3. The Renormalisation Group

We've built up the technology of field theory and path integrals, and I've promised you that this is sufficient to understand what happens at a second order phase transition. But so far, we've made little headway. All we've seen is that as we approach the critical point, fluctuations dominate and the Gaussian path integral is no longer a good starting point. We need to take the interactions into account.

Sometimes in physics, you can understand a phenomenon just by jumping in and doing the right calculation. And we will shortly do this, using perturbation theory to understand how the ϕ^4 terms change the critical exponents. However, to really understand second order phase transitions requires something more: we will need to set up the right framework in which to think of physics at various length scales. This set of ideas was developed in the 1960s and 1970s, by people like Leo Kadanoff, Michael Fisher and, most importantly, Kenneth Wilson. It goes by the name of the *Renormalisation Group*.

3.1 What's the Big Idea?

Let's start by painting the big picture. As in the previous section, we're going to consider a class of theories based around a single scalar field $\phi(\mathbf{x})$ in d dimensions. (We will consider more general set-ups in Section 4.) The free energy takes the now familiar form,

$$F[\phi] = \int d^dx \, \left[\frac{1}{2}\nabla\phi\cdot\nabla\phi + \frac{1}{2}\mu^2\phi^2 + g\phi^4 + \ldots\right] \tag{3.1}$$

In what follows, we will look at what happens as we approach the critical point from above $T \to T_c$. All the important temperature dependence in (3.1) is sitting in the quadratic term, with

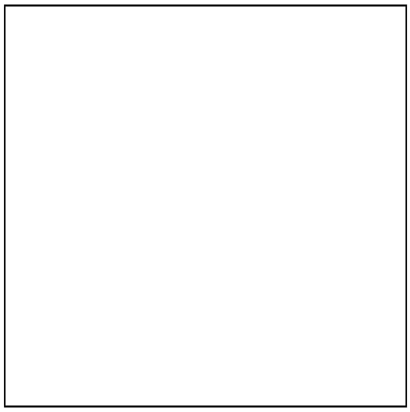
$$\mu^2 \sim T - T_c \tag{3.2}$$

In contrast to the previous section, we will allow μ^2 to take either sign: $\mu^2 > 0$ in the disordered phase, and $\mu^2 < 0$ in the ordered phase where $\langle\phi\rangle \neq 0$.

There is one important change in convention from our earlier discussion: we have rescaled the coefficient of the gradient term to be 1/2; we will see the relevance of this shortly. All other terms have arbitrary coefficients.

So far we've focussed on just a few couplings, as shown in the free energy (3.1). Here we're going to expand our horizons. We'll consider all possible terms in the free energy, subject to a couple of restrictions. We'll insist that the free energy is analytic around $\phi = 0$, so has a nice Taylor expansion, and we will insist on the $\mathbf{Z}_2$ symmetry $\phi \to -\phi$, so that only even powers of ϕ arise. This means, for example, that we will include the term ϕ^6 and $(\nabla^2\phi)^2$ and ϕ^{14} and $\phi^{137}(\nabla\phi \cdot \nabla\phi)\nabla^2\phi$ and so so. Each of these terms comes with its own coupling constant. However, we don't include terms like ϕ^{17} because this violates the $\mathbf{Z}_2$ symmetry, nor $1/\phi^2$ because this is not analytic at $\phi = 0$.

Next, consider the infinite dimensional space, parameterised by the infinite number of coupling constants. We will call this the *theory space* (although I should warn you that this isn't standard terminology). You should have in your mind something like this:



But possibly bigger.

As we've seen, our interest is in computing the partition function

$$Z = \int \mathcal{D}\phi \; e^{-F[\phi]} \tag{3.3}$$

Note that I've written the exponent as e^{-F} rather than $e^{-\beta F}$. This is because the overall power of β does nothing to affect the physics; all the relevant temperature dependence is in the coefficient (3.2) while, for the quantities of interest near the critical point, this overall factor can be set to $\beta \approx 1/T_c$. You can think that we've simply rescaled this into the field ϕ.

