3 EFT and radiative corrections

Up to now we have ignored quantum corrections in our effective theory. A Lagrangian such as eq. (56) 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 :

$$\mathcal{L}_{\rm UV} = \frac{1}{2}\left((\partial\phi)^2 - m^2\phi^2 + (\partial S)^2 - M^2 S^2 - \kappa\phi^2 S\right) \tag{89}$$

The parameter κ has dimension of mass, and I will assume $\kappa < M$. Never mind that the vacuum energy is unbounded below; 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 S field "integrated out". 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.

One performs the matching in a loop 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 one-loop diagrams agree with each other, then two-loop, etc. Why is this justified? Consider a graph with P propagators, V vertices and E external legs. Euler's formula tells us that $L = P - V + 1$. Furthermore, for a theory with a single type of interaction vertex involving n fields, it must be that $(E + 2P) = nV$, 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 these two equations together we have $V = (2L + E - 2)/(n - 2)$, which shows that for a given number of external lines, the number of vertices grows with the number of loops, so a loop expansion is justified if a perturbative expansion is justified. It is also worth noting that a loop expansion is an expansion in $\hbar$: 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 path integrand for the EFT is $\exp(iS_{\rm eft}/\hbar$, the L-loop matching involves contributions to $\mathcal{L}_{\rm eft}$ at $O(\hbar^L)$. It is convenient therefore to count powers of $\hbar$, keeping in mind that this is justified when perturbation theory is justified.

So we match the UV theory eq. (89) to the EFT order by order in a loop expansion $= \hbar$ expansion; since the EFT is expressed in terms of local operators, the matching also involves performing an expansion in powers of external momenta, translating to derivatives acting on fields in the EFT. 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, with the result

$$\mathcal{L}^0_{\rm eff} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - c_0\frac{\kappa^2}{M^2}\frac{\phi^4}{4!} - d_0\frac{\kappa^2}{M^4}\frac{(\partial\phi)^2\phi^2}{4} + \dots\ , \tag{90}$$

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 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 full 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

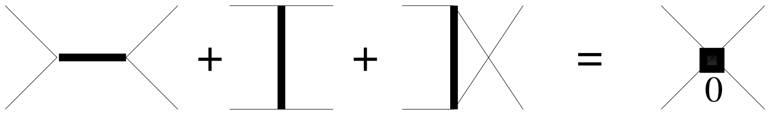

Figure 7: *Matching at $O(\hbar^0)$ between the UV theory and the EFT: on the left, integrating out the heavy scalar S (dark propagator); on the right, all contributions of four-point vertices the tree level EFT $\mathcal{L}^0$. Equating the two sides allows on to solve for this vertices.*

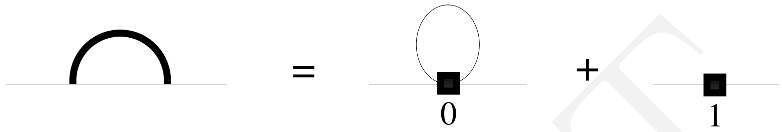

Figure 8: *Matching the 2-point function in the EFT at $O(\hbar)$. On the left, the 1-loop 1LPI graph contributing in the full theory, and on the right, graphs from the EFT include 1-loop graphs involving the 4-point vertices from $\mathcal{L}^0_{EFT}$, as well as $O(\hbar)$ tree-level contributions 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.*

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

$$\begin{aligned}
\mathcal{L}^1_{\rm eft} \;=\; & \frac{1}{2}\left(1 + a_1 \frac{\kappa^2}{16\pi^2 M^2}\right)(\partial\phi)^2 - \frac{1}{2}\left(m^2 + b_1 \frac{\kappa^2}{16\pi^2}\right)\phi^2 \\
& - \left[c_0\left(\frac{\kappa^2}{M^2}\right) + c_1\left(\frac{\kappa^4}{16\pi^2 M^4}\right)\right]\frac{\phi^4}{4!} \\
& - \left[d_0\left(\frac{\kappa^2}{M^4}\right) + d_1\left(\frac{\kappa^4}{16\pi^2 M^6}\right)\right]\frac{(\partial\phi)^2\phi^2}{4} + \ldots
\end{aligned} \tag{91}$$

where the coefficients a, b, c, d are going to be $\mathcal{O}(1)$. 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.

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 is not proportional to m^2, so that it is "unnatural" for the physical mass to be $\ll \frac{\kappa^2}{16\pi^2}$ — that 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).

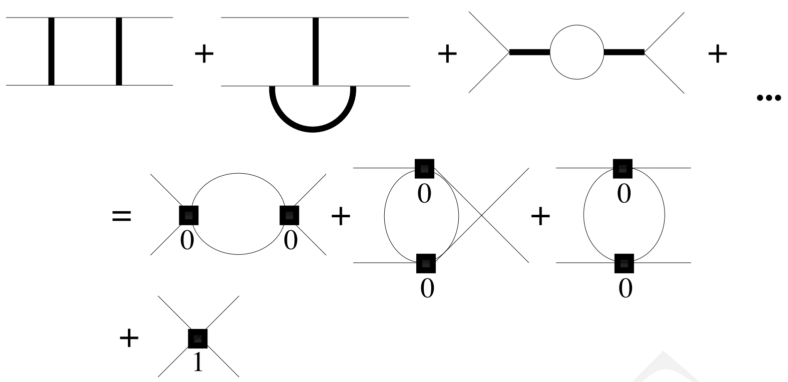

Figure 9: *Matching the 4-point function in the EFT at $O(\hbar)$. On the left, the 1LPI graphs in the full theory (with the ellipsis indicating other topologies), and on the right the $O(\hbar)$ contribution from the EFT, including 1-loop graphs involving the 4-point vertices from $\mathcal{L}^0_{EFT}$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

- In the matching conditions the graphs in both theories have pieces depending non-analytically on light particle masses and momenta (eg, $\ln m^2/M^2$ or $\ln p^2/M^2$)...these terms cancel on both sides of the matching condition so that the interactions in $\mathcal{L}_{\rm eft}$ have a local expansion in inverse powers of $1/M$. This is an important and generic property of effective field theories.

