is forbidden because it is not gauge invariant. In the quark sector the lowest dimension operator one can write which is hypercubic symmetric but Lorentz violating is

$$\sum_{\mu=1}^{4} \bar{\psi}\gamma_\mu D_\mu^3 \psi \tag{81}$$

which is dimension six and therefore irrelevant. Thus Lorentz symmetry is automatically restored in the continuum limit.

Accidental symmetries in the SM notably include baryon number B and lepton number L: if one writes down all possible dimension ≤ 4 gauge invariant and Lorentz invariant operators in the SM, you will find they all preserve B and L. It is possible to write down dimension five $\Delta L = 2$ operators and dimension six $\Delta B = \Delta L = 1$ operators, however. That means that no matter how completely B and L are broken in the UV, at our energies these irrelevant operators become...irrelevant, and B and L appear to be conserved, at least to high precision. So perhaps B and L are not symmetries of the world at all – they just look like good symmetries because the scale of new physics is very high, so that the irrelevant B and L violating operators have very little effect at accessible energies. We will look at these different operators briefly in turn.

2.2.1 BSM physics: neutrino masses

The most important irrelevant operators that could be added to the SM are dimension 5. Any such operator should be constructed out of the existing fields of the SM and be invariant under the $SU(3) \times SU(2) \times U(1)$ gauge symmetry. Recall that the matter fields in the SM have the gauge quantum numbers

$$
\begin{aligned}
\text{LH fermions:} \quad & Q = (3,2)_{\frac{1}{6}} \ , \quad U = (\bar{3},1)_{-\frac{2}{3}} \ , \quad D = (\bar{3},1)_{\frac{1}{3}} \ , \quad L = (1,2)_{-\frac{1}{2}} \ , \quad E = (1,1)_1 \ , \\
\text{Higgs:} \quad & H = (1,2)_{-\frac{1}{2}} \ , \quad \widetilde{H} = (1,2)_{\frac{1}{2}} \ ,
\end{aligned}
\tag{82}
$$

where $\widetilde{H}_i = \epsilon_{ij} H_j^*$ is not an independent field. The gauge fields transform as adjoints under their respective gauge groups.

The only gauge invariant dimension 5 operator one can write down is the $\Delta L = 2$ operator (violating lepton number by two units) [5]:

$$
\mathcal{L}_{\Delta L=2} = -\frac{1}{\Lambda}(L\widetilde{H})(L\widetilde{H}) \ , \qquad L = \begin{pmatrix} \nu \\ \ell^- \end{pmatrix} \ , \qquad \widetilde{H} = \begin{pmatrix} h^+ \\ h^0 \end{pmatrix} \ , \qquad \langle \widetilde{H} \rangle = \frac{v}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ , \tag{83}
$$

where $v = 250$ GeV. There is only one independent operator (ignoring lepton flavor) since the two Higgs fields ($\widetilde{H}\widetilde{H}$) cannot be antisymmetrized and therefore must be in an $SU(2)$-triplet. An operator coupling LL in a weak triplet to HH in a weak triplet can be rewritten in the above form, where the combination $(L\widetilde{H})$ is a weak singlet,

$$
(L\widetilde{H}) = \left(\nu h^0 - \ell^- h^+\right) \quad \longrightarrow \quad \frac{\nu v}{\sqrt{2}} \ . \tag{84}
$$

Therefore after spontaneous symmetry breaking by the Higgs, the operator gives a contribution to the neutrino mass,

$$
\mathcal{L}_{\Delta L=2} = -\frac{1}{2} m_\nu \nu\nu \ , \qquad m_\nu = \frac{v^2}{\Lambda} \ , \tag{85}
$$

a $\Delta L = 2$ Majorana mass for the neutrino. A mass of $m_\nu = 10^{-2}$ eV corresponds to $\Lambda = 6 \times 10^{15}$ GeV, an interesting scale, being near the scale of GUT models, and far beyond the reach of accelerator experiments. Or: if $\Lambda = 10^{19}$ GeV, the Planck scale, then $m_\nu = 10^{-5}$ eV. This operator provides a possible and rather compelling explanation for the smallness of observed neutrino masses: they arise as Majorana masses because lepton number is not a symmetry of the universe, but are very small because lepton number becomes an accidental symmetry below a high scale. Of course, we could have the spectrum of the low energy theory wrong: perhaps there is a light right-handed neutrino and neutrinos only have L-preserving Dirac masses like the charged leptons, small simply because of a very small Yukawa coupling to the Higgs. Neutrinoless double beta decay experiments are searching for lepton number violation in hopes of establishing the Majorana mass scenario.

In any case, it is interesting to imagine what sort of UV physics could give rise to the operator in eq. (83). Three possibilities present themselves for how such an operator could

[5] One might expect to be able to write down magnetic dipole operators of the form $\bar{\psi}\sigma_{\mu\nu}F^{\mu\nu}\psi$, but such operators have the chiral structure of a mass term and require an additional Higgs field to be gauge invariant, making them dimension 6.

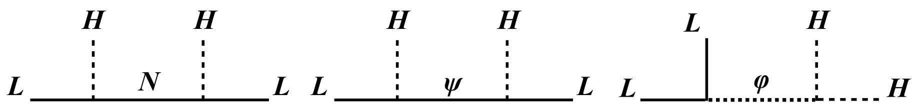

Figure 6: *Three ways the dimension 5 operator for neutrino masses in eq. (83) could arise from tree level exchange of a heavy $SU(2)\times U(1)$ singlet fermion N, a heavy $SU(2)\times U(1)$ triplet fermion ψ, or else from exchange of a massive $SU(2)$ triplet scalar ϕ.*

arise from a high energy theory at tree level, shown in Fig. 6 – either through exchange of a heavy $SU(2)\times U(1)$ singlet fermion N (a "right handed neutrino"), through exchange of a heavy $SU(2)\times U(1)$ triplet fermion ψ, or else via exchange of a heavy scalar with quantum number 3_1 under $SU(2)\times U(1)$. The fact that the resultant light neutrino mass is inversely proportional to the new scale of physics (called the "see-saw mechanism") simply results from the fact that a neutrino mass operator in the SM is an irrelevant dimension-5 operator. Note that just as G_F is proportional to g^2/M_W^2, and therefore knowing G_F was not sufficient for predicting the W mass, the scale Λ is not necessarily the mass of a new particle, as it will be inversely proportional to coupling constants about which we know nothing except in the context of some particular UV candidate theory.