There is one more ingredient that we need to make sense of the path integral (3.3). This is the UV cut-off Λ. Recall that, implicit in our construction of the theory is the requirement that the Fourier modes $\phi_{\mathbf{k}}$ vanish for suitably high momenta

$$\phi_{\mathbf{k}} = 0 \quad \text{for } k > \Lambda$$

This arises because, ultimately our spins sit on some underlying lattice which, in turn, was coarse-grained into boxes of size a. The UV cut-off is given by $\Lambda \sim 1/a$.

Until now, the UV cut-off has taken something of a back seat in our story, although it was needed to render some of the path integral calculation in the previous section finite. Now it's time for Λ to move centre stage. As we will explain, we can use the cut-off to define a *flow* in the space of theories.

Suppose that we only care about physics on long distance scales, L. Then we have no real interest in the Fourier modes $\phi_{\mathbf{k}}$ with $k \gg 1/L$. This suggests that we can write down a different theory, that has a lower cut-off,

$$\Lambda' = \frac{\Lambda}{\zeta}$$

for some ζ. As long as $\Lambda' \gg 1/L$, the scale of interest, our theory can tell us everything that we need to know. Moreover, we know, at least in principle, how to construct such a theory. We write our Fourier modes as

$$\phi_{\mathbf{k}} = \phi^-_{\mathbf{k}} + \phi^+_{\mathbf{k}}$$

where $\phi^-_{\mathbf{k}}$ describe the long-wavelength fluctuations

$$\phi^-_{\mathbf{k}} = \begin{cases} \phi_{\mathbf{k}} & k < \Lambda' \\ 0 & k > \Lambda' \end{cases}$$

and $\phi^+_{\mathbf{k}}$ describe the short-wavelength fluctuations that we don't care about

$$\phi^+_{\mathbf{k}} = \begin{cases} \phi_{\mathbf{k}} & \Lambda' < k < \Lambda \\ 0 & \text{otherwise} \end{cases}$$

There are several other names for these variables that are used interchangeably. The modes $\phi^-_{\mathbf{k}}$ and $\phi^+_{\mathbf{k}}$ are also referred to as low- and high-energy modes or, importing language from quantum mechanics, slow and fast modes, respectively. In a rather quaint nod to the electromagnetic spectrum, the short-distance, microscopic physics that we care little about is often called the *ultra-violet*; the long-distance physics that we would like to understand is the *infra-red.*

Similarly, we decompose the free energy, written in Fourier space, as

$$F[\phi_{\mathbf{k}}] = F_0[\phi^-_{\mathbf{k}}] + F_0[\phi^+_{\mathbf{k}}] + F_I[\phi^-_{\mathbf{k}}, \phi^+_{\mathbf{k}}]$$

Here $F_I[\phi^-_{\mathbf{k}}, \phi^+_{\mathbf{k}}]$ involves the terms which mix the short and long-wavelength modes. The partition function (3.3) can then be written as

$$Z = \int \prod_{k<\Lambda} d\phi_{\mathbf{k}} \; e^{-F} = \int \prod_{k<\Lambda'} d\phi^-_{\mathbf{k}} \; e^{-F_0[\phi^-_{\mathbf{k}}]} \int \prod_{\Lambda'<k<\Lambda} d\phi^+_{\mathbf{k}} \; e^{-F_0[\phi^+_{\mathbf{k}}]} e^{-F_I[\phi^-_{\mathbf{k}}, \phi^+_{\mathbf{k}}]}$$

We write this as

$$Z = \int \mathcal{D}\phi^- \; e^{-F'[\phi^-]}$$

where $F'[\phi]$ is known as the *Wilsonian effective free energy.* (In fairness, this term is rarely used: you're more likely to hear "Wilsonian effective action" to describe the analogous object in a path integral describing a quantum field theory.) We're left with a free energy which describes the long-wavelength modes, but takes into account the effects of the short wavelength modes. It is defined by

$$e^{-F'[\phi^-]} = e^{-F_0[\phi^-_{\mathbf{k}}]} \int \prod_{\Lambda'<k<\Lambda} d\phi^+_{\mathbf{k}} \; e^{-F_0[\phi^+_{\mathbf{k}}]} e^{-F_I[\phi^-_{\mathbf{k}}, \phi^+_{\mathbf{k}}]} \tag{3.4}$$