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 Landau liquid versus BCS instability

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 [4]. 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

$$S_{free} = \int dt \int d^3p \sum_{s=\pm\frac{1}{2}} \Big[\psi_s(p)^\dagger i\partial_t\psi_s(p) - (\epsilon(p) - \epsilon_F)\psi_s^\dagger(p)\psi_s(p)\Big] \tag{92}$$

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

$$\mathbf{p} = \mathbf{k} + \boldsymbol{\ell} \tag{93}$$

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 $\ell \ll k$,

$$\epsilon(\mathbf{p}) - \epsilon_F = \boldsymbol{\ell} \cdot \mathbf{v}_F(\mathbf{k}) + \mathcal{O}(\ell^2) \ , \tag{94}$$

and so $[\epsilon - \epsilon_f] = 1$ and $[\partial_t] = 1$. Given that the action eq. (92) isn't supposed to change under this scaling,

$$[\psi] = -\frac{1}{2} \ . \tag{95}$$

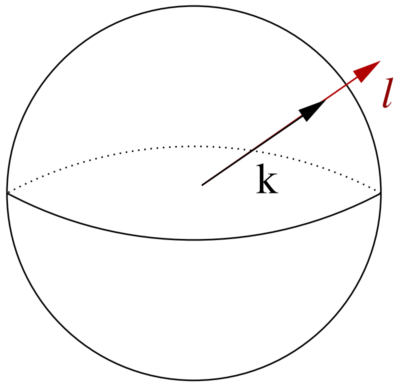

Figure 10: *The momentum* $\mathbf{p}$ *of an excitation is decomposed as* $\mathbf{p} = \mathbf{k} + \boldsymbol{\ell}$, *where* $\mathbf{k}$ *lies on the Fermi surface (which does not have to be a sphere), and* $\boldsymbol{\ell}$ *is perpendicular to the Fermi surface. Small* $|\boldsymbol{\ell}|$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

Now consider an interaction of the form

$$S_{int} = \int dt \int \prod_{i=1}^{4} (d^2\mathbf{k}_i d\ell_i)\delta^3(\mathbf{P}_{tot})C(\mathbf{k}_1, \ldots, \mathbf{k}_4)\psi_s^\dagger(\mathbf{p}_1)\psi_s(\mathbf{p}_2)\psi_{s'}^\dagger(\mathbf{p}_3)\psi_{s'}(\mathbf{p}_4) \ . \tag{96}$$

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 $\delta(0)$

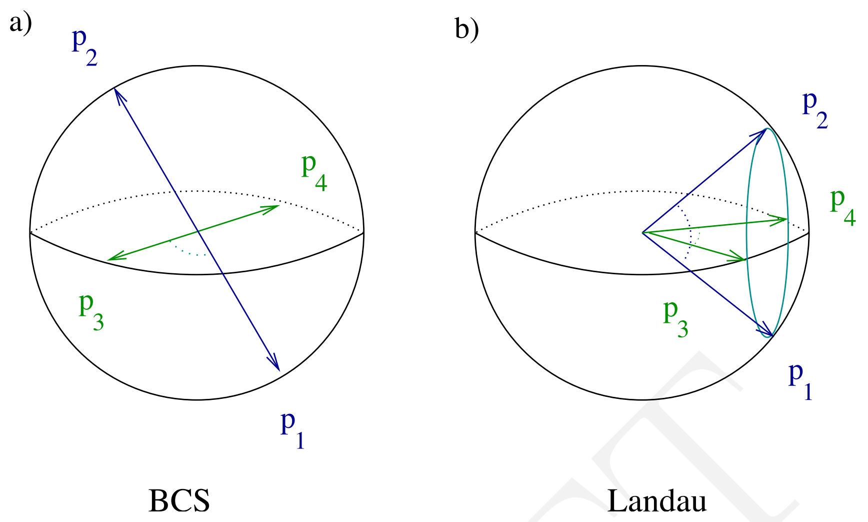

Figure 11: *Fermions scattering near the Fermi surface. (a) Head-on collisions: With* $\mathbf{k}_1+\mathbf{k}_2 = 0$*, only two degrees of freedom in the outgoing momenta* $\mathbf{k}_3$ *and* $\mathbf{k}_4$ *are constrained, as they can point to any two opposite points on the Fermi surface. (b) The 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. [5], 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.

3.3 Relevant operators and naturalness

We saw earlier that a scalar field mass typically gets large 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 a Higgs with a weak scale mass...a natural size for the Higgs mass^2 would be $\alpha/4\pi\Lambda^2$. That makes people optimistic that the cutoff in the SM is at only a few TeV, above which scale new physics is operative, such as SUSY or some new strong interactions. Another relevant operator that is problematic is the unit operator, which is called the vacuum energy or cosmological constant. Loop corrections to this are expected to be $O(\Lambda^4)$ in an EFT with cutoff Λ. Fermion masses look relevant and therefore problematic, but typically are not because of chiral symmetry (see problem III.1).

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 anthropoid 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 [6], and anthropic arguments for the axion in an inflationary universe (see [7] and references therein).

Another idea is that the world is fine-tuned because of its dynamical evolution. A very creative idea along these lines recently appeared in [8]. See also the parable I have reprinted below from my 1997 TASI lectures; I wrote that after being at a conference where I thought that SUSY advocates were being unreasonably smug!

3.4 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...

3.5 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?