A fourth possibility for neutrino masses does not arise in EFT at all: the neutrino mass might be Dirac, with right-handed neutrino we have not detected (since it is neutral under gauge symmetries) which couples to the lepton doublet L via an extremely small Yukawa coupling to the Higgs. In this case lepton symmetry is not violated. Seeing neutrinoless double β decay would be a signal of a $\Delta L = 2$ process and would be evidence in favor of a seesaw origin for neutrino masses.

2.2.2 BSM physics: proton decay

At dimension 6 one can write down very many new operators in the SM, including interesting CP violating electric dipole moment operators for fermions. One particularly interesting set of dimension 6 operators in the SM are those that violate B; they all consist of three quark fields (for color neutrality) and a lepton field, and are therefore all $\Delta B = 1$, $\Delta L = 1$ operators which conserve the combination $B - L$. These are very interesting because (i) B is a particularly good symmetry, since the proton appears to be very stable, and (ii) B violation is a prerequisite for any theory of baryogenesis. Below the QCD scale one needs to match the three quark operator onto hadron fields. An example of such an operator would be

$$\frac{1}{\Lambda^2}\epsilon_{abc}\epsilon_{\alpha\beta}\epsilon_{\gamma\delta}(d_L^{a\alpha}u_L^{b\beta})(u_L^{c\gamma}e_L^{\delta} - d_L^{c\gamma}\nu_L^{\delta})\ , \tag{86}$$

where a, b, c are color indices and $\alpha, \beta, \gamma, \delta$ are $SU(2)$ Lorentz indices for the left-handed Weyl spinors; the terms in parentheses are weak $SU(2)$ singlets, and the whole operator is neutral under weak hypercharge. Below the QCD scale one has to match the three-quark operator onto hadrons fields. Thus roughly speaking $uud \to Z_1 p + Z_2(p\pi^0 + n\pi^+) + \ldots$. We can assume that the Z factors are made up of pure numbers times the appropriate powers of the strong interaction scale, such as $f_\pi \simeq 100$ MeV, the pion decay constant. The Z_1 term will lead to positron-proton mixing and cannot lead to proton decay, but the Z_2 term can via the processes $p \to e^+\pi^0$, or $p \to \pi^+\nu$. We can make a crude estimate of the width (inverse lifetime) to be

$$\Gamma \simeq \frac{M_p^5}{\Lambda^4}\frac{1}{8\pi} \tag{87}$$

where I used dimensional analysis to estimate the M_p^5/Λ^4 factor, assuming that the strong interaction scale in Z_2 as well as powers of momenta from phase space integrals could be approximated by the proton mass M_p , and I inserted a typical 2-body phase space factor of $1/8\pi$. For a bound on the proton lifetime of $\tau_p > 10^{34}$ years, this crude estimate gives us $\Lambda \gtrsim 10^{16}$ GeV, not so far off the bound one finds from a more sophisticated calculation. If proton decay is discovered, that will tell us something about the scale of new physics, and then the task will be to construct the full UV theory from what we learn about proton decay, much as the SM was discovered starting from the Fermi theory.

2.3 BSM physics: "partial compositeness"

This next topic does not have to do with accidental symmetry violation, but instead picks up on an interesting feature of the baryon number violating interaction we just discussed, as it suggests a mechanism for quarks and leptons to acquire masses without a Higgs. In estimating the effects of the dimension six $\Delta B = 1$ operator in the previous section I said that the 3-quark operator could be expanded as $uud \to Z_1 p + Z_2(p\pi^0 + n\pi^+) + \ldots$, and then focussed on the Z_2 term. But what about the Z_1 term? By dimensions, $Z_1 \sim \Lambda_{QCD}^3$, and so that term gives rise to a peculiar mass term of the form

$$\frac{\Lambda_{QCD}^3}{\Lambda^2} pe + h.c. \tag{88}$$

which allows a proton to mix with a positron, and the anti-proton to mix with the electron. This is not experimentally interesting, but it is an interesting phenomenon for a theorist to contemplate. Imagine eliminating the Higgs doublet from the SM. The proton would still get a mass from chiral symmetry breaking in QCD even though the quarks would remain massless, and due to the above term, and even though there would not be an electron mass directly from the weak interactions, there would be a above contribution to the proton mass due to the above positron-proton mixing. For the two component system one would have a mass matrix looking something like

$$\begin{pmatrix} M_p & \frac{\Lambda_{QCD}^3}{\Lambda^2} \\ \frac{\Lambda_{QCD}^3}{\Lambda^2} & 0 \end{pmatrix} \tag{89}$$

and to the extent that $\Lambda \gg \Lambda_{QCD}$ we find the mass eigenvalues to be

$$m_1 \simeq M_p \ , \qquad m_2 \simeq \frac{\Lambda_{QCD}^6}{M_p \Lambda^4} \tag{90}$$

so for $\Lambda = 10^{16}$ GeV the positron gets a mass of $m_e \simeq 10^{-64}$ GeV. Yes, this is a ridiculously small mass of no interest, but it is curious that the positron (electron) got a mass at all, without there being any Higgs field! It must be that QCD when QCD spontaneously breaks chiral symmetry, it has also broken $SU(2) \times U(1)$, without a Higgs, and that this proton decay operator has somehow taken the place of a Higgs Yukawa coupling. Quark fields and QCD has assumed both of the the roles that the Higgs field plays in the SM. Therefore it is worth asking whether this example be modified somehow to obtain more interesting masses for quarks and leptons?

In a later lecture we will examine how QCD breaks the weak interactions, and how a scaled up version called technicolor, with the analogue of the pion decay constant f_π being up at the 250 GeV scale instead of 93 MeV, could properly account for the spontaneous breaking of $SU(1) \times U(1)$ without a Higgs. Here I will just comment that such a theory would be expected to have TeV mass "technibaryons", which could carry color and charge. With an appropriate dimension 6 operator such as our proton decay operator, but with techniquarks in place of quarks, and all the standard model fermions in place of the positron field, in principle one could give masses to all the SM fermions through their mixing with the technibaryons. This is the idea of "partial compositeness", which in its original formulation [1] was not especially useful for model building, but which has become more interesting in the context of composite Higgs [2] – more about composite Higgs later too.