In subsequent sections, we'll put some effort into calculating this object. However, at the end of the day the new free energy $F'[\phi^-]$ must take the same functional form as the original free energy (3.1), simply because we started from the most general form possible. The only effect of integrating out the high-momentum modes is to shift the infinite number of coupling constants, so we now have

$$F'[\phi] = \int d^dx \; \left[\frac{1}{2}\gamma' \nabla\phi \cdot \nabla\phi + \frac{1}{2}\mu'^2\phi^2 + g'\phi^4 + \ldots\right] \tag{3.5}$$

We would like to compare the new free energy (3.5) with the original (3.1). However, we're not quite there yet because, the two theories are different types of objects – like apples and oranges – and shouldn't be directly compared. This is because the theory is defined by both the free energy *and* the UV cut-off and, by construction our two theories have different cut-offs. This means that the original theory $F[\phi]$ can describe things that the new theory $F'[\phi^-]$ cannot, namely momentum modes above the cut-off Λ'.

It is straightforward to remedy this. We can place the two theories on a level playing field by rescaling the momenta in the new theory. We define

$$\mathbf{k}' = \zeta \mathbf{k}$$

Now k' takes values up to Λ, as did k in the original theory. The counterpart of this scaling in real space is

$$\mathbf{x}' = \frac{\mathbf{x}}{\zeta}$$

This means that all lengths scales are getting smaller. You can think of this step as zooming out, to observe the system on larger and larger length scales. As you do so, all features become smaller.

There is one final step that we should take. The new theory $F'[\phi]$ will typically have some coefficient $\gamma' \neq 1$ in front of the leading, quadratic gradient term. To compare with the original free energy (3.1), we should rescale our field. We define

$$\phi'_{\mathbf{k}'} = \sqrt{\gamma'}\,\phi^-_{\mathbf{k}}$$

Which, in position space, reads

$$\phi'(\mathbf{x}') = \sqrt{\gamma'}\,\phi^-(\mathbf{x}) \tag{3.6}$$

Now, finally, our free energy takes the form

$$F_\zeta[\phi'] = \int d^dx\ \left[\frac{1}{2}\nabla\phi'\cdot\nabla\phi' + \frac{1}{2}\mu^2(\zeta)\phi'^2 + g(\zeta)\phi'^4 + \ldots\right] \tag{3.7}$$

We see that this procedure induces a continuous map from $\zeta \in [1,\infty)$ onto the space of coupling constants. Our original coupling constants in (3.1) are those evaluated at $\zeta = 1$. As we increase ζ, we trace out curves in our theory space, that look something like the picture shown in Figure 22

We say that the coupling constants *flow*, where the direction of the flow is telling us what the couplings look like on longer and longer length scales. The equations which describe these flows – which we will derive shortly – are known, for historic reasons, as *beta functions*.

These, then are the three steps of what is known as the *renormalisation group* (RG):

- Integrate out high momentum modes, $\Lambda/\zeta < k < \Lambda$.
- Rescale the momenta $\mathbf{k}' = \zeta\mathbf{k}$.
- Rescale the fields so that the gradient term remains canonically normalised.

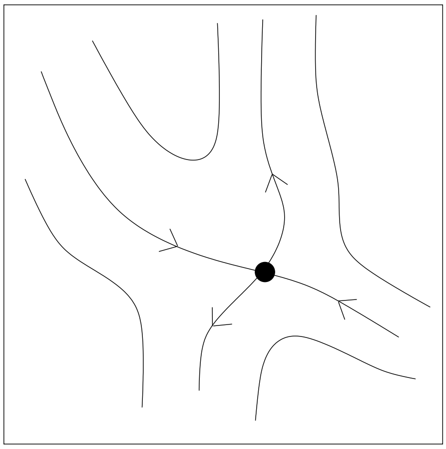


Figure 22: Flows in theory space; the arrows are in the direction of increasing ζ.

