Typeset using JHEP styles

Regularization and Renormalization
A Pedagogical (Re-)Introduction

Carl Turner

Department for Applied Maths and Theoretical Physics,
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge, CB3 OWA, UK

E-mail: courses@suchideas.com, cpt39@cam.ac.uk

Abstract: A note on renormalization from a different perspective to how it is usually introduced. See http://suchideas.com/courses/asides/ for updates! Please send corrections, suggestions and notes to courses@suchideas.com The author's homepage for all courses may be found on his website at SuchIdeas.com. **These notes have not yet been proof-read.**

Contents

1 Introduction

In this note, we explore the concepts of *regularization* and *renormalization*, aiming to connect what might initially seem a very arbitrary and unmathematical trick to 'cancel infinities' against each other with some perfectly reasonable, simple examples of essentially the same things at work. We'll then try to justify the procedures of Quantum Field Theory (QFT) with the ideas of the Wilsonian renormalization and the *Exact Renormalization Group* (ERG).

The core ideas we want to focus on are as follows:

- When we construct most quantum field theories, like the Standard Model, we are really constructing an *Effective Field Theory* (EFT). An EFT is a prescription for making 'low energy' predictions based on a finite number of observations to fit parameters in your model. 'Low energy' here means with respect to the scale at which the model breaks down. The Planck scale is certainly an upper bound for the regime where the Standard Model is likely to be valid.

- *Regularization* is the process of taking a first, naïve approximation to what a low energy theory might look like - which in fact diverges, or more generally fails to encode the interesting physics in the limit of interest - and adding an extra parameter which cuts off the divergence (or similar) to see exactly how and why the theory diverges.

- *Renormalization* is the process of taking a (regularized if necessary) theory and finding the correct set of physical parameters to describe the low energy theory, and then taking the low energy limit again.

- Until we genuinely claim to have a unified theory of everything (and there is no way the Standard Model is such a theory), we should always view our models as approximations, and fix parameters to observables at energy scales actually accessible to us.

We will begin by seeing examples of regularization and renormalization in classical and simple quantum contexts, before moving on to discuss how to visualize renormalization in the bog standard QFT example of ϕ^4 theory.

See also:

- Polchinski on effective Lagrangians, http://www.sciencedirect.com/science/article/pii/0550321384902876, and also effective field theory, http://arxiv.org/abs/hep-th/9210046.

- Sonoda on Wilsonian renormalization in quantum perturbation theory, http://arxiv.org/abs/hep-th/0603151.

- Delamotte on a 'hint of renormalization', http://arxiv.org/abs/hep-th/0212049.

- Olness and Scalise on the freshman EM example, http://arxiv.org/abs/0812.3578.

- Neumaier on 'renormalization without infinities', http://www.osti.gov/eprints/topicpages/documents/record/312/1635173.html.

- Luty on renormalization in QFT, http://www.physics.umd.edu/courses/Phys851/Luty/notes.html.

2 Prototypes

The examples we look at in this section are as follows:

- The 'freshman EM' example of regularization. Here we will see a first illustration of the principles enshrined in the introduction in the context of calculating the potential due to a wire. The aim is to make clear the difference between interesting divergences and uninteresting divergences.
- We will then look at two classic QM problems to illustrate how divergences can arise from claiming knowledge of the behaviour of all momentum modes incorrectly. Firstly, we look at low-momentum scattering off an odd potential localized at the origin; secondly, we look at low-energy s-wave scattering off a pole at the origin.
- Finally, we will look at an example of classical field theory (the 1D Ising model) which illustrates 'renormalization without infinities' to emphasize that renormalization is a concept which exists independently of the need for regularization.

2.1 Freshman EM

The Problem - Statement 1 Suppose we have an infinitely long wire lying along the y-axis, with uniform charge per unit length λ. What is the potential a distance r from the wire?

This is of the form of a problem undergraduate students are typically taught to solve in their first course on electromagnetism. The problem is essentially Poisson's equation

$$\nabla^2\phi = -\frac{\rho}{4\pi\epsilon_0}$$

which has a standard solution in terms of Green's functions

$$\phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{x}')\,\mathrm{d}^3\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|}$$

so that a distance r from the wire one finds

$$\phi(r) = \frac{\lambda}{4\pi\epsilon_0}\int_{-\infty}^{\infty}\frac{dy}{\sqrt{r^2+y^2}} \tag{2.1}$$

However, this is a problematic integral. As $y\to\infty$ we see that the integrand goes as $1/y$ in both directions. Since $\int^L \mathrm{d}y/y \sim \log L$ we see that we find

$$\boxed{\phi = \infty}$$

... which isn't so useful for computing the electric field $\mathbf{E} = -\boldsymbol{\nabla}\phi$ for example.

One might simply hope that the problem only arises because there is 'too much energy' in the electrostatic problem as stated. But that transpires to be a rather odd objection, since one can find this physical setup as a standard example in textbooks, with $\mathbf{E}\propto\hat{\mathbf{r}}/r$ scaling as $1/r$ as r increases. (The computation is done by considering the integral form of the equation $\boldsymbol{\nabla}\cdot\mathbf{E}\propto\rho$ with a surface given by a cylinder surrounding the wire.)

Something we might notice when we see $|\mathbf{E}|\propto 1/r$ is that it suggests what ϕ should look like. Since $\mathbf{E}$ is a derivative of ϕ is is obvious that in fact there should be an answer $\phi\propto\log r$. But now we are led to an interesting observation; we aren't allowed to take the log of a dimensionful quantity.[1]

[1] Why? Well, what units should it have? $\log x = \sum(-x)^n/n$ clearly does not have well-defined units! Looked at another way, $\log(1\,\text{metre}) = \log(100\,\text{centimetres}) = \log 100 + \log(1\,\text{centimetre})$ looks a little odd.

(Obviously r has dimensions of length here.) There is a simple remedy we could implement, namely to introduce some arbitrary r_0 and take $\log(r/r_0)$, which is totally legitimate. However, where could this r_0 come from in our attempt to solve the problem? We are short of a length scale!

More carefully, let us take

$$v(r) = 4\pi\epsilon_0/\lambda \times \phi(r), \qquad e(r) = -\partial_r v(r)$$

Then v is dimensionless, e has dimensions of inverse length, and r has dimensions of length. Consequently, dimensional analysis tells us that $e(r) \propto 1/r$ but that $v(r)$ cannot depend on r!

Remark. In some sense, this shows that finding $v = \infty$ was completely inevitable. Why? If v is constant in r, but $e \neq 0$ so that $\partial_r v \neq 0$, then the only function which can be changed by a finite amount and remain constant is... ∞.

This shows that our attempt to solve this problem was in some ways doomed to failure. We tried to take a sensible limit by going straight to the limiting case without thinking about how we made the approximation, and the reward was that we got no useful information.

Now let us acknowledge that our statement of the problem is actually a little disingenuous. When we posed this problem, it was obvious that the picture we *secretly* had was actually of a *really long* wire which we were sufficiently close to that it seemed infinitely long. We were really asking for the long-wire or near-wire approximation for the form of the potential. Let's rephrase the problem to respect this.

The Problem - Statement 2 (Regularized) Suppose we have a wire of length Λ lying along the y-axis with uniform charge per unit length λ. Find an expression for the potential a distance r from the wire valid in the large Λ limit.

Okay - now the analogous expression to (2.1) can be computed exactly:

$$v(r;\Lambda) = \int_{-\Lambda}^{\Lambda} \frac{dy}{\sqrt{r^2+y^2}} = \log\left(\frac{+\Lambda + \sqrt{\Lambda^2+r^2}}{-\Lambda + \sqrt{\Lambda^2+r^2}}\right)$$

Before we start thinking about expanding this for large Λ, we note that firstly in fact

$$v = \log\left(\frac{+1 + \sqrt{1+(r/\Lambda)^2}}{-1 + \sqrt{1+(r/\Lambda)^2}}\right) = f(r/\Lambda)$$

as was necessary on the grounds of dimensional analysis. Also, we note that physical quantities like the potential difference

$$\delta v(r_1, r_2; \Lambda) = v(r_2;\Lambda) - v(r_1;\Lambda) = 2\log\frac{r_1}{r_2} \quad + \mathcal{O}\left(r_i^2/\Lambda^2\right)$$

and so on are now rendered finite, and indeed have well-defined $\Lambda \to \infty$ limits.

So what about the form of the potential in this limit?

$$v(r;\Lambda) = -2\log\frac{r}{\Lambda} + \mathcal{O}\left(r^2/\Lambda^2\right)$$

Now we see clearly that

- v does indeed have the predicted logarithmic dependence on r.

- Λ provides the 'missing length scale' we wanted.