,

2.4 Problems for lecture II

II.1) What is the dimension of the operator ϕ^{10} in a $d = 2$ relativistic scalar field theory?

II.2) One defines the "critical dimension" d_c for an operator to be the spacetime dimension for which that operator is marginal. How will that operator behave in dimensions d when $d > d_c$ or $d < d_c$? In a theory of interacting relativistic scalars, Dirac fermions, and gauge bosons, determine the critical dimension for the following operators:

1. A ϕ^3 interaction;
2. A gauge coupling to either a fermion or a boson through the covariant derivative in the kinetic term;
3. A Yukawa interaction, $\phi\bar{\psi}\psi$;
4. An anomalous magnetic moment coupling $\bar{\psi}\sigma_{\mu\nu}F^{\mu\nu}\psi$ for a fermion;
5. A four fermion interaction, $(\bar{\psi}\psi)^2$.

II.3) Derive the analogue of Fermi's theory in eq. (72) for tree level Z exchange, expressing your answer in terms of G_F using the fact that $M_Z^2 = M_W^2/\cos^2\theta_w$.

II.4) How would one write down an electric dipole operator in QED, and what dimension does it have? What would you have to do to make a gauge invariant electric dipole operator in the SM that is invariant under the full $SU(3) \times SU(2) \times U(1)$ gauge symmetry?

II.5) Show that the operator

$$\epsilon_{\alpha\beta}\left(L_{\alpha i}(\sigma_2\sigma^a)_{ij}L_{\beta j}\right)\left(H_k(\sigma_2\sigma^a)_{k\ell}H_\ell\right)$$

is equivalent up to a factor of two to

$$\epsilon_{\alpha\beta}(L_{\alpha i}(\sigma_2)_{ij}H_j)(L_{\beta k}(\sigma_2)_{k\ell}H_\ell)$$

where α, β are Weyl spinor indices, and i, j, k, ℓ are the $SU(2)$ gauge group indices. Write down the two high energy theories that could give rise to the neutrino mass operator as in Fig. 6. How do I see that these theories break lepton number by two units?

3 EFT and radiative corrections

Up to now we have ignored quantum corrections in our effective theory. A Lagrangian such as eq. (57) is what used to be termed a "nonrenormalizable" theory, and to be shunned. The problem was that the theory needs an infinite number of counterterms to subtract all infinities, and was thought to be unpredictive. In contrast, a "renormalizable" theory contained only marginal and relevant operators, and needed only a finite number of counterterms, one per marginal or relevant operator allowed by the symmetries. (A "superrenormalizable" theory contained only relevant operators, and was finite beyond a certain order in perturbation theory.) However Wilson changed the view of renormalization. In a perturbative theory, irrelevant operators are renormalized, but stay irrelevant. On the other hand, the coefficients of relevant operators are renormalized to take on values proportional to powers of the cutoff, unless forbidden by symmetry. Thus in Wilson's view the relevant operators are the problem, since giving them small coefficients requires fine-tuning – unless a symmetry forbids corrections that go as powers of the cutoff. Relevant operators protected by symmetry include fermion masses and Goldstone boson masses, but for a general interacting scalar, the natural mass is $m^2 \simeq \alpha\Lambda^2$ — which means one should never see such scalars in the low energy theory.

In this lecture I discuss the techniques used to create top-down EFTs beyond tree level, as well as an example of an EFT with a marginal interaction with asymptotic freedom and an exponentially small IR scale.

3.1 Matching

I will consider a toy model for UV physics with a light scalar ϕ and a heavy scalar S :

$$\mathcal{L}_{\rm UV} = \frac{1}{2}\left((\partial\phi)^2 - m^2\phi^2 + (\partial S)^2 - M^2 S^2 - \kappa\phi^2 S\right) \tag{91}$$

The parameter κ has dimension of mass, and I will assume $\kappa \lesssim M$ and that $\langle S\rangle = 0$. This is a pretty sill model, but it is useful as an example since the Feynman diagrams are very simple; never mind that the vacuum energy is unboundedbelow, as one won't see this in perturbation theory. Suppose we are interested in $2\phi \to 2\phi$ scattering at energies much below the S mass M, and want to construct the EFT with the heavy S field "integrated out". Formally:

$$e^{-i\frac{\mathcal{S}_{\rm eft}(\phi)}{\hbar}} = \int [dS] e^{-i\frac{\mathcal{S}(\phi,S)}{\hbar}} \tag{92}$$

where I use $\mathcal{S}$ to denote the action, to distinguish it from the field S. There are good reasons for doing so: if you try to compute observables in this theory at some low momentum $k << M$ you are typically going to run into large logarithms of the form $\ln k^2/M^2$ that will spoil perturbation theory. They are easily taken care of in an EFT where you integrate out S at the scale $\mu = M$, matching the EFT to the full theory to ensure that you are reproducing the same physics. Then within the EFT you run the couplings from $\mu = M$

down to $\mu = k$ before doing your calculation. The renormalization group running sums up these large logs for you.

It is convenient to perform the matching in an $\hbar$ expansion, meaning that first you make sure that tree diagrams agree in the two theories, as in our derivation of the Fermi theory of weak interactions from the SM. Then you make sure that the two theories agree at $O(\hbar)$, etc. What sort of graphs does this matching entail, and why is this justified? Consider the original theory, in which $\hbar$ only appears in the explicit factor in $\exp -i\mathcal{S}/\hbar$, and look at a graph with P propagators, V vertices and E external legs. Euler's formula tells us that $L = P - V + 1$ [6]. Since $\hbar$ enters the path integral through $\exp(iS/\hbar)$, every propagator brings a power of $\hbar$ and every vertex brings a power of $\hbar^{-1}$; it follows that a graph is proportional to $\hbar^{P-V} = \hbar^{L-1}$. Since the graph is providing a vertex in $S_{\rm eft}/\hbar$, the L-loop matching involves contributions to $\mathcal{L}_{\rm eft}$ at $O(\hbar^L)$. An $\hbar$ expansion is always justified when a perturbative expansion is justified. To see that, consider a graph in a theory with a single type of interaction vertex involving n fields. In this case we have $(E+2P) = nV$, since one end of every external line and two ends of every internal line must end on a vertex, and there must be n lines coming in to each vertex. Putting this together with Euler's equation we have $V = (2L+E-2)/(n-2)$, which shows that for a given number of external lines, the number of vertices (and hence the power of the coupling constant) grows with the number of loops, so a loop expansion (or equivalently, an $\hbar$ expansion) is justified if a perturbative expansion is justified. This can be generalized to a theory with several types of vertices.