You may wonder why we didn't just include a coupling constant $\gamma(\zeta)$ for the gradient term, and watch that change too. The reason is that free energies $F[\phi]$ which differ only by a multiplicative constant have partition functions which differ only by a constant. This means that they describe the same physics and, from our perspective, should be consider equivalent theories. To remove this ambiguity, we should pin down the value of one of the coupling constants, and the gradient term $(\nabla\phi)^2$ is the most convenient choice. If we ever find ourselves in a situation where $\gamma(\zeta) = 0$ for some ζ then we would have to re-evaluate this choice. (We'll actually come across an example where it's sensible to make a different choice in Section 4.3.)

The "renormalisation group" is not a great name. It has a hint of a group structure, because a scaling by ζ_1 followed by a scaling by ζ_2 gives the same result as a scaling by $\zeta_1\zeta_2$. However, unlike for groups, there is no inverse: we can only integrate out fields, we can't put them back in. A more accurate name would be the "renormalisation semi-group".

The Renormalisation Group in Real Space

The procedure we've described above is the renormalisation group in momentum space: to get an increasingly coarse-grained description of the physics, we integrate out successive momentum shells. This version of the renormalisation group is most useful when dealing with continuous fields and will be the approach we will focus on in this course.

There is a somewhat different, although ultimately equivalent, phrasing of the renormalisation group which works directly in real space. This approach works best when dealing directly with lattice systems, like the Ising model. As we explained rather briefly in Section 1.3, one constructs a magnetisation field $m(\mathbf{x})$ by coarse-graining over boxes of size a, each of which contains many lattice sites. One can ask how the free energy changes as we increase a, a procedure known as *blocking*. Ultimately this leads to the same picture that we built above.

3.1.1 Universality Explained

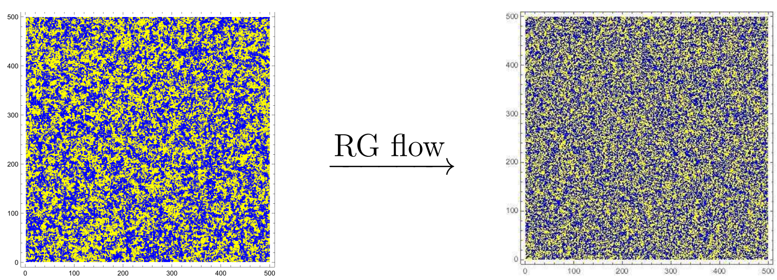
Even before we do any calculations, there are general lessons to be extracted from the framework above. Let's suppose we start from some point in theory space. This can be arbitrarily complicated, reflecting the fact that it contains informations about all the microscopic, short-distance degrees of freedom.

Of course, we care little about most of these details so, in an attempt to simplify our lives we perform a renormalisation group transformation, integrating out short distance degrees of freedom to generate a new theory which describes the long wavelength physics. And then we do this again. And then we do this again. Where do we end up?

There are essentially two possibilities: we could flow off to infinity in theory space, or we could converge towards a *fixed point*. These are points which are invariant under a renormalisation group transformation. (One could also envisage further possibilities, such as converging towards a limit cycle. It turns out that these can be ruled out in many theories of interest.)

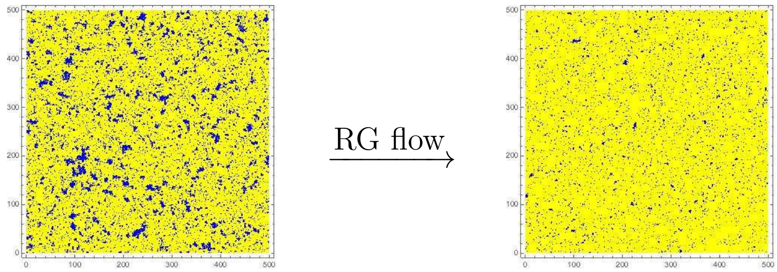
Our interest here lies in the fixed points. The second step in the renormalisation group procedure ensures that fixed points describe theories that have no characteristic scale. If the original theory had a correlation length scale ξ, then the renormalised theory has a length scale $\xi' = \xi/\zeta$. (We will derive this statement explicitly below when we stop talking and start calculating.) Fixed points must therefore have either $\xi = 0$ or $\xi = \infty$.