How could we try to take the large Λ limit sensibly here? Since we have seen the problem is a missing length scale - and the one we have introduced will unhelpfully go to infinity in this limit - we instead define $\Lambda = \mu/\epsilon$ for some fixed $\mu > 0$ and $\epsilon \ll 1$. Then we find

$$v(r;\mu,\epsilon) = \underbrace{-2\log\epsilon}_{\text{Divergent... but irrelevant!}} - 2\log\frac{r}{\mu} + \mathcal{O}\left(\epsilon^2\right)$$

We have now divided the effects of the Λ parameter into two parts. The large magnitude of Λ relative to where we make the observation is encoded in the parameter ϵ, whilst the presence of a length-scale somewhere in the problem is enshrined in the mysterious parameter μ.

The advantage here is that we see that in fact the divergence due to $\log\epsilon$ is irrelevant in the sense that the expression for v is physically equivalent to one without the divergent term. This is because constants in v have no physical significance. Put another way,

$$v_{\lim}(r;\mu) = \lim_{\epsilon\to 0}\left(v(r;\mu,\epsilon) + 2\log\epsilon\right) = -2\log\frac{r}{\mu}$$

is a physically equivalent potential defined in the large Λ limit.

We could rephrase this in various other ways too. For example, we could say that we make observations of $v(r;\Lambda)$ relative to some fixed $r = \mu$ and adopt the convention that $v(r=\mu;\Lambda) = 0$ always holds for all Λ. (This particular choice is permissible due to the irrelevance of constants in v and is *not* a feature we will see in later examples. Later, we will replace this 0 with a physically observable quantity $g(\mu)$ which we hold fixed in the large Λ limit.) This would lead to the same $v_{\lim}$ in the large Λ limit. This way of looking at the setup is most closely connected to the examples we will will see below.

We can then happily use this limiting form of v to calculate physical observables. 'But wait!' I hear you cry. 'How do I know what μ should be when I make predictions?' The answer is: it can't matter. *Any* choice of μ will do, so long as it is adopted consistently through the calculation. How is this manifested here? Because constants in v are irrelevant, and μ appears only in a logarithm, any rescaling of μ has no effect. Alternatively, all physical quantities depend on derivatives of or differences in v, so the μ will always drop out of the final answer.

Take Home Message We have described ways of reaching the large Λ limit and obtaining useful answers by *changing the way we take the limit*. We did this by

1. regularizing by adding a new parameter parametrizing how close we are to the limit;
2. spotting what it was about the way we took the limit that led to the divergence;
3. changing the way we used parameters in the problem to be centered on a physically sensible convention (finiteness of the potential at some fixed distance);
4. taking the limit with respect to these new conventions.

2.2 1D Potential Scattering

The Problem Suppose there is an odd potential $V(x)$ localized near the origin $x = 0$. We scatter incoming particles of momentum k at negative infinity off the potential, and ask what the transmission amplitude is for $x \to \infty$. That is, find a solution to

$$E\Psi = \left(-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right)\Psi \tag{2.2}$$

with an asymptotic form $\Psi \sim \exp(ikx) + R\exp(-ikx)$ as $x \to -\inf$ and $\Psi \sim T\exp(ikx)$ as $x \to \infty$. What is $T = T(k)$ for small k, given $V(x) = -V(-x)$?

This time, we've been careful to state the problem in terms of the limits we want to take. But first, let's imagine we make a naïve setup model. A generic (not odd) potential is usually modeled as $V(x) \sim V_0\delta(x)$. The idea is that taking long wavelengths is the same as scaling the x coordinate down, so that we 'zoom out' on the potential. This actually works just fine, and solving (2.2) with this $\sim \delta(x)$ potential is a standard introductory exercise in quantum mechanics (QM) courses. Roughly, we expect that V should only interact with the value of Ψ at 0, and the δ function ensures we only have sensitivity to $\Psi(0)$.

This heuristic picture might motivate us to model an odd potential, for which the $\sim \delta(x)$ part of the expansion of V (for an expansion of V it loosely speaking is[2]) vanishes, by $\delta'(x)$. Setting $\hbar = 1$ and multiplying through by m we find

$$\frac{1}{2}k^2\Psi = \left(-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \alpha^{-1/2}\delta'(x)\right)\Psi \tag{2.3}$$

where α is dimensionless (with the power just for later convenience) and $k = \sqrt{2mE}$.

However, attempting to solve this in the same way as the $\delta(x)$ case - by considering the jump conditions obtained by integrating (2.3) over $(-\epsilon, \epsilon)$ - leads to inconsistencies. This isn't surprising, as $\delta'(x)$ is very singular.

More intriguingly, if one attempts to solve this via a perturbation expansion, one discovers a 'UV divergence'. What do we mean by this? One can expand S-matrix elements in the interaction picture as a series in V, with 'sums' (really integrals) over orthonormal states $|p\rangle$ - if one does this, one finds a divergence coming from the high $|p|$ part of the integration.

Remark. The reader is advised to look this up elsewhere, but essentially one finds

$$\langle p_f|S|p_i\rangle = \delta(p_f - p_i) + 2\pi\delta(E_f - E_i)\left[\langle p_f|V|p_i\rangle + \int \mathrm{d}p\ \langle p_f|V|p\rangle \frac{i}{E - \frac{1}{2}p^2 + i\epsilon}\langle p|V|p_i\rangle + \cdots\right]$$

and then calculating the matrix elements $\langle p'|V|p\rangle \propto (p' - p)$ for our potential implies that the integrand is asymptotically constant in this term.

To try and fix up the problem, we move to a regularized version

$$\frac{k^2}{2}\Psi = \left(-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{\alpha^{-1/2}}{2\Delta}\left[\delta(x+\Delta) - \delta(x-\Delta)\right]\right)\Psi$$

Assuming my calculation is correct, we find that the transmission amplitude T is given by

$$T = \frac{4(k\Delta)^2}{\alpha^{-1}(1 - \exp(4i\cdot k\Delta)) + 4(k\Delta)^2} \sim i\cdot k\Delta\cdot\alpha \quad \text{for small } k\Delta \tag{2.4}$$

Here - especially looking at the full answer - we see a much more striking manifestation of the problems we have discussed before.

$|T|$ ranges from 0 to 1 according to how large k is; more accurately, T (at fixed α) is a dimensionless function of the dimensionless quantity $k\Delta$. But in the absence of such a length scale Δ, how would

[2] Consider a matrix element $\langle\chi|V|\psi\rangle = \int \chi^\star(x)\psi(x)V(x)$. Write $f = \chi^\star\psi$. Then this element $\sim \int (f(0) + xf'(0) + \cdots)V(x) \sim f(0)\int V + f'(0)\int xV + \cdots$ so that V is equivalent to $c_1\delta(x) + c_2\delta'(x) + \cdots$ where $c_1 = \int V$ and $c_2 = -\int xV$ and so on.

we tell whether k was small or not?! Clearly our δ' formulation, with its absence of a length-scale, was doomed to failure - there was no way we could hope to find a 'low energy' approximation since our problem didn't have any idea what a 'low energy' was.

Remark. As a side note, in the $\Delta \to 0$ limit, the wavefunction between the two δ functions is proportional to $\alpha^{1/2}$, which diverges in the *weak* interaction limit(!) This is another illustration of the fact that there is clearly a lot of sensitivity to the order in which limits are taken in this sort of problem.

Anyhow, we see that *the amplitude T is still dependent crucially on what we thought was a 'microscopic' or 'irrelevant' parameter of the potential* - we assumed that the above Δ would play no role, but in fact it affects the leading order term in T. It is at this point that we come to the first real example of renormalization.

We stress that the point here is to *rewrite the theory before taking the limit* in such a way that *the physical observables stay fixed* but we are left with *a simplified description of the system.* We will do this here in the conventional manner of replacing α in (2.4) with some function $\alpha(\Delta)$ which changes as we take the $\Delta \to 0$ limit so that (the leading term of) $T = T(k; \alpha(\Delta), \Delta)$ stays constant. That is, having first *introduced* an extra parameter, we *reduce* the two parameter problem with two unknowns α, Δ to a simpler system reflecting the fact that there is redundancy in our description - this is effectively a change of coordinates such that Δ becomes irrelevant (to leading order). This enables us to take the limit safely.

Concretely, we see that (2.4) gives us

$$\Delta \frac{\mathrm{d}T(k; \alpha(\Delta), \Delta)}{\mathrm{d}\Delta} = \mathcal{O}\left((k\Delta)^2\right)$$

precisely if

$$\boxed{\alpha(\Delta) = \frac{\tilde{\alpha}}{\Delta}}$$

for some new physical constant $\tilde{\alpha}$, which we observe has dimensions of *length*, unlike our original - apparently aphysical - parameter α. (Usually, the differential equation $\Delta \mathrm{d}T/\mathrm{d}\Delta$ is expanded using the chain rule to find an equation for $\Delta \mathrm{d}\alpha/\mathrm{d}\Delta$ but we can find the solution here by inspection. Had we formed this differential equation, it would have taken the form of a 'renormalization group equation' describing the 'running of the coupling' α; however, since this is the analogue of the 'bare' coupling, we might not use this language. There is no need to concern ourselves overmuch with this terminology for the moment. Focus on the concreteness of this example.)