So we match the UV theory eq. (91) to the EFT order by order in an $\hbar$ expansion with

$$\mathcal{L}_{\rm eft} = \mathcal{L}^0_{\rm eft} + \hbar\mathcal{L}^1_{\rm eft} + \hbar^2\mathcal{L}^2_{\rm eft} + \ldots \qquad (93)$$

What makes this interesting is that we start introducing powers of $\hbar$ into the coupling constants of the EFT, so in the EFT the powers of $\hbar$ in a graph can be higher than the number of loops; the example below should make this clear. Since the EFT is expressed in terms of local operators, the matching also involves performing an expansion in powers of external momenta, with a contribution at p^n matching onto an n-derivative operator. We only match amplitudes involving light particles on external legs.

Tree level matching. At $\hbar^0$ we have to match the two theories at tree level. There are an infinite number of tree level graphs one can write down in the full theory, but the only ones we have to match are those that do not fall apart when I cut a light particle propagator...these I will call "1LPI" diagrams, for "1 Light Particle Irreducible". The other graphs will be automatically accounted for in the EFT by connecting the vertices with light particle propagators. That means we can fully determine the EFT by computing the three tree diagrams of the UV theory on the left side of Fig. 7. Because we are doing a momentum expansion, these will determine an infinite number of operator coefficients in the EFT. To compute the 4-point vertices in the EFT at this order, we equate the graphs shown in Fig. 7. Here I do not compute the graphs, but just indicate their general size,

[6] One way to derive this is to think of some Green function represented by a Feynman graph: it involves L integrations over loop momenta and one overall momentum conserving δ-function, $\delta^4(p_{\rm tot})$; on the other hand it also equals one momentum integration for each propagator and a momentum conserving δ-function at each vertex. From that observation one finds $(\int d^4p)^L\delta^4(p) \sim (\int d^4p)^P(\delta^4(p))^V$ or $(L-1) = (P-V)$.

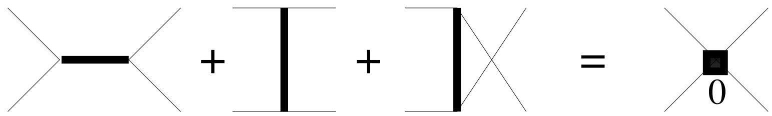

Figure 7: *Matching at $O(\hbar^0)$ between the UV theory and the EFT: on the left, integrating out the heavy scalar S (dark propagator); on the right, all contributions of four-point vertices the tree level EFT $\mathcal{L}^0_{eft}$. Equating the two sides allows on to solve for this vertices.*

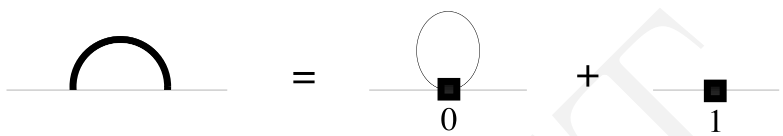

Figure 8: *Matching the 2-point function in the EFT at $O(\hbar)$. On the left, the 1-loop 1LPI graph contributing in the full theory, and on the right, graphs from the EFT include 1-loop graphs involving the 4-point vertices from $\mathcal{L}^0_{eft}$, as well as $O(\hbar)$ tree-level contributions from ϕ^2 operators in $\mathcal{L}^1_{eft}$, including the mass and kinetic term, as well as the infinite number of operators induced at this order with more derivatives. When working to a given order in a low momentum expansion, one does not need to compute all of these higher derivative operator coefficients.*

with the result

$$\mathcal{L}^0_{\rm eft} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - c_0\frac{\kappa^2}{M^2}\frac{\phi^4}{4!} - d_0\frac{\kappa^2}{M^4}\frac{(\partial\phi)^2\phi^2}{4} + \dots \ , \tag{94}$$

where c_0, d_0 etc. are going to be $O(1)$ dimensionless numbers and the ellipses refers to operators with four powers of ϕ and more powers of derivatives. The subscript 0 indicates that these coupling constants are $O(\hbar^0)$. The factors of κ^2 comes from the two vertices on the LHS of Fig. 7, and expanding the heavy scalar propagator in powers of the light field's momentum gives terms of the form $(p^2/M^2)^n \times 1/M^2$.

One loop matching. At $O(\hbar^1)$ we have to compute all 1-loop 1PLI graphs in the UV theory with arbitrary numbers of external legs, in a Taylor expansion in all powers of external momenta, and equate the result to all diagrams in the EFT that are order $\hbar$; the latter include (i) all 1-loop diagrams from $\mathcal{L}^0_{\rm eft}$ (since its couplings are $O(\hbar^0)$ and a loop brings in a power of $\hbar$), plus (ii) all tree diagrams from $\mathcal{L}^1_{\rm eft}$, since the couplings of $\mathcal{L}^1_{\rm eft}$ are $O(\hbar)$. The matching conditions for the two-point functions are shown in Fig. 8, and those for the four-point functions are shown in Fig. 9. All loop diagrams are most easily renormalized using the $\overline{MS}$ with renormalization scale set to the matching scale, e.g. $\mu = M$, so that the $\ln M^2/\mu^2$ terms that will arise vanish. The result one will find is

$$\hbar\mathcal{L}^1_{\rm eft} \quad = \quad \frac{1}{2}\left(a_1\frac{\kappa^2}{16\pi^2M^2}\right)(\partial\phi)^2 - \frac{1}{2}\left(b_1\frac{\kappa^2}{16\pi^2} + b_1'\frac{m^2}{16\pi^2}\right)\phi^2$$

$$-c_1\left(\frac{\kappa^4}{16\pi^2 M^4}\right)\frac{\phi^4}{4!} - d_1\left(\frac{\kappa^4}{16\pi^2 M^6}\right)\frac{(\partial\phi)^2\phi^2}{4} + \ldots \tag{95}$$

where the coefficients a_1, b_1, b_1', c_1, d_1 are going to be $\mathcal{O}(\hbar)$. In the above expression the b_1 term arises from the loop on the left of the equal sign in Fig. 8, while the b_1' term arises from the loop to the right of the equal sign. In addition at this order there are higher n-point vertices generated in the EFT, such as ϕ^6, $(\phi\partial^2\phi)^2$, etc. This Lagrangian can be used to compute $2\phi \to 2\phi$ scattering up to 1 loop. One can perform an a_1-dependent rescaling of the ϕ field to return to a conventionally normalized kinetic term.