In the disordered phase, with $T > T_c$, enacting an RG flow reduces the correlation length. Pictorially, we have



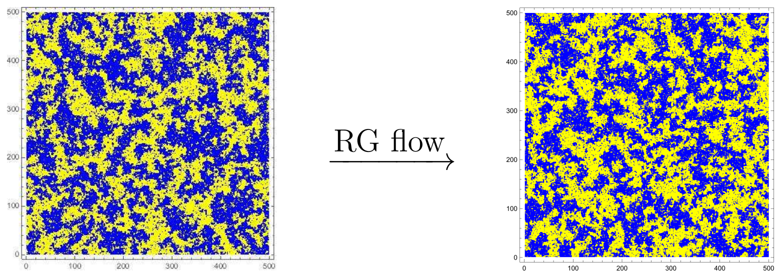
In this case, shrinking the correlation length is equivalent to increasing the temperature. The end point of the RG flow, at $\xi = 0$, is the infinite temperature limit of the theory. This is rather like flowing off to infinity in theory space. As we will see, it is not uncommon to end up here after an RG flow. But it is boring.

Similarly, in the ordered phase the RG flow again reduces the correlation length,



Now the end point at $\xi = 0$ corresponds to the zero temperature limit; again, it is a typical end point of RG flow but is dull.

Theories with $\xi = \infty$ are more interesting. As we saw above, this situation occurs at a critical point where the theory contains fluctuations on all length scales. Now, if we do an RG flow, the theory remains invariant. In terms of our visual configurations,



Note that the configuration itself doesn't stay the same. (It is, after all, merely a representative configuration in the ensemble.) However, as the fluctuations on small distance scales shrink away due to RG, they are replaced by fluctuations coming in from larger distance scales. The result is a theory which is scale invariant. For this reason, the term "critical point" is often used as a synonym for "fixed point" of the RG flow.

This picture is all we need to understand the remarkable phenomenon of universality: it arises because many points in theory space flow to the same fixed point. Thus, many different microscopic theories have the same long distance behaviour.

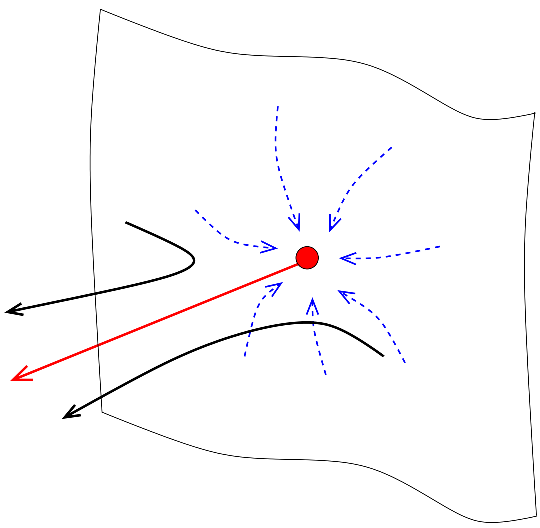


Figure 23: The critical surface. Irrelevant deformations from the critical point are shown as dotted blue lines; the relevant deformation is is red.

Relevant, Irrelevant or Marginal

It is useful to characterise the properties of fixed points by thinking about the theories in their immediate neighbourhood. Obviously, there are an infinite number of ways we can move away from the fixed point. If we move in some of these directions, the RG flow will take us back towards the fixed point. These deformations are called *irrelevant* because if we add any such terms to the free energy we will end up describing the same long-distance physics.

In contrast, there will be some directions in which the RG flow will sweep us away from the fixed point. These deformations are called *relevant* because if we add any such terms to the free energy, the long-distance physics will be something rather different. Examples of relevant and irrelevant deformations are shown in Figure 23. Much of the power of universality comes from the realisation that the vast majority of directions are irrelevant. For a given fixed point, there are typically only a handful of relevant deformations, and an infinite number of irrelevant ones. This means that our fixed points have a large basin of attraction, huge slices of the infinite dimensional theory space all converging to the same fixed point. The basin of attraction for a particular fixed point is called the *critical surface*.