Now in terms of our new variables, we find

$$T = ik \cdot \tilde{\alpha} + \mathcal{O}\left((k\Delta)^2\right) \to ik \cdot \tilde{\alpha} \tag{2.5}$$

in the low energy/localized potential limit. (These are of course the same limit, by dimensional analysis if nothing else!) This demonstrates a striking aspect of how badly our original parameter α described our problem; our new parameter, which in (2.5) we see captures the essential physics, has a completely different dimension! We will see how to make things look a little less odd next.

The key observation here is very physical in nature. (2.5) shows that *one observation* of T is sufficient to determine $\tilde{\alpha}$; then we know all parameters in our model, and can happily make predictions for other low energies. For better comparison with later work, let us think almost experimentally. How would we fit our parameters? We would pick an arbitrary scale μ at which to make some measurement. Suppose we observe

$$T(k = \mu) = C_\mu$$

Now we might note $T(\mu) = i \cdot \mu\Delta \cdot \alpha$ and attempt to absorb the (in the limit of interest) small quantity $\mu\Delta$ into the α-type coupling term in order to make the α-type term more relevant to the scale of physics we are actually doing. In this spirit, let

$$T(k = \mu) \equiv i \cdot \alpha_R(\mu)$$

define a quantity $\alpha_R(\mu)$ where we explicitly note that this definition depends upon the scale at which we arbitrarily chose to make a measurement.[3]

What form does T take in terms of this *new* variable?

$$T(k; \alpha_R(\mu), \mu) \sim i \cdot k\Delta \cdot \alpha = i \cdot \frac{k}{\mu}\mu\Delta \cdot \alpha = \boxed{i \cdot \frac{k}{\mu}\alpha_R(\mu)}$$

The subscript $\cdot_R$ we have smuggled in is there because α_R has essentially interpretation of a *renormalized* coupling constant.

The key idea here is that where there used to be a 'small' length scale Δ, we have instead got a more reasonable, experiment-scale length scale μ^{-1}. Similarly, we have replaced the parameter α with a (much smaller, as it happens) parameter α_R which is relevant to the scale of the interaction. Just as in the EM example, we have isolated the two aspects of the regularization process: the dependence on the new length scale has been made into a dependence on a fixed (arbitrary) scale μ, whilst the smallness of the new length scale has been smuggled away into the definition of α_R.

Note that this prescription for finding an expression for T works for *any* μ we care to choose. In fact, the formula we have found must be entirely *independent* of μ. That is, $\alpha_R(\mu)$ must vary with the scale we happened to make an observation at in precisely the right way to counteract the appearance of μ in T. Indeed, we could now go away and write down a 'renormalization group equation' telling us how the relevant physical quantity $\alpha_R(\mu)$ varies with μ from the principle that T is independent of μ, exactly as we thought of doing for the parameters Δ, α above. (It obviously amounts to $\alpha_R \propto \mu$.)

Clearly, this doesn't tell us anything we didn't already know, though it perhaps says it in a different way. The fact that α_R increases with μ (in keeping with QFT calculations, we would note $\mu\,\mathrm{d}T/\mathrm{d}\mu = 0$ which implies that the beta function $\beta(\alpha_R(\mu)) = \mu\,\mathrm{d}\alpha_R/\mathrm{d}\mu = 1$ is positive) says that the effective potential seen by higher energy particles has a larger α-type parameter, that is, a smaller magnitude. This is essentially a restatement of the fact that $T(k)$ increases with k; there is less reflection for higher-momentum particles. However, in more sophisticated problems, this is by far the easiest way of obtaining information about how a problem depends on the energy scale.

2.3 2D QM

Having essentially exhausted all there is to say about that example, let us briefly review a second quantum mechanical example to see the same concepts at play. Consider

$$\left(-\frac{1}{2}\nabla^2 + c\delta^2(x)\right)\Psi = \frac{1}{2}k^2\Psi$$

for c dimensionless. One can ask how circularly symmetric s-waves scatter off such a spiked potential; a particularly common variable which is of interest is the *phase shift* δ_0 experienced by particles incident upon such a potential. This, it transpires, also has a 'UV-divergent' nature. Accordingly, we regularize

[3] We could have arrived at this prescription by calculating $\tilde{\alpha}$ in terms of an initial condition at scale μ for the differential equation $\alpha(\Delta)$ obeys, but this presentation is arguably more natural, and more in the spirit of later QFT calculations.

using the prescription $\delta(r) \to \delta(r - \Delta)$ for the radial part of the scattering. In order to make contact with later QFT calculations, we will define

$$\Lambda \stackrel{\text{def}}{=} \frac{1}{\Delta}$$

which has the interpretation of a high energy scale rather than a short distance scale. Then one finds that with this convention

$$\cot \delta_0 (k; c, \Lambda) = -\frac{2}{c} + \frac{2}{\pi}\left(\log \frac{k}{2\Lambda} + \gamma\right) + \mathcal{O}\left(\frac{k^2}{\Lambda^2}\right) \tag{2.6}$$

This shows a common feature of QFT calculations - the existence of a 'large logarithm', very much like that encountered in the electromagnetic example. Whilst the series expansion in k/Λ is here a reassuring one, taking the logarithm of this *small* quantity gives a *large (negative)* contribution to the answer. We would like to eliminate this unpleasant expression, and also acknowledge that once more a bad decision in our parametrization of the problem - the choice of $c \cdot \delta^2(x)$ - should be fixed up by at least tweaking $c \to c(\Lambda)$ so that δ_0 remains fixed to leading order. This, once can show by differentiating (2.6), would lead to an equation

$$\Lambda \frac{\mathrm{d}}{\mathrm{d}\Lambda} \frac{1}{c(\Lambda)} = -1/\pi + \cdots$$

Solving this, one would inevitably find that the initial condition one imposed would lead to a renormalized coupling at some scale.

Alternatively, we could adopt the trick used in the EM example to write $\Lambda = \mu/\epsilon$ and find

$$-\frac{1}{2}\cot \delta_0 (k; c, \Lambda) = \frac{1}{c} - \frac{1}{\pi}\left(\log \frac{k}{2\mu} + \log \epsilon + \gamma\right) + \mathcal{O}\left(\epsilon^2 \times \frac{k^2}{\mu^2}\right)$$

and jiggle the $\ln \epsilon$ into the c term

Instead, however, we take the approach pioneered at the end of the last section, and think more physically - we attempt to use an observation at scale μ to find a better quantity than c to hold constant in the limit. We find

$$\cot \delta_0 (\mu; c, \Lambda) = -\frac{2}{c} + \frac{2}{\pi}\left(\ln \frac{\mu}{2\Lambda} + \gamma\right) + \mathcal{O}\left(\frac{\mu^2}{\Lambda^2}\right)$$

and accordingly define

$$\boxed{\frac{1}{c_R(\mu)} \stackrel{\text{def}}{=} -\frac{1}{2}\cot \delta_0 (\mu; c_{\text{phys}}, \Lambda_{\text{phys}})}$$

where we emphasize that the observation is made at the physical values of $c_{\text{phys}}, \Lambda_{\text{phys}}$ - which we don't know, but which correspond to whatever the microscopic description really secretly is. We then proceed to attempt to forget about these parameters! The $-1/2$ factor is, of course, simply to make c_R appear in the same place as c with the same numerical factor. This helps to emphasize the self-similarity of the problem. (In general, one thinks of c_{phys} as being c_R but evaluated at a scale $\mu \to \Lambda_{\text{phys}}$ - we will later explore the notion of effectively 'lowering the cutoff' Λ and obtaining similar equations, but with renormalized coefficients c_R.)

We can now go back to our general expression for the phase shift, and write everything in terms of our new favourite parameters

$$-\frac{1}{2}\cot \delta_{0,\text{phys}}(k) = \boxed{-\frac{1}{2}\cot \tilde{\delta}_0 (k; c_R(\mu), \mu) = \frac{1}{c_R} - \frac{1}{\pi}\log \frac{k}{\mu} + \mathcal{O}(\epsilon^2)}$$

where we expect $\epsilon = \mu/\Lambda$. However, we should think of this as an expression good for k near to μ - the large logarithms, or other terms involving k/μ, may ruin the expansion for other k.

Of course, it is by now no surprise that we can figure out how $c_R(\mu)$ depends on its scale by enforcing the independence of physical observables on the scale we make the parameter-fitting observation at - inevitably, it turns out to have the same form as the above equation for $c(\Lambda)$:

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\frac{1}{c_R(\mu)}\right) = -\frac{1}{\pi} + \mathcal{O}\left(\epsilon^2\right)$$

The core physical idea is worth reiterating once more: we make an observation (which in some sense is at a *low* energy scale μ) to fit a parameter c_R (which of course depends on μ) and then express our physical predictions in terms of this parameter c_R (and μ). We are still thinking of the original (albeit regularized) model as being (an approximation to) the correct high energy form of the theory, but we do not really think about what the 'bare' parameters at this high energy scale are. Then we can happily make predictions for what observations at arbitrary k near to the scale μ should give, based on our approximation to the high energy theory but with all calculations expressed in terms of low energy, accessible quantities.