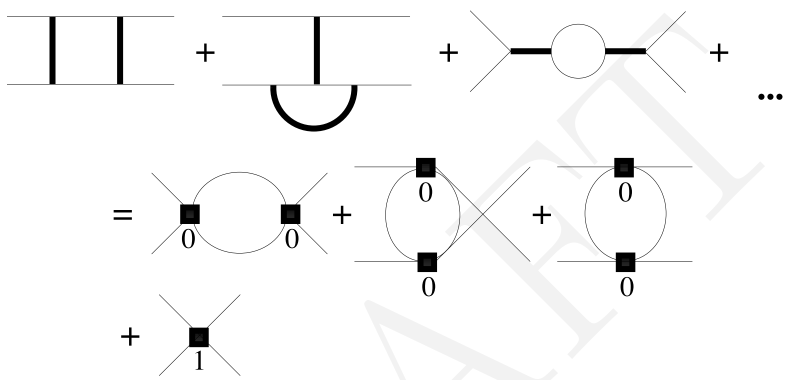

Figure 9: *Matching the 4-point function in the EFT at $O(\hbar)$. On the left, the 1LPI graphs in the full theory (with the ellipsis indicating other topologies), and on the right the $O(\hbar)$ contribution from the EFT, including 1-loop graphs involving the 4-point vertices from $\mathcal{L}^0_{EFT}$ and tree level contributions from 4-point vertices in $\mathcal{L}^1_{EFT}$, which are determined from this matching condition.*

Let me close this section with several comments about the above example:

- Notice that the loop expansion is equivalent to an expansion in $(\kappa^2/16\pi^2 M^2)$. To the extent that this is a small number, perturbation theory and the loop expansion make sense.
- We see that the matching correction to the scalar mass^2 includes a term proportional to κ^2 instead of m^2, so that even if we took $m \ll \kappa$, it is "unnatural" for the physical mass to be $\ll \frac{\kappa^2}{16\pi^2}$. In fact, in order for the meson to have a very light physical mass would require a finely tuned conspiracy between m^2 and κ^2.
- The coefficients of operators in the effective field theory are regularization scheme dependent. Their values differ for different schemes, but physical predictions do not (e.g, the relative cross sections for $2\phi \to 2\phi$ at two different energies).
- In the matching conditions the graphs in both theories have pieces depending non-analytically on light particle masses and momenta (eg, $\ln m^2$ or $\ln p^2$)...these terms cancel on both sides of the matching condition so that the interactions in $\mathcal{L}_{\text{eft}}$ have a local expansion in inverse powers of $1/M$. This is an important and generic property of effective field theories.

One can now use the effective theory one has constructed to compute low energy $\phi - \phi$ scattering. If one is interested in physics at scales far below the cutoff of the EFT one might have to renormalization group improve the answer. For example, in a GUT one integrates out particles at the scale $M_{\rm GUT} \sim 10^{15}$ GeV and then computes scattering at energies around 1 GeV. Without RG improvement, the perturbative expansion will involve terms such as $\alpha(M_{\rm GUT}) \ln 10^{15} \sim 0.6$. So one eliminates the log by computing in terms of $\alpha(1 \text{ GeV})$. To do that one just needs to compute the β functions in the EFT, ignoring the heavy particles integrated out at $M_{\rm GUT}$. This procedure requires using a mass-independent renormalization scheme so one can solve the RG equations. Now since both the proton and top quark exist in the EFT, one risks getting terms such as $\alpha(1 \text{ GeV}) \ln m_t/m_p$, for example, which also isn't so good if α is the strong coupling. So one can construct a ladder of EFTs, integrating out GUT scale particles at M_{GUT}, the top quark at the scale m_t, the W and Z at their mass scale, etc. At each stage in this ladder of EFTs one uses the β function appropriate for the light particles that remain as explicit degrees of freedom in the EFT.

Matching computations like this are used for predicting the low energy gauge couplings in the SM as predicted by Grand Unified Theories (GUTS), integrating out the heavy particles at the GUT scale $M_{\rm GUT}$ and matching onto the SM as the EFT. At tree level matching, the gauge couplings in the EFT at the scale $\mu = M_{\rm GUT}$ are equal (when suitably normalizing the $U(1)$ coupling), and then one runs them down to low energy, each gauge coupling running in the SM with its own 1-loop β-function. This is the classic calculation of Georgi, Quinn and Weinberg [3] and can be used to predict $\alpha_s(M_Z)$, since the input are two unknowns (the scale $M_{\rm GUT}$ and the GUT gauge coupling $g(\mu)$ at $\mu = M_{\rm GUT}$) while the output are the three parameter of the SM α, $\alpha_s(M_Z)$, and $\sin^2\theta_w$. However, if you want greater precision you must match the GUT to the EFT at one loop, which generates small and unequal shifts in the SM gauge coupling at $\mu = M_{\rm GUT}$, and then one scales them down using the 2-loop β-functions.

3.2 Relevant operators and naturalness

We have seen that a scalar field mass typically gets large additive quantum corrections, so that in a theory with new physics at scales much larger than the weak scale (e.g. any GUT theory, and probably any theory with gravity!) it seems unnatural to have light scalars in the low energy effective theory. This would seem to be a potential problem for the SM, where we know there exists a a Higgs with a weak scale mass, far below the Planck and GUT scales. A natural size for the Higgs mass^2 would be $\frac{\alpha}{4\pi}\Lambda^2$ where Λ is the cutoff of the ET we call the SM. But what is Λ? We have seen that for proton decay and lepton violation we expect $\Lambda \gtrsim 10^{15}$ GeV; if that is the appropriate Λ to use in the Higgs mass estimate, then the SM must be fine-tuned to ~ 13 orders of magnitude! In fact, the Higgs mass appears not to be fine-tuned only if $\Lambda \sim 1$ TeV – which is one reason to think the LHC might still discover new physics.