Finally, it's possible that our fixed point is not a point at all, but a line or a higher dimensional surface living within theory space. In this case, if we deform the theory in the direction of the line, we will not flow anywhere, but simply end up on another fixed point. Such deformations are called *marginal*; they are rare, but not unheard of.

Why High Energy Physics is Hard

Universality is a wonderful thing if you want to understand the low-energy, long-wavelength physics. It tells you that you can throw away many of the microscopic details because they are irrelevant for the things that you care about.

In contrast, if you want to understand the high-energy, short distance physics then universality is the devil. It tells you that you have very little hope of extracting any information about microscopic degrees of freedom if you only have access to information at long distances. This is because many different microscopic theories will all give the same answer.

As we saw in Section 2.3, quantum field theory is governed by the same mathematical structure as statistical field theory, and the comments above also apply. Suppose, for example, that you find yourself living in a technologically adolescent civilisation that can perform experiments at distance scales of 10^{-16} cm or so, but no smaller. Yet, what you really care about is physics at, say, 10^{-32} cm where you suspect that something interesting is going on. The renormalisation group says that you shouldn't pin your hopes on learning anything from experiment.

The renormalisation group isn't alone in hiding high-energy physics from us. In gravity, cosmic censorship ensures that any high curvature regions are hidden behind horizons of black holes while, in the early universe, inflation washes away any trace of what took place before. Anyone would think there's some kind of conspiracy going on....

3.2 Scaling

The idea that second order phase transitions coincide with fixed points of the renormalisation group is a powerful one. In particular, it provides an organising principle behind the flurry of critical exponents that we met in Section 1.

As we explained above, at a fixed point of the renormalisation group any scale must be washed away. This is already enough to ensure that correlation functions must take the form of a power-law,

$$\langle \phi(\mathbf{x})\phi(0)\rangle \sim \frac{1}{r^{d-2+\eta}} \tag{3.8}$$

Any other function would require a scale on dimensional grounds. The only freedom that we have is in the choice of exponent which we have chosen to parameterise as η. One of the tasks of the RG procedure is to compute η, and we will see how this works in Section 3.5.

However, even here there's something of a mystery because usually we can figure out the way things scale by doing some simple dimensional analysis. (If you would like to refresh your memory, some examples of dimensional analysis can be found in Chapter 3 of the lectures on Dynamics and Relativity.) What does that tell us in the present case?

We will measure dimension in units of inverse length. So, for example, $[x] = -1$ while $[\partial/\partial x] = +1$. The quantity $F[\phi]$ must be dimensionless because it sits in the exponent of the partition function as e^{-F}. The first term is

$$F[\phi] = \int d^dx \; \frac{1}{2}\nabla\phi \cdot \nabla\phi + \dots$$

From this we learn that

$$[\phi] = \frac{d-2}{2} \tag{3.9}$$

Which, in turn, tells us exactly what the exponent of the correlation function must be: $\eta = 0$.

This is sobering. Dimensional analysis is one of the most basic tools that we have, and yet it seems to be failing at critical points where experiment is showing that $\eta \neq 0$. What's going on?

A better way to think about dimensional analysis is to think in terms of scaling. Suppose that we rescale all length as $\mathbf{x} \to \mathbf{x}' = \mathbf{x}/\zeta$. How should other quantities scale so that formula remains invariant? Stated this way, it's clear that there's a close connection between dimensional analysis and RG. The correlation function (3.8) is telling us that we should rescale $\phi(\mathbf{x}) \to \phi'(\mathbf{x}') = \zeta^{\Delta_\phi}\phi(\mathbf{x})$, where

$$\Delta_\phi = \frac{d-2+\eta}{2} \tag{3.10}$$

This is called the *scaling dimension*. It differs from the naive "engineering dimension" $[\phi]$ by the extra term $\eta/2$ which is referred to as the *anomalous dimension*.