2.4 Classical 'Field' Theories

Finally, before we move on to look at quantum theories, we briefly discuss a topic in classical theory which - at first sight - might seem to have nothing to do with the previous ideas.

2.4.1 1D Ising Model

The setup we will consider is the 1D Ising model, which consists of $N \gg 1$ spins (± 1, or $\uparrow$ and $\downarrow$) in a large circle - that is, with periodic boundary conditions - for which there are energy contributions according to (a) whether or not each spin is aligned with a background field, and (b) whether or not each spin is aligned with its neighbours. Adding in (c), a constant for each spin site (or chemical potential for N) we form a fairly general 1D 'nearest-neighbour' model. The Hamiltonian H is reduced to $\bar{H} = -\beta H$ where

$$\bar{H} = \underbrace{h\sum s_i}_{\text{(a)}} + \underbrace{K\sum_i s_i s_{i+1}}_{\text{(b)}} + \underbrace{CN}_{\text{(c)}}$$

and the partition function is (up to a constant)

$$\bar{Z} = \prod_i \left[\frac{1}{2}\sum_{s_i=\pm 1}\right] e^{\bar{H}[s_1,\ldots,s_N]}$$

The principle of the *renormalization group*[4] is to 'zoom out' on this model (just as we did in the quantum mechanical model) and see what the effective theory would be *if there were only half as many spins*. To do this, we do the summation in $\bar{Z}$ over *every other spin*.

Suppose some site s has adjacent spins $s_\pm$. Then the relevant term in $\bar{Z}$ is

$$\frac{1}{2}\sum_{s=\pm 1} \cdots \exp\left(K s_- s + \frac{1}{2}h\left(s_- + s\right) + C\right) \exp\left(K s_+ s + \frac{1}{2}h\left(s + s_+\right) + C\right) \cdots$$

[4] As ever, we emphasize that there is really not anything to the name 'group' here. The structure is not really anything more than gradually varying a parameter to obtain different couplings, with something like the structure of a semigroup, an entirely trivial observation.

(noting we have split the e^{hs} type terms consistently between the factors). An unenlightening series of algebraic steps lead to the conclusion that

$$\bar{Z} = \prod_{\text{even } i} \left[\frac{1}{2}\sum_{s_i=\pm 1}\right] \cdots \exp\left(K' s_- s_+ + \frac{1}{2}h'\left(s_- + s_+\right) + C'\right) \cdots$$

for some new variables $K' = K'\left(K,h\right)$, $h' = h'\left(K,h\right)$ and $C' = 2C + \tilde{C}'\left(K,h\right)$.

This is actually quite a remarkable result, on some reflection. We observe that the new expression for the theory, with partition function (assuming N even) written as

$$\bar{Z} = \prod_{\text{even } i} \left[\frac{1}{2}\sum_{s_i=\pm 1}\right] e^{\bar{H}'\left[s_2, s_4, \ldots, s_{N/2}\right]}$$

takes *exactly* the form of the old theory, except that $N \to N/2$, $K \to K'$, $h \to h'$ and $C \to C'$.

For simplicity assuming $N = 2^k$ for some large k, one can generate a series of expressions for $\bar{Z}$ with $N \to N^{(i)} = N/2^i$, $K \to K^{(i)} = \left(\cdots\left(\left(K\right)'\right)'\cdots\right)'$ and so forth. This self-similarity property is a discrete version of what we have discussed up until this point - we see that changing the scale associated with an observation (previously μ, or here something associated to N) alters the effective coupling constants one infers (previously α, c and here K, h, C).

Taking $N \to \infty$ means that one is actually left with the *same* system with changed couplings only.

2.4.2 Continuum Limits

THIS SECTION IS NEW AND ALSO INCOMPLETE!

An important concept in classical field theory is the concept of the *continuum limit* of a field theory defined on a lattice. Consider an (infinite) 2D Ising model in the absence of a external field; we take the action to be

$$S = -K \sum_{\mathbf{n}=(n_1,n_2)} \left[\sigma_{\mathbf{n}}\sigma_{\mathbf{n}+(0,1)} + \sigma_{\mathbf{n}}\sigma_{\mathbf{n}+(1,0)}\right]$$

where K takes the role of a dimensionless inverse temperature, and $\sigma_{\mathbf{n}} = \pm 1$ is a spin at each lattice site $\mathbf{n} \in \mathbb{Z}^2$. Then we have partition function

$$Z\left(K\right) = \sum_{\sigma_{\mathbf{n}}=\pm 1} e^{-S}$$

and correlation functions

$$\left\langle \sigma_{\mathbf{n}_1} \cdots \sigma_{\mathbf{n}_m} \right\rangle_K = \frac{\sum_{\sigma_{\mathbf{n}}=\pm 1} \sigma_{\mathbf{n}_1} \cdots \sigma_{\mathbf{n}_m} e^{-S}}{Z\left(K\right)}$$

We measure space in lattice units (so that $\mathbf{n} \in \mathbb{Z}^2$ is a sensible convention), where the lattice spacing a is a length. Thus physical masses go as $1/\left(\xi a\right)$ for some dimensionless number ξ called the *correlation length* which is a property of the theory. If one imagines taking a continuum limit $a \to 0$, it is now obvious that a *finite mass* theory must have an *infinite correlation length* to compensate for the fact that $a \to 0$.

A lattice theory which has an infinite correlation length is called a *critical* theory. Thus we expect that we will have to tune the single parameter K in our theory to some *special* value K_c in order to achieve a limit with sensible macroscopic properties. That is, we are imagining shrinking the lattice

(zooming out, $a \to 0$) such that the *observable* Compton wavelength of particle excitations in the continuum limits remains finite; to do this, we choose K_c carefully. (In fact, K_c will be the dividing line between two phases of the lattice theory - in the *low* temperature phase $K > K_c$ the $\mathbb{Z}_2$ symmetry $\sigma \to -\sigma$ is broken, and a vacuum expectation value $|\langle \sigma_{\mathbf{n}} \rangle_K| = s(K) \neq 0$ will appear.)

Forgetting about the interpretation of these correlation lengths, we *define* $\xi = \xi(K)$ by

$$\langle \sigma_{\mathbf{n}} \sigma_{\mathbf{0}} \rangle_K \sim \exp(-|\mathbf{n}|/\xi(K))$$

for large $|\mathbf{n}|$.

One can ...

2.4.3 Scaling and Effective Field Theories

In order to obtain an heuristic understanding of the concepts involved in effective field theories, we are going to do a quick calculation which illustrates the picture you should have in the back of your mind.

We imagine a field theory which arises as the 'low energy' limit (quotation marks as ever because this can be compared to very large scales like the Planck energy) of a theory we are not pretending to know about. What should this look like? Suppose that we integrate out all momenta above Λ, and are left with an *effective* Lagrangian which is schematically

$$\mathcal{L}_{\text{eff}} = \cdots + \frac{g_n}{\Lambda^n} O^{(n+4)} + \cdots$$

where each $O^{(n+4)}$ is a pure operator has mass dimension $n+4$. Note that inserting the Λ^n term in this way guarantees that (since $S = \int \mathrm{d}^4 x \, \mathcal{L}$ must be dimensionless if $Z = \int \mathrm{D}[\psi] \exp(iS)$ is to be well-defined so $\mathcal{L}$ must have mass dimension 4) all constants g_n are dimensionless.

As an example, for a scalar field with a canonically normalized kinetic term

$$\mathcal{L}_{\text{eff}} \sim (\partial\phi)^2 + \cdots$$

one discovers ϕ has mass dimension 1. Hence ϕ^4 also appears with no Λ terms, whilst a mass-type term ϕ^2 is expected to appear with a Λ^2 term. In general, we expect that each g_n is $\mathcal{O}(1)$, since we have no reason to suspect otherwise.[5]

The idea is to now just see how the action behaves for ϕ in the low momentum limit. To explore this in a simple way, we consider a given scalar field $\phi(x)$ (with some typical momentum scale if you like), and construct a family $\phi_\xi(x) := \phi(\xi x)$ for $\xi \to 0$. We are smearing out the field in space. Writing $x' = \xi x$ and forgetting about the cutoff (assuming that the field has no high-momentum excitations)

$$S\left[\phi_\xi(x); \ldots g^{(n)}, \ldots\right] \sim \int \mathrm{d}^4 x \left[(\partial_x \phi_\xi)^2 + \cdots + \frac{g^{(n)}}{\Lambda^n} \phi_\xi^{n+4} + \cdots\right]$$

We now introduce $\phi' = \xi^{-1}\phi$, so that

$$S\left[\phi_\xi(x); \ldots g^{(n)}, \ldots\right] \sim \int \mathrm{d}^4 x' \left[(\partial_{x'} \phi')^2 + \cdots + \frac{g^{(n)}}{\Lambda^n} \xi^n (\phi')^{n+4} + \cdots\right]$$

[5] There are obvious exceptions to this; in particular if the original theory has a symmetry (which is preserved under quantization) then we will expect $g_n = 0$. We might also imagine that if the original theory had even *higher* energy scales Λ' in it, then terms $h_n/\Lambda'^n = [h_n (\Lambda/\Lambda')^n]/\Lambda^n$ will occur, with $h_n = \mathcal{O}(1)$ so that $g_n = \mathcal{O}((\Lambda/\Lambda')^n) \ll 1$. This only works if symmetry-type arguments prevent a $1/\Lambda^n$ term from dominating this smaller term. This is not particularly relevant to our interests. The most physically important consequence of this expectation is the *naturalness problem* for scalar fields - if $\mathcal{L}_{\text{eff}}$ contains a term $m^2\phi^2$ then we expect $m \propto \Lambda$. The fact that the Higgs field has a weak-scale ϕ^2 coefficient rather than, say, a Planck-scale coefficient is a little troubling.