Before thinking about what could explain the light Higgs mass, it is interesting to ask whether there are other relevant operaotrs in the SM. Fermion masses in the SM arise from dimension 4 operators (Yukawa interactions) and are hence marginal above the weak scale. You might think they are relevant below the weak scale, appearing as dimension 3 operators $m\bar\psi\psi$ – but that is misleading. Since a fermion mass term breaks the chiral symmetry $\psi \to e^{i\theta\gamma_5}\psi$, a fermion mass can only be renormalized multiplicatively: any divergent diagram must contain chiral symmetry breaking, and hence must be proportional to m. Thus renormalizations of the fermion mass are not additive: ($\delta m \sim \alpha\Lambda$) but multiplicative: ($\delta m \sim m \ln \Lambda$). Thus light fermion masses are natural since a logarithmic dependence on the cutoff is very weak and the cutoff may be exponentially larger than the fermion mass without requiring fine tuning. Similarly, since gauge boson masses break gauge symmetry, they too do not receive additive mass contributions. However there is one other relevant operator which is problematic: the operator 1, the vacuum energy or cosmological constant. This operator's coefficient receives additive contributions proportional to Λ^4, while we know that the cosmological constant needs to be $\sim (10^{-3}\text{ eV})^4$ to be consistent with cosmological observations – a scale much lower than many known physical scales. A number of new TeV physics scenarios have been proposed to solve or partially solve the Higgs fine-tuning problem: technicolor, supersymmetry, composite Higgs, extra dimensions... however no one has found a dynamical theory to explain the small cosmological constant.

It could be that new physics is around the corner, but one has to wonder whether the naturalness argument isn't missing something, especially because we see a small cosmological constant, and we see a light Higgs mass, but we don't see a host of not-very-irrelevant higher dimension operators that one might expect to be generated by new physics, and whose effects we would expect to have seen already if its scale were low. Various ideas have been suggested for alternatives to naturalness. One popular one is the anthropic principle, the idea that there are many places in the universe with different parameters, most of which are "natural" but in which life is impossible. Therefore we exist in those very peculiar fine-tuned places where life *is* possible and we shouldn't worry that it looks like a bizarre world. To make these arguments sensible you have to (i) have a UV theory for the possible values and correlations between parameters (for example: is it possible to find a place where the up quark is heavy and the down quark is light, or do they have to scale together?), as well as a sensible theory for the a priori probability distribution that they take; then (ii) one has to have a good understanding about how these parameters affect our existence. I have only seen two examples where these two criteria are met at all: anthropic arguments for the cosmological constant [4], and anthropic arguments for the axion in an inflationary universe (see [5] and references therein). And in fact, the anthropic solution is the only plausible explanation for the small cosmological constant that has been proposed. I will return to the anthropic axion in my last lecture.

Another idea is that the world is fine-tuned because of its dynamical evolution. Whereas a creationist might say that human eye is a miraculous finely-tuned apparatus, the evolutionist says it is the product of a billion years of evolution. Perhaps the light Higgs is due to some dynamical path that the universe has taken since the Big Bang? A very creative idea along these lines recently appeared in [6], again involving the axion. See also the parable I have reprinted below (but will not include in my lecture) from my 1997 TASI lectures; I wrote that after being at a conference where I thought that SUSY advocates were being way too smug about the theory!

3.3 Aside – a parable from TASI 1997

I used to live in San Diego near a beach that had high cliffs beside it. The cliffs were composed of compressed sand, and sand was always sprinkling down to the beach below.

At the base of these cliffs there was always a little ramp of sand. One day I was walking down the beach with a physicist friend of mine, and she remarked on the fact that each of these ramps of sand was at precisely the same angle.

"How peculiar!" she said, and I had to agree, but thought no further of it. However, she had a more inquiring mind than I, and called me up that evening:

"I've been conducting an experiment," she said. "I take a box of sand and tilt it until it avalanches, which occurs at an angle θ_c. You won't believe it, but θ_c is precisely the same angle as the ramps of sand we saw at the beach! Isn't that amazing?"

"Indeed," I said. "Apparently someone has performed the same experiment as you, and has sent someone out to adjust all the sand piles to that interesting angle. Perhaps it's the Master's project of some Fine Arts student."

"That's absurd!" she said. "I could believe it if the artist did this *once*, but just think — the wind is always blowing sand onto some piles and off of other piles! The artist would have to fix the piles continuously, day and night!"

As unlikely as this sounded, I had to insist that there seemed to be no alternative. However, the next morning when we met again, my friend was jubilant.

"I figured out what is going on!" she exclaimed. "I have deduced the existence of *swind*. Every time the wind blows and moves the sand, swind blows and moves it back! I call this 'Swindle Theory'. "

"That's absurd!" I exclaimed. "Wind is made of moving molecules, what in the world is swind made of, and why haven't we seen it?"

"Smallecules," she replied, "they're too small to see."

"But you still need the art student to come by in the beginning and fix the sand at precisely the angle θ_c, right?" I asked.

"True"

"So why should I believe in Swindle Theory? You've hardly explained anything!"

"Well look," she retorted, "see how beautiful the Navier-Stokes equations become when generalized to include swind?"

Indeed, the equations were beautiful, so beautiful that I felt compelled to believe in Swindle Theory, although I occasionally still have my doubts...

3.4 Landau liquid versus BCS instability

We have discussed irrelevant operators and rare symmetry violating processes; relevant operators and the naturalness problem; let us now turn to the fascinating case of asymptotically free marginal interactions.

In our discussion of 2D quantum mechanics we encountered asymptotic freedom and the dynamical generation of an exponentially small scale in the IR. This is a possibility in theories with marginal interactions that are pushed into relevancy by small radiative corrections; however it is known that for relativistic QFTs it only actually happens in nonabelian gauge theories – the most famous example being QCD. Asymptotic freedom explains why $\Lambda_{\rm QCD}$ is naturally so much smaller than the GUT or Planck scales. However, the same physics is responsible for the large Cooper pairs found in superconducting materials, which I describe here, following the work of Polchinski [7]. I like this example because it emphasizes that you should not have a fixed idea what an EFT has to look like, but should be able to adapt its use to widely different theories.

A condensed matter system can be a very complicated environment; there may be various types of ions arranged in some crystalline array, where each ion has a complicated electron shell structure and interactions with neighboring ions that allow electrons to wander around the lattice. Nevertheless, the low energy excitation spectrum for many diverse systems can be described pretty well as a "Landau liquid", whose excitations are fermions with a possibly complicated dispersion relation but no interactions. Why this is the case can be simply understood in terms of effective field theories, modifying the scaling arguments to account for the existence of the Fermi surface.