We still haven't yet explained why the scaling dimension differs from engineering dimension. The culprit turns out to be the third step of the RG procedure (3.6) where the field ϕ gets rescaled. In real space, this is viewed as coarse-graining ϕ over blocks of larger and larger size a. As we do so, it dresses ϕ with this UV cut-off scale $\Lambda \sim 1/a$, often in a complicated and non-intuitive way. This means that the correlation function (3.8) is actually

$$\langle \phi(\mathbf{x})\phi(0)\rangle \sim \frac{a^\eta}{r^{d-2+\eta}}$$

which is in full agreement with naive dimensional analysis. We can work with usual engineering dimensions if we keep track of this microscopic distance scale a. But it is much more useful to absorb this into ϕ and think of a coarse-grained observable, with dimension Δ_ϕ, that is the appropriate for measuring long distance correlations.

3.2.1 Critical Exponents Revisited

The critical exponents that we met in Section 1.2.3 are all a consequence of scale invariance, and dimensional analysis based on the scaling dimension. Let's see how this arises.

We know that as we move away from the critical point by turning on $\mu^2 \sim T - T_c$, we introduce a new length scale into the problem. This is the correlation length, given by

$$\xi \sim t^{-\nu} \quad \text{with} \quad t = \frac{|T - T_c|}{T_c} \tag{3.11}$$

Here t is called the *reduced temperature*, while ν is another critical exponent that we will ultimately have to calculate. Since ξ is a length scale, it transforms simply as $\xi \to \xi/\zeta$. In other words, it has scaling dimension $\Delta_\xi = -1$. The meaning of the critical exponent ν is that the reduced temperature scales as $t \to \zeta^{\Delta_t} t$, with

$$\Delta_t = \frac{1}{\nu} \tag{3.12}$$

In what follows, our only assumption is that the correlation length ξ is the only length scale that plays any role.

We start with the thermodynamic free energy, $F_{\text{thermo}}(t)$, evaluated at $B = 0$. This takes the form

$$F_{\text{thermo}}(t) = \int d^dx \ f(t)$$

Because $F_{\rm thermo}$ is scale invariant at the fixed point, $f(t)$ must have scaling dimension d, which immediately tells us that

$$f(t) \sim t^{d\nu}$$

There is an intuitive way to understand this. At T close to T_c, the spins are correlated over distances scales ξ, and can be viewed as moving as one coherent block. The free energy $F_{\rm thermo}$ is extensive, and so naturally scales as $F \sim (L/\xi)^d \sim t^{d\nu}$.

From the thermodynamic free energy, we can compute the singular contribution to the heat capacity near $t = 0$. It is

$$c = \frac{\partial^2 f}{\partial t^2} \sim t^{d\nu - 2} \sim t^{-\alpha}$$

where the second relationship is there to remind us that we already had a name for the critical exponent related to heat capacity. We learn that

$$\alpha = 2 - d\nu \tag{3.13}$$

This is called the *Josephson relation* or, alternatively, the *hyperscaling* relation.

The next critical exponent on the list is β. Recall that this relates the magnetisation in the ordered phase – which we used to call m and have now called ϕ – to the temperature as

$$\phi \sim t^\beta$$

But the scaling dimensions of this equation only work if we have

$$\Delta_\phi = \beta \Delta_t \quad \Rightarrow \quad \beta = \nu \Delta_\phi = \frac{(d - 2 + \eta)\nu}{2} \tag{3.14}$$

The next two critical exponents require us to move away from the critical point by turning on a magnetic field B. This is achieved through the addition of a linear term $\int d^dx \; B\phi$ in the free energy. (We didn't include such a linear term in our previous discussion of RG, but it can be added without changing the essence of the story.) The scaling dimensions of this term must add to zero, giving

$$\Delta_\phi + \Delta_B = d \quad \Rightarrow \quad \Delta_B = \frac{d + 2 - \eta}{2}$$