But this now has the form of the original action evaluated with a different set of coefficients and fields:

$$S\left[\phi_\xi(x);\ldots g^{(n)},\ldots\right]=S\left[\xi^{-1}\phi(x);\ldots\xi^n g^{(n)},\ldots\right]$$

This is interesting, because now in the $\xi\to 0$ limit we note that it is likely possible to compute the right-hand side more straightforwardly than the left - the reason is that the terms with $n>0$ are suppressed by ξ^n. These terms are called *irrelevant*. Similarly, $n<0$ terms are *relevant* whilst $n=0$ terms are *marginal*.

More concretely, if one takes ϕ to be a wavepacket with some definite energy scale k then one finds that the effects of $O^{(n+4)}$ are multiplied by $(k/\Lambda)^n$.

This observation, though very crude, is key to the logic we will explore shortly. The idea to take away is that at low energies, one can assume the high energy Lagrangian only had relevant and marginal couplings present at all, since the other can have only transient effects. (The effects are to influence the running of the other couplings, but since we only observe *renormalized* couplings at low energies, and not the high energy *bare* couplings, this is not something we have to worry about. We can just assume the bare couplings are 'just so' that we get the results observed.) Another important way we will look at this is to say that at low energies, the Lagrangian is *attracted to a finite-dimensional family of Lagrangians* which one could parametrize by *either* an arbitrary set of observable renormalized couplings, or a set of bare couplings which are exactly the relevant and marginal ones.

Hence for scalar field theory, we expect a *three-dimensional* space of possible Lagrangians with coordinates corresponding to (in the bare theory at least, though their low-energy analogues can also be used) m^2,λ and the field's normalization.

3 Quantum Field Theories

Before we look at pursuing the above foray into field theory, we will look briefly at a typical QFT loop calculation to make contact with the original ideas we developed for the EM and QM examples. Then, we will return to considering the nature of effective field theories and flows in the space of Lagrangians.

3.1 A Divergence

Let's consider the usual toy model of ϕ^4 theory...

$$\mathcal{L}=-\frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi)-\frac{1}{2\hbar^2}m^2\phi^2-\frac{g}{4!\hbar}\phi^4,\qquad Z=\int D[\phi]\exp\left(i\int \mathrm{d}^4x\,\mathcal{L}/\hbar\right)$$

(Don't worry, we'll set $\hbar=1$ faster than you can blink.)

For a vertex with momenta k_1,k_2 in & k_3,k_4 out, we find

$$\mathcal{M}(\{k_i\})=-ig+\frac{1}{2}(-ig)^2\sum_{p=k_1+k_2,\ldots}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\frac{i}{k^2+m^2-i\epsilon}\frac{i}{(p-k)^2+m^2-i\epsilon}+\mathcal{O}(g^3)$$

where the sum is over $p=k_1+k_2,k_1-k_3,k_1-k_4$. This integral is immediately divergent, being of the form $\int \mathrm{d}^4k/k^4$. Let's impose a momentum cutoff. We find

$$\mathcal{M}(\{k_i\})=-ig+iCg^2\left[\log\frac{\Lambda^2}{s}+\log\frac{\Lambda^2}{t}+\log\frac{\Lambda^2}{u}\right]+\mathcal{O}(g^3)$$

where the momentum cutoff is Λ, and we have let $s = (k_1 + k_2)^2$, $t = (k_1 - k_3)^2$, $u = (k_1 - k_4)^2$ and neglected the mass terms ($m \to 0$).

In the spirit of the previous calculations, we should storm ahead and imagine an observation at some typical scale. Here, however, we have *four* inputs to our experiment, which are all *four*-vectors (possibly subject to a mass shell constraint) k_1, k_2, k_3, k_4. We could easily fix all *sixteen* of these to be different quantities in our experiment, and then have four quantities $\mu_1, \mu_2, \ldots, \mu_{16}$. For simplicity, however, we choose our k_is such that $s = t = u = \mu^2$.[6] Call these variables $\tilde{k}_i(\mu)$. Then we can go ahead and define some renormalized couplings g_R by

$$-ig_R(\mu) = \mathcal{M}\left(\{\tilde{k}_i(\mu)\}\right)$$

Then on the grounds that $g_R(\mu)$ is a physical quantity, so that we want $\frac{\mathrm{d}g_R}{\mathrm{d}\Lambda} = 0$, one would allow $g = g(\Lambda)$. Then substitution gives:

$$\mathcal{M}(\{k_i\}) = -ig_R(\mu) + iCg_R(\mu)^2\left[\log\frac{\mu^2}{s} + \log\frac{\mu^2}{t} + \log\frac{\mu^2}{u}\right] + \mathcal{O}\left(g_R(\mu)^3\right)$$

so that our theory now allows to make predictions for arbitrary momenta k_i, exactly as we desired, based on one observation![7] (By substitution, we mean solving $-ig_R(\mu) = \mathcal{M}\left(\{\tilde{k}_i(\mu)\}\right)$ for g order by order, and then substituting this solution into our original expression for $\mathcal{M}(\{k_i\})$. Then all the remaining Λ go away! One should be able to rephrase this in terms of order-by-order counterterms in roughly the usual manner.)

However, this picture conceals lots of complexity. What about higher-order terms? What if we had to impose $g = g(\Lambda, k_i)$ to fix $\mathrm{d}g_R/\mathrm{d}\Lambda$? How do we know that all the *other* integrals in the theory are rendered finite by this choice of g, and possibly a choice of m if we take this to be non-zero? (There are lots of complicated issues to do with *sub*divergences in diagrams.)

Perhaps most worryingly, we haven't looked too closely at the 'smallness' of the constants we are expanding in. We have error terms $\mathcal{O}\left(g^3\right)$ and $\mathcal{O}\left({g_R}^3\right)$, but in general, are both small? Can we invert these series?

Indeed, a standard observation in QFT is that the series one obtains in solutions are almost always at best *asymptotic*, and essentially never form convergent power series. This isn't very surprising; it's common that setting a coupling to be slightly *negative* instead of slightly *positive* causes the theory to blow up completely, so we shouldn't ever expect analytic power series. The prototype for this particular example is that

$$\int_{-\infty}^{+\infty}\mathrm{d}x\, e^{-\frac{1}{2}x^2 - \frac{1}{4}gx^4} \sim \sum_{n\geq 0}(-1)^n\sqrt{2}\frac{\Gamma\left(2n+\frac{1}{2}\right)}{n!}g^n$$

which is very definitely not a convergent series. (The coefficients diverge factorially, worse than n^n.) Similarly, ϕ^4 has an unstable vacuum if the ϕ^4 coefficient has the wrong sign, and QED is unstable to the addition of weak couplings (as seen in the Cooper instability to pair formation, as present in superconductivity).

Yet the above amounts to roughly what one does in standard calculations... Why does it work? At least some of the reason is that everything above is best viewed as a *formal* power series. But

[6] This point is not *kinematically accessible* - there is no set of physical (on-shell) momenta k_i which gives these values of s, t, u. This shouldn't worry you too much; you can choose different variables if you want. Note that there are also many possible choices of $(k_i)^\mu$ satisfying these constraints.

[7] Even if the above point is not kinematically accessible, we can infer (to some order in g_R) what g_R should be from a *physical* observation using this renormalized expression.

it is difficult to answer all of the questions about the endlessly huge number of possible divergent calculations one can form in the QFT. To address these issues satisfactorily, we clearly need to look not at individual diagrams and divergences, but at the *action* and *Lagrangian* themselves.