Let us assume that the low energy spectrum of the condensed matter system has fermionic excitations with arbitrary interactions above a Fermi surface characterized by the fermi energy ϵ_F ; call them "quasi-particles". Ignoring interactions at first, the action can be written as

$$S_{free} = \int dt \int d^3p \sum_{s=\pm\frac{1}{2}} \Big[\psi_s(p)^\dagger i\partial_t\psi_s(p) - (\epsilon(p) - \epsilon_F)\psi_s^\dagger(p)\psi_s(p)\Big] \tag{96}$$

where an arbitrary dispersion relation $\epsilon(p)$ has been assumed.

To understand how important interactions are, we wish to repeat some momentum space version of the scaling arguments I introduced in the first lecture. In the present case, a low energy excitation corresponds to one for which $(\epsilon(p) - \epsilon_F)$ is small, which means that $\mathbf{p}$ must lie near the Fermi surface. So in momentum space, we will want our scaling variable to vary the distance we sit from the Fermi surface, and not to rescale the overall momentum $\mathbf{p}$. After all, here a particle with $\mathbf{p} = 0$ is a high energy excitation.

This situation is a bit reminiscent of HQET where we wrote $p_\mu = mv_\mu + k_\mu$, with k_μ being variable that is scaled, measuring the "off-shellness" of the heavy quark. So in the present case we will write the momentum as

$$\mathbf{p} = \mathbf{k} + \boldsymbol{\ell} \tag{97}$$

where $\mathbf{k}$ lies on the Fermi surface and $\boldsymbol{\ell}$ is perpendicular to the Fermi surface (shown in Fig. 10 for a spherical Fermi surface). Then $\boldsymbol{\ell}$ is the quantity we vary in experiments and so we define the dimension of operators by how they must scale so that the theory is unchanged when we change $\boldsymbol{\ell} \to r\boldsymbol{\ell}$. If an object scales as r^n, then we say it has dimension n. Then $[k] = 0$, $[\ell] = 1$, and $[\int d^3p = \int d^2k d\ell] = 1$. And if we define the Fermi velocity as $\mathbf{v}_F(\mathbf{k}) = \boldsymbol{\nabla}_k\epsilon(\mathbf{k})$, then for $\ell \ll k$,

$$\epsilon(\mathbf{p}) - \epsilon_F = \boldsymbol{\ell}\cdot\mathbf{v}_F(\mathbf{k}) + \mathcal{O}(\ell^2)\ , \tag{98}$$

and so $[\epsilon - \epsilon_f] = 1$ and $[\partial_t] = 1$. Given that the action eq. (96) isn't supposed to change under this scaling,

$$[\psi] = -\frac{1}{2}\ . \tag{99}$$

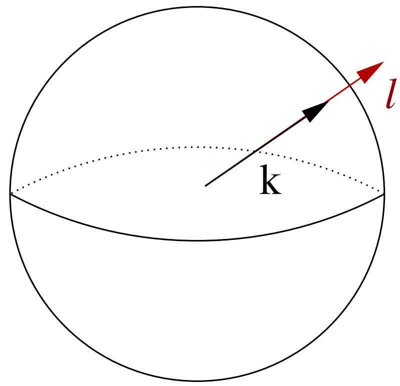

Figure 10: *The momentum* $\mathbf{p}$ *of an excitation is decomposed as* $\mathbf{p} = \mathbf{k} + \boldsymbol{\ell}$*, where* $\mathbf{k}$ *lies on the Fermi surface (which does not have to be a sphere), and* $\boldsymbol{\ell}$ *is perpendicular to the Fermi surface. Small* $|\boldsymbol{\ell}|$ *corresponds to a small excitation energy.*

Now consider an interaction of the form

$$S_{int} = \int dt \int \prod_{i=1}^{4} (d^2\mathbf{k}_i d\ell_i) \delta^3(\mathbf{P}_{tot}) C(\mathbf{k}_1, \ldots, \mathbf{k}_4) \psi_s^\dagger(\mathbf{p}_1)\psi_s(\mathbf{p}_2)\psi_{s'}^\dagger(\mathbf{p}_3)\psi_{s'}(\mathbf{p}_4) \ . \qquad (100)$$

This will be relevant, marginal or irrelevant depending on the dimension of C. Apparently we have the scaling dimension $[\delta^3(\mathbf{P}_{tot})C] = -1$. So how does the δ function by itself scale? For generic $\mathbf{k}$ vectors, $\delta(\mathbf{P}_{tot})$ is a constraint on the $\mathbf{k}$ vectors that doesn't change much as one changes ℓ, so that $[\delta^3(\mathbf{P}_{tot})] = 0$. It follows that $[C] = -1$ and that the four fermion interaction is irrelevant...and that the system is adequately described in terms of free fermions (with an arbitrary dispersion relation). This is why Landau liquid theory works and is related to why in nuclear physics Pauli blocking allows a strongly interacting system of nucleons to have single particle excitations.

This is not the whole story though, or else superconductivity would never occur. Let us look more closely at the conclusion above $[\delta^3(\mathbf{P}_{tot})] = 0$. Consider the case when all the $\boldsymbol{\ell}_i = 0$, and therefore the $\mathbf{p}_i = \mathbf{k}_i$ and lie on the Fermi surface. Suppose we fix the two incoming momenta $\mathbf{k}_1$ and $\mathbf{k}_2$. The $\delta^3(\mathbf{P}_{tot})$ then constrains the sum $\mathbf{k}_3 + \mathbf{k}_4$ to equal $\mathbf{k}_1 + \mathbf{k}_2$, which generically means that the vectors $\mathbf{k}_3$ and $\mathbf{k}_4$ are constrained up to point to opposite points on a circle that lies on the Fermi surface (Fig. 11b). Thus one free parameter remains out of the four independent parameters needed to describe the vectors $\mathbf{p}_3$ and $\mathbf{p}_4$. So we see that in this generic case, $\delta^3(\mathbf{P}_{tot})$ offers three constraints, even when $\boldsymbol{\ell}_i = 0$. Therefore $\delta^3(\mathbf{P}_{tot}) = \delta^3(\mathbf{K}_{tot})$ is unaffected when $\boldsymbol{\ell}$ is scaled, and we find the above assumption $[\delta^3(\mathbf{P}_{tot})] = 0$ to be true, and Landau liquid theory is justified.