Now we can look at the various relationships. The behaviour of the susceptibility near the critical point is

$$\chi = \left.\frac{\partial \phi}{\partial B}\right|_T \sim t^{-\gamma}$$

Once again, the scaling dimensions are enough to fix γ to be

$$\Delta_\phi - \Delta_B = -\frac{\gamma}{\nu} \quad \Rightarrow \quad \gamma = \nu(2-\eta) \tag{3.15}$$

which is sometimes called *Fisher's identity.* Once again, there is an intuitive way to understand this. The meaning of ξ is that the spins are no longer correlated at distances $r \gg \xi$. This can be seen, for example, in our original formula (2.29). Using our earlier expression (2.28) for the susceptibility, we have

$$\chi \sim \int_0^\xi d^dx \, \frac{1}{r^{d-2+\eta}} \sim \xi^{2-\eta} \sim t^{-\nu(2-\eta)}$$

which again gives $\gamma = \nu(2-\eta)$.

The final critical exponent relates the magnetisation ϕ to the magnetic field B when we sit at the critical temperature $t = 0$. It should come as little surprise by now to learn that this is again fixed by scaling analysis

$$\phi \sim B^{1/\delta} \quad \Rightarrow \quad \delta = \frac{\Delta_B}{\Delta_\phi} = \frac{d+2-\eta}{d-2+\eta} \tag{3.16}$$

We end up with four equations, relating α (3.13), β (3.14), γ (3.15) and δ (3.16) to the critical exponents η and ν. For convenience, let's recall what values we claimed this exponents take:

	α	β	γ	δ	η	ν
MF	$\frac{4-d}{2}$	$\frac{1}{2}$	1	3	0	$\frac{1}{2}$
$d=2$	0	$\frac{1}{8}$	$\frac{7}{4}$	15	$\frac{1}{4}$	1
$d=3$	0.1101	0.3264	1.2371	4.7898	0.0363	0.6300

where we've used the result (2.14), including quadratic fluctuations, for the mean field value of α. We see that the relations are satisfied exactly for $d = 2$ and to within the accuracy stated for $d = 3$. However, there's a wrinkle because they only agree with the mean field values when $d = 4$!

This latter point is an annoying subtlety and will be explained in Section 3.3.2. Our main task is to understand why the mean field values don't agree with experiment when $d < 4$.

3.2.2 The Relevance of Scaling

The kind of dimensional analysis above also determines whether a given interaction is relevant, irrelevant or marginal.

Consider an interaction term $\mathcal{O}(\mathbf{x})$ in the free energy,

$$F[\phi] \sim \int d^d x \; g_{\mathcal{O}} \, \mathcal{O}(\mathbf{x}) \tag{3.17}$$

Here $\mathcal{O}$ can be ϕ^n or $\phi^m(\nabla\phi)^2$ or any of the other infinite possibilities. In a spillover from quantum field theory, the different interaction terms $\mathcal{O}(\mathbf{x})$ are referred to as *operators.*

We're interested in operators which, in the vicinity of a given point, transform simply under RG. Specifically, suppose that, under the rescaling $\mathbf{x} \to \mathbf{x}' = \mathbf{x}/\zeta$, the operator has a well defined scaling dimension, transforming as

$$\mathcal{O}(\mathbf{x}) \to \mathcal{O}'(\mathbf{x}') = \zeta^{\Delta_{\mathcal{O}}} \mathcal{O}(\mathbf{x}) \tag{3.18}$$

You can think of such operators as eigenstates of the RG process. From the free energy (3.17), the scaling dimension of the coupling is

$$\Delta_{g_{\mathcal{O}}} = d - \Delta_{\mathcal{O}}$$

Under an RG flow, these couplings scale as $g_{\mathcal{O}} \to \zeta^{d-\Delta_{\mathcal{O}}} g_{\mathcal{O}}$. We can see immediately that $g_{\mathcal{O}}$ either diverges or vanishes as we push forwards with the RG. Invoking our previous classification, $\mathcal{O}$ is:

- Relevant if $\Delta_{\mathcal{O}} < d$
- Irrelevant if $\Delta_{\mathcal{O}} > d$
- Marginal if $\Delta_{\mathcal{O}} = d$

The tricky part of the story is that it's not always easy to identify the operators $\mathcal{