3.2 $\mathcal{L}$-Space

We're going to think a bit about *Lagrangian space*, or $\mathcal{L}$-space for short in this section. Specifically, we can write down the most general possible Lagrangian $\mathcal{L}$ (which can be generated by loop corrections, with no 'renormalizability' criterion), which requires an infinite number of coefficients, and use these coefficients as coordinates. For example, for a scalar field theory with a $\phi \to -\phi$ $\mathbb{Z}_2$ symmetry, we might write

$$\mathcal{L} = c_1\phi^2 + c_2\left(\partial\phi\right)^2 + c_3\phi^4 + c_4\left(\partial\phi\right)^2\phi^2 + c_5\phi^6 + \cdots$$

so that $\mathcal{L}$-space has coordinates $(c_1, c_2, c_3, c_4, c_5, \cdots)$. (You could easily make these dimensionless by introducing a mass scale if you like.)

Now the ideas we are going to use all focus on the idea of *integrating out high-momentum modes* - first of all, we want to make clear what this means.

The idea is actually seen most clearly (or physically, at least) in terms of Feynman diagrams. Suppose you have a diagram with three 'low-momentum' incoming and three outgoing lines, connected in the usual ϕ^4 theory with cutoff Λ_{old} in the obvious way, with all three incoming lines going into a ϕ^4 vertex, a propagator leaving this and reaching a second ϕ^4 vertex from which the three outgoing lines originate. But now suppose the three incoming momenta add up to make the central line 'high-momentum', schematically '$\Lambda_{\text{new}} < p < \Lambda_{\text{old}}$' - then if we were to consider a theory where we integrate only up to Λ_{new}, we could not reproduce the contribution of this diagram naïvely.

Instead, we imagine 'zooming out' a bit - what do we see? We see (a) three lines going into a vertex, (b) something magic happens, (c) three lines come out. But this middle stage is invisible at our scale, so in fact we see a *six-point vertex*! The idea is that $\mathcal{L}_{\text{old}} \to \mathcal{L}_{\text{new}}$ under this integration, where $\mathcal{L}_{\text{new}}$ contains c_4 and/or c_5 non-zero. (The dependence of this diagram and permutations thereof etc. upon momenta determines which vertex this contributes to.)

We can think of constructing paths in $\mathcal{L}$-space parametrized by *cutoffs*. What do we mean by this?

- Start with some high energy effective theory with Lagrangian $\mathcal{L}_0$ and a (real!) cutoff Λ_0.
- Compute the effects of modes with momenta $\Lambda' < p \leq \Lambda_0$ on lower energy modes.
- Rephrase this in terms of a new Lagrangian $\mathcal{L}'$.
- Thus we draw a path connecting $(\mathcal{L}_0, \Lambda_0) \rightsquigarrow (\mathcal{L}', \Lambda')$.
- In general, we have a path $\mathcal{L}(\Lambda)$.

This is interesting. Physically, we imagine that - at some large energy scale Λ_0, like $\Lambda_0 = \Lambda_{\text{Planck}}$ - the True Wonderful Theory of Everything becomes an effective field theory with Lagrangian $\mathcal{L}_0$. This is a realistic theory, but one might imagine difficult to compute with, since even if all of our momenta are at some reasonable scale $\mu \ll \Lambda_0$, we still have to compute integrals $\sim \int^{\Lambda_0}$ with large contributions at large Λ_0 in order to work out physical amplitudes.

More specifically, $\mathcal{L}_0$ has *infinitely many unknown coordinates* which are not directly related to physical observables, since we can mostly only do calculations perturbatively in c_j, but we might get

series with some very large terms like $\log(\mu/\Lambda_0)\, c_j^2$ which are almost completely useless for computing the c_j with.

Morally, what we want to do is follow the path $\mathcal{L}(\Lambda)$ all the way down from $\mathcal{L}_0 = \mathcal{L}(\Lambda_0)$ down to $\mathcal{L}_{\text{reasonable}} = \mathcal{L}(\Lambda_{\text{reasonable}})$, where $\Lambda_{\text{reasonable}}$ is just slightly larger than the momenta we want to do our calculations with, and then do computations *using this new Lagrangian* and the *nice low cutoff* $\Lambda_{\text{reasonable}}$.

Note that *this description involves no infinities* as stated. In fact, there isn't even a whiff of regularization. This is all because we started off by assuming that some genuine, true magical-physical theory thing looks, at the scale Λ_0, like a field theory with a cutoff. There was never a field theory with no cutoff present in our description. This is all fine with us; in fact, it might make us very pleased with ourselves!

But there is a reason for wanting Λ_0 to be very large, arising from the idea discussed in section 2.4.3 that at low energies, many couplings are (technically) *irrelevant*. How will this manifest itself?

The argument we will attempt to make is that *as we scale down an EFT from a high cutoff, the Lagrangian path is attracted towards a finite dimensional submanifold with dimension equal to the number of relevant and marginal couplings.* So for $\phi \to -\phi$ scalar field theory, we expect an attracting three-dimensional manifold, since c_1, c_2, c_3 are the only couplings with mass dimension ≥ 0.

We need to make what we mean slightly more precise. Since we want to have a theory at a fixed energy scale $\Lambda_{\text{reasonable}}$ - just above our observations - we take the limit 'the wrong way round'. To avoid any ambiguity, $\mathcal{L}(\Lambda; \mathcal{L}_0, \Lambda_0)$ is defined to be the Lagrangian reached by following the path from $\mathcal{L}_0$ at cutoff Λ_0 down to Λ. Then we can consider

$$\text{the set } U \subset \mathcal{L}\text{-space such that } \lim_{\Lambda_0 \to \infty} \text{dist}\left(\mathcal{L}(\Lambda; \mathcal{L}_0, \Lambda_0), U\right) = 0 \text{ for all (physical) } \mathcal{L}_0$$

where $\text{dist}(x, U)$ measures the distance between x and the nearest point $y \in U$ to x, measured according to some metric like $\sup_i |x_i - y_i|$ where x_i, y_i are the coefficients of x, y. Here we claim that U is a finite-dimensional submanifold, and that trajectories 'squash up' to U as Λ decreases - but in order to emphasize that our scale of observation is the physical scale of relevance, we hold that fixed, and so the limit arises from $\Lambda_0 \to \infty$.

Remark. For a very simple example of this, the equations $\partial_x f(x) = -f(x)$ and $\partial_x g(x) = 0$ have solutions $f = f_0 e^{-(x-x_0)}$ and $g(x) = g_0$, and so as $x \to \infty$ one approaches the manifold $f = 0, g =$ const. - but we can instead say that $x_0 \to -\infty$ at fixed x and achieve the same result.

We use the example Polchinski uses to illustrate the type of situation arising in renormalization.

3.3 Polchinski's Example

Suppose we have one *relevant* and one *marginal* coupling, respectively the dimensionless coupling g_4 and the dimension $(\text{mass})^2$ coupling g_6. (The suggestive names draw a parallel with 4- and 6-point functions in a 4D scalar theory.) Then the couplings satisfy renormalization group equations (equations governing their dependence on a mass scale or cutoff)

$$\Lambda \frac{\mathrm{d}g_4}{\mathrm{d}\Lambda} = \beta_4\left(g_4, \Lambda^2 g_6\right) \qquad \Lambda \frac{\mathrm{d}g_6}{\mathrm{d}\Lambda} = \Lambda^2 \beta_6\left(g_4, \Lambda^2 g_6\right) \tag{3.1}$$

on dimensional grounds. We define $\lambda_4 = g_4, \lambda_6 = \Lambda^2 g_6$ to be non-dimensionalized versions of these couplings so that

$$\Lambda \frac{\mathrm{d}\lambda_4}{\mathrm{d}\Lambda} = \beta_4(\lambda_4, \lambda_6) \qquad \Lambda \frac{\mathrm{d}\lambda_6}{\mathrm{d}\Lambda} - 2\lambda_6 = \beta_6(\lambda_4, \lambda_6)$$

Now we consider a particular solution of these evolution equations given by some functions $\left(\bar{\lambda}_4(\Lambda), \bar{\lambda}_6(\Lambda)\right)$. Let $\epsilon_i = \lambda_i - \bar{\lambda}_i$ be a small perturbation to this orbit. We find its evolution with Λ to be (linearized in ϵ)

$$\Lambda \frac{\mathrm{d}\epsilon_4}{\mathrm{d}\Lambda} \approx \overline{\frac{\partial \beta_4}{\partial \lambda_4}}\epsilon_4 + \overline{\frac{\partial \beta_4}{\partial \lambda_6}}\epsilon_6 \qquad \Lambda \frac{\mathrm{d}\epsilon_6}{\mathrm{d}\Lambda} - 2\epsilon_6 \approx \overline{\frac{\partial \beta_6}{\partial \lambda_4}}\epsilon_4 + \overline{\frac{\partial \beta_6}{\partial \lambda_6}}\epsilon_6$$

where the barred derivatives signify evaluation at $\left(\bar{\lambda}_4, \bar{\lambda}_6\right)$.