However now look at the special case when the collisions of the incoming particles are nearly head-on, $\mathbf{k}_1 + \mathbf{k}_2 = 0$. Now $\delta^3(\mathbf{P}_{tot})$ constrains the outgoing momenta to satisfy $\mathbf{k}_3 + \mathbf{k}_4 = 0$. But as seen in Fig. 11a, this only constrains $\mathbf{k}_3$ and $\mathbf{k}_4$ to lie on opposite sides of the Fermi surface. Thus $\delta^3(\mathbf{P}_{tot})$ seems to be only constraining two degrees of freedom, and could be written as $\delta^2(\mathbf{k}_3 + \mathbf{k}_4)\delta(0)$. This singularity obviously arose because the set the $\boldsymbol{\ell}_i = 0$, and so $\delta^3(\mathbf{P}_{tot})$ must be scaling as an inverse power of $\boldsymbol{\ell}$. For nonzero $\boldsymbol{\ell}$ the $\delta(0)$

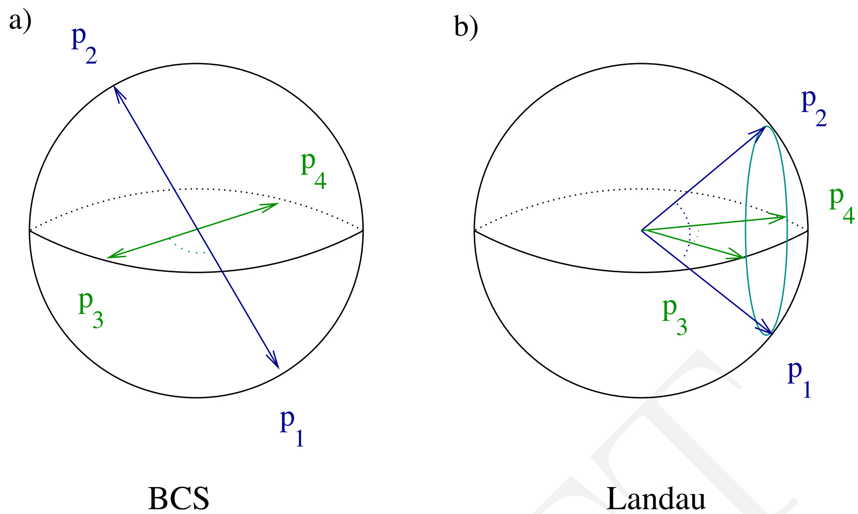

Figure 11: *Fermions scattering near the Fermi surface. (a) Head-on collisions: With* $\mathbf{k}_1 + \mathbf{k}_2 = 0$*, only two degrees of freedom in the outgoing momenta* $\mathbf{k}_3$ *and* $\mathbf{k}_4$ *are constrained, as they can point to any two opposite points on the Fermi surface. (b) The generic Landau liquid case, where the incoming particles do not collide head-on, and three degrees of freedom in the outgoing momenta* $\mathbf{k}_3$ *and* $\mathbf{k}_4$ *are constrained, as they must point to opposite sides of a particular circle on the Fermi surface. Figure from ref. [8], courtesy of Thomas Schäfer.*

becomes $\delta(\boldsymbol{\ell}_{tot})$, and as a result, the δ function scales with $\boldsymbol{\ell}^{-1}$: $[\delta^3(\mathbf{P}_{tot})] = -1$. But since $[\delta^3(\mathbf{P}_{tot})C] = -1$, it follows that for these head-on collisions we must have $[C] = 0$, and the interaction is marginal!

We have already seen that quantum corrections make a marginal interaction either irrelevant or relevant; it turns out that for an attractive interaction, the interaction becomes relevant, and for a repulsive interaction, it becomes irrelevant, just as we found for the δ-function interaction in two dimensions.

Therefore, an attractive contact interaction between quasiparticles becomes strong exponentially close to the Fermi surface (since the coupling runs logarithmically), and can lead to pairing and superconductivity just as the asymptotically free QCD coupling leads to quark condensation and chiral symmetry breaking. The BCS variational calculation shows that the pairing instability does indeed occur; the effective field theory analysis explains why Cooper pairs are exponentially large compared to the lattice spacing in superconductors. The difference between superconductors and metals that behave as Landau liquids depends on the competition between Coulomb repulsion and phonon mediated attraction in the particular material, which determines the sign of the C coupling.

3.5 Problems for lecture III

III.1) A small fermion mass can be considered natural, in contrast to a small scalar mass. This has to do with the fact that if a fermion becomes massless, usually the symmetry of the theory is enhanced by a $U(1)$ chiral symmetry $\psi \to e^{i\alpha\gamma_5}\psi$. Thus at $m = 0$, there cannot be any renormalization of the fermion mass. A corollary is that at nonzero mass m, any renormalization must be proportional to m. Can you explain why this makes the fermion mass behave like a marginal operator rather than a relevant one? Can you construct an example of a theory where it is *not* natural to have a light fermion?

References

[1] D. B. Kaplan, *Flavor at SSC energies: A New mechanism for dynamically generated fermion masses*, *Nucl.Phys.* **B365** (1991) 259–278.

[2] K. Agashe, R. Contino, and A. Pomarol, *The Minimal composite Higgs model*, *Nucl.Phys.* **B719** (2005) 165–187, [hep-ph/0412089].

[3] H. Georgi, H. R. Quinn, and S. Weinberg, *Hierarchy of Interactions in Unified Gauge Theories*, *Phys.Rev.Lett.* **33** (1974) 451–454.

[4] S. Weinberg, *Anthropic Bound on the Cosmological Constant*, *Phys.Rev.Lett.* **59** (1987) 2607.

[5] D. B. Kaplan and A. E. Nelson, *Inflationary Axion Cosmology Beyond Our Horizon*, arXiv:0809.1206.

[6] P. W. Graham, D. E. Kaplan, and S. Rajendran, *Cosmological Relaxation of the Electroweak Scale*, *Phys. Rev. Lett.* **115** (2015), no. 22 221801, [arXiv:1504.07551].

[7] J. Polchinski, *Effective Field Theory and the Fermi Surface*, hep-th/9210046.

[8] T. Schaefer, *Quark Matter*, hep-ph/0304281.