Remember that β functions are loop phenomena; they are generically $\mathcal{O}\left(\lambda^2\right)$.[8] Suppose we evolve towards a *lower* cutoff (Λ decreases from a large Λ_0) and assume we head towards a region where λ is reasonably small. Then the $\overline{\partial\beta/\partial\lambda}$ terms become small, so the ϵ_6 equation suggests that ϵ is suppressed by powers of $(\Lambda/\Lambda_0)^2$. This supports the previous analysis, though in a *slightly* more rigorous manner.

Hence our trajectory is converging towards the $\bar{\lambda}$ trajectory in the λ_6 direction! But *before* $\epsilon_6 \to 0$, it affects (in a generically leading order manner) the running of ϵ_4 in a non-trivial manner.

Now if we want to explore whether both the trajectory and its perturbation are indeed being attracted towards some manifold, we want to see if these are both attracted towards each other *as lines* - we don't care about the parameter Λ. That is, since the running of ϵ_4 is at leading order different according to the values of λ_6, we might expect to attract towards the manifold at an offset in the λ_4 direction. We have no reason to compare $\bar{\lambda}(\Lambda)$ with $\left(\bar{\lambda} + \epsilon\right)(\Lambda)$ specifically.

To look at this another way: suppose $\epsilon_6, \bar{\lambda}_6$ magically vanish exactly at and below some Λ'. Then at this point, the two evolution equations reduce to simply the running of λ_4. But now the 'initial' condition $\lambda_4(\Lambda')$ is all that determines the running, and the equations are autonomous in $\ln\Lambda$ (that is, as a differential equation with respect to the variable $\ln\Lambda$, the equations do not feature Λ) - hence even though the two trajectories have different 'initial' conditions they are probably the *same* trajectory, at different parameter values.

So what are ϵ_4 and ϵ_6 telling us? They describe the deviation from a trajectory at fixed parameter values. We want a deviation from the *nearest* point on the manifold, not some other random point.

To fix this, we subtract off something parallel (at leading order) to the proposed manifold. A simple attempt might be

$$\begin{aligned}
\xi_4 &= \epsilon_4 - \frac{\mathrm{d}\bar{\lambda}_4}{\mathrm{d}\Lambda} \times \left[\epsilon_4 / \frac{\mathrm{d}\bar{\lambda}_4}{\mathrm{d}\Lambda}\right] \equiv 0 \\
\xi_6 &= \epsilon_6 - \frac{\mathrm{d}\bar{\lambda}_6}{\mathrm{d}\Lambda} \times \left[\epsilon_4 / \frac{\mathrm{d}\bar{\lambda}_4}{\mathrm{d}\Lambda}\right]
\end{aligned}$$

where the vector $\mathrm{d}\left(\bar{\lambda}_i\right)/\mathrm{d}\Lambda$ is parallel to the trajectory and the coefficient is chosen to make the λ_4-deviation vanish identically for simplicity. Then one computes the evolution of ξ_6 from

$$\Lambda \frac{\mathrm{d}\xi_6}{\mathrm{d}\Lambda} - 2\xi_6 \approx \left[\overline{\frac{\partial \beta_6}{\partial \lambda_4}} + \overline{\frac{\partial \beta_4}{\partial \lambda_4}} - \Lambda \frac{\mathrm{d}}{\mathrm{d}\Lambda} \ln \bar{\beta}_4\right] \xi_6$$

with the solution

$$\xi_6(\Lambda) \approx \xi_6(\Lambda_0) \left(\frac{\Lambda}{\Lambda_0}\right)^2 \left(\frac{\bar{\beta}_4(\Lambda_0)}{\bar{\beta}_4(\Lambda)}\right) \exp \int_{\Lambda_0}^{\Lambda} \frac{\mathrm{d}\Lambda'}{\Lambda'} \left[\overline{\frac{\partial \beta_6}{\partial \lambda_4}} + \overline{\frac{\partial \beta_4}{\partial \lambda_4}}\right]$$

Hence provided the β function runs sufficiently slowly, we gave $\xi_6 \sim (\Lambda/\Lambda_0)^2$ at low cutoffs, and it is suppressed! (The constraint on β functions is equivalent to saying that the *anomalous dimension* of

[8] This depends on the choice of λ and various other things, but we'll not worry about it.

λ_4 does not overwhelm the canonical one, which is due to the two mass dimensions.) Also this really *is* an irrelevant parameter, in the sense that its initial value does not affect either of the low energy parameters ξ_4, ξ_6 (or rather $\bar{\lambda}_i + \epsilon_i$).

This demonstrates that there is a one-dimensional manifold to which we are being attracted. Practically, as soon as we know λ_4 we know λ_6 at least to an accuracy $\sim (\Lambda/\Lambda_0)^2$.

To make contact with our above $\lim_{\Lambda_0\to\infty} \mathrm{dist}\left(\mathcal{L}\left(\Lambda;\mathcal{L}_0,\Lambda_0\right),U\right) = 0$ statement, we now rewrite what we are doing in a different language. Up until now, we have been imagining a 'bare' theory at Λ_0 with parameter λ^0 (perhaps with $\lambda_6^0 = 0$ for simplicity, though this is fairly arbitrary) which is evolved down to $\Lambda_R \ll \Lambda_0$ to give a 'renormalized' theory $\lambda_{4,6}^R$. Now, we consider a *fixed* scale Λ_R with a single *fixed* renormalized coupling λ_4^R; then we look for a theory at cutoff Λ_0 such that we recover our 'renormalized' theory at Λ_R. We choose this theory to have $\lambda_6^0 = 0$ for simplicity. (We only really need it to be bounded for large Λ_0.) This *defines* $\lambda_4^0 = \lambda_4^0\left(\Lambda_0;\lambda_4^R,\Lambda_R\right)$.

Note that the reason we do not construct this to depend on λ_6^R is because we expect $\lambda_6^R \equiv \lambda_6^R\left(\lambda_4^R\right)$ in the limit of $\Lambda_0/\Lambda_R \gg 1$. Specifically, its dependence on Λ_0 goes as $(\Lambda_R/\Lambda_0)^2$.

Now we are in a position to take the limit $\Lambda_0 \to \infty$, 'taking the cutoff to infinity', whilst adjusting the 'bare' couplings λ_i^0 to reproduce some physically observed λ_4^R. Then

- The notion of *renormalizability* is that we can take such a limit[9] (adjusting the bare couplings as needed) and recover a *finite* renormalized Lagrangian with *finite* couplings λ_4^R, λ_6^R.
- The prediction of the theory is that $\lambda_6^R \equiv \lambda_6^R\left(\lambda_4^R\right)$ in the particular way obtained by our arbitrary $\lambda_6^0 = 0$ prescription.

Reformulating the above arguments (in terms of ϵ, ξ) within our $\Lambda_0 \to \infty$ framework, we consider first the quantity $\Lambda_0\left(\partial\lambda_i^R/\partial\Lambda_0\right)$ which describes how the renormalized couplings

$$\lambda_i^R \equiv \lambda_i^R\left(\Lambda_R;\lambda_4^0,\Lambda_0\right) \equiv \lambda_i^R\left(\Lambda_R;\lambda_4^0\left(\Lambda_0;\lambda_4^R,\Lambda_R\right),\Lambda_0\right)$$

vary as we take our infinite cutoff limit $\Lambda_0 \to \infty$. Then, in order to fix the issue discussed above, one considers

$$v_i = \Lambda_0 \frac{\partial\lambda_i^R}{\partial\Lambda_0} - \frac{\partial\lambda_i^R}{\partial\lambda_4^0} \times \left[\left(\frac{\partial\lambda_i^R}{\partial\lambda_4^0}\right)^{-1} \Lambda_0 \frac{\partial\lambda_4^R}{\partial\Lambda_0}\right]$$

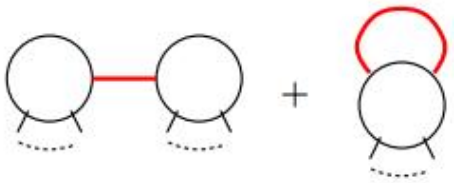
where $\partial\lambda_i^R/\partial\lambda_4^0$ is another vector nearly parallel to the claimed attracting manifold. The advantage of this choice is that

$$v_i = \Lambda_0 \frac{\mathrm{d}}{\mathrm{d}\Lambda_0}\lambda_i^R \equiv \Lambda_0 \frac{\mathrm{d}}{\mathrm{d}\Lambda_0}\lambda_i^R\left(\Lambda_R;\lambda_4^0\left(\Lambda_0;\lambda_4^R,\Lambda_R\right),\Lambda_0\right)$$

where the derivative is effectively at fixed λ_4^R. Then one computes that $v_i \sim \Lambda_R^2/\Lambda_0^2$ up to slowly varying functions and hence λ_i^R have sensible limits.

Remark. If one wished to recreate the usual (and from this point of view, perverse) computations in perturbative QFT by solving the original equations (3.1) order by order in the *bare* coupling g_4^0 ($g_6^0 = 0$) then one would find divergences which cancel when one converts to expressing everything in terms of the renormalized g_4^R.

[9] In perturbation theory it is possible to specify λ_4^0 as a formal power series in λ_4^R, but in general it is possible that $\lambda^0 \to \infty$ as Λ_0 approaches some finite critical value.

$$\frac{\partial}{\partial t}\mathcal{V}_{2n}(t;p_1,\cdots,p_{2n}) =$$ 

3.4 A Concrete Cutoff

To implement a cutoff in momentum space, we will smoothly modify the propagator term in the Lagrangian to

$$K\left(p/\Lambda\right)/\left(p^2+m^2\right)$$

where we require $K\left(0\leq x\leq 1\right)=1$ and $K\left(x\to\infty\right)\to 0$ rapidly. The partition function is

$$Z\left[J;g_\Lambda,\Lambda\right]=\int D\left[\phi\right]\exp\left(\int \mathrm{d}^4x\, J\phi+\underbrace{\frac{1}{2}i\int \mathrm{d}^4p\,\frac{p^2+m^2}{K\left(p/\Lambda\right)}\phi\left(p\right)\phi\left(-p\right)+i\int \mathrm{d}^4x\,\mathcal{L}_{\text{int}}\left[\phi;g_\Lambda,\Lambda\right]}_{S[\phi;g_\Lambda,\Lambda]}\right)$$

where J contains no high-momentum modes. We will parametrize changes in the cutoff with t for $\Lambda=\mu e^{-t}$, separating out the two aspects of the cutoff as before.

We want to essentially find $\partial Z/\partial t=0$, so that physical predictions are independent of scale, say, $\left\langle\phi\left(\mathbf{p}_1\right)\cdots\phi\left(\mathbf{p}_n\right)\right\rangle_{S(t)}=\left\langle\phi\left(\mathbf{p}_1\right)\cdots\phi\left(\mathbf{p}_n\right)\right\rangle_{S(t+\Delta t)}$.

In this point of view, we seek to *compensate a change in the propagator by a change in* $\mathcal{L}_{int}$.

To work out this in detail, we split the propagator up as

$$K\left(\frac{p}{\mu e^{-t}}\right)=K\left(\frac{p}{\mu e^{-t-\Delta t}}\right)+\left[-\frac{\partial}{\partial t}K\left(\frac{p}{\mu e^{-t}}\right)\right]\Delta t$$

where

$$\left[-\frac{\partial}{\partial t}K\left(\frac{p}{\mu e^{-t}}\right)\right]=:\Delta\left(\frac{p}{\mu e^{-t}}\right)$$

represents a special term multiplying the propagator, which essentially only gives contributions for $p\in\left(\mu e^{-t-\Delta t},\mu e^{-t}\right)$. Lines due to this term in the μe^{-t} theory do not appear in the $\mu e^{-t-\Delta t}$ theory! Hence we might draw diagrams with red lines for these propagators, say.

The ambition is to *modify the vertices* to take account for the *effect of integrating out these lines*. For concreteness, impose $\phi\to-\phi$ symmetry (which is preserved without *anomalies* under renormalization group flow, just like e.g. the charge conjugation symmetry in QED) so that $\mathcal{L}_{\text{int}}$ may be decomposed as

$$-\sum_{n=1}^{\infty}\frac{1}{(2n)!}\int_{\sum p_i=0}\mathcal{V}_{2n}\left(t;p_1,\ldots,p_{2n}\right)\phi\left(p_1\right)\cdots\phi\left(p_{2n}\right)$$

where $\mathcal{V}_{2n}$ describes the effects of a particular $2n$-vertex. The following image depicts, schematically, the two effects that these special lines create, where the circles represent vertices (in the renormalized theory at the lower cutoff).

Firstly, one obtains modifications where the external lines are partition into two sets and connected to two vertices joined by the high-momentum line; secondly, one obtains terms with a high-momentum loop on a single $(2n+2)$-vertex.

The general result of this is more concretely expressed as

$$\partial_t \left(-S_{\text{int}}\right) = \frac{1}{2}\int \mathrm{d}^4 p \, \frac{\Delta\left(\frac{p}{\mu e^{-t}}\right)}{p^2+m^2}\left[\frac{\delta\left(-S_{\text{int}}\right)}{\delta\phi\left(p\right)}\frac{\delta\left(-S_{\text{int}}\right)}{\delta\phi\left(-p\right)} + \frac{\delta^2\left(-S_{\text{int}}\right)}{\delta\phi\left(p\right)\delta\phi\left(-p\right)}\right]$$

This achieves exactly what we wanted! We are given a prescription for calculating how to adjust the action to account for the effect of high-momentum lines.

The effective action, at lower cutoffs, contains many complicated vertices (of arbitrary valence vertices) but the structure is simple: these vertices arise purely through the type of diagrams we have just described! This effectively encapsulates the entire behaviour we sought to understand.

Example. At $t=0$, suppose the small number λ is the 4-point vertex. (We also assume m and the field normalization are chosen such that the zeroth order approximations to the propagator are the usual ones.) Then (at first order in λ) $\mathcal{V}_4^{(1)} = -\lambda$.

One may compute first-order (in λ) corrections to the two-point function using

$$\partial_t \mathcal{V}_2^{(1)}\left(t;p,-p\right) = \frac{1}{2}\int_q \frac{\Delta\left(\frac{q}{\mu e^{-t}}\right)}{q^2+m^2}\underbrace{\mathcal{V}_4^{(1)}\left(t;\cdots\right)}_{-\lambda} = -\frac{\lambda}{2}\int_q \frac{\Delta\left(\frac{q}{\mu e^{-t}}\right)}{q^2+m^2}$$

where $\int_q$ denotes the usual integral $\propto \int \mathrm{d}^4 q$. Asymptotically - at large t, in the 'bare' limit - we find

$$\int_q \frac{\Delta\left(\frac{q}{\mu e^{-t}}\right)}{q^2+m^2} = \mu^2 e^{-2t}\int_{\tilde{q}} \frac{\Delta\left(\tilde{q}\right)}{\tilde{q}^2+\frac{m^2}{\mu^2 e^{-2t}}} \sim \mu^2 e^{-2t}\int_{\tilde{q}} \frac{\Delta\left(\tilde{q}\right)}{\tilde{q}^2} - m^2\int_{\tilde{q}} \frac{\Delta\left(\tilde{q}\right)}{\tilde{q}^4}$$

Then we can compute the integral $\int_{-\infty}^{t}$ of our equation for $\partial_t \mathcal{V}_2^{(1)}$ straightforwardly:

$$\begin{aligned}
\mathcal{V}_2^{(1)} = -\frac{\lambda}{2}\int_q \frac{\Delta\left(\frac{q}{\mu e^{-t}}\right)}{q^2+m^2} &= -\frac{\lambda}{2}\left[\int_q \Delta\left(\frac{q}{\mu e^{-t}}\right)\left(\frac{1}{q^2+m^2}-\frac{1}{q^2}+\frac{m^2}{q^4}\right)\right. \\
&\quad \left. -\frac{\mu^2 e^{-2t}}{2}\int_q \frac{\Delta\left(q\right)}{q^2} - tm^2\int_q \frac{\Delta\left(q\right)}{q^4}\right] \\
&= -\frac{\lambda}{2}\left[\int_q \Delta\left(\frac{q}{\mu e^{-t}}\right)\frac{m^4}{q^4\left(q^2+m^2\right)}\right. \\
&\quad \left. -\frac{\mu^2 e^{-2t}}{2}\int_q \frac{\Delta\left(q\right)}{q^2} - tm^2\int_q \frac{\Delta\left(q\right)}{q^4}\right]
\end{aligned}$$

If one sets e.g. $K\left(x\right) = \theta\left(1-x\right)$ (even though this is a singular choice, with Δ containing δ functions) one may compute $\mathcal{V}_2^{(1)}\left(t\right)$ for varying t. Accordingly, the self-energy correction for instance is given by

$$-\frac{\lambda}{2}\int_{q<\mu e^{-t}} \frac{1}{q^2+m^2} + \mathcal{V}_2^{(1)}\left(t\right) = -\frac{\lambda}{\left(4\pi\right)^2} m^2 \ln\frac{m^2}{\mu^2}$$

which is bare-cutoff (t) independent, in agreement with the usual result and completely free of infinities in the argument, since we stayed away from the 'bare' $t\to-\infty$ limit. The μ dependence specifies how

varying the *low-energy* cutoff affects the result. (Thus, we can take $\Lambda_0 = \mu e^{-t} \to \infty$ and eliminate it, but $\Lambda_R = \mu$ has stuck around.)

One could nonetheless still consider the asymptotic limit discussed in the Polchinski case; in accordance with his arguments, $t \to -\infty$ gives divergent bare quantities (for some standard choice of bare quantities like $\mathcal{V}_{2n\geq 6} \to 0$, $\mathcal{V}_4 \to$ const., etc.) but a finite set of functions $\mathcal{V}_{2n}\,(t = 0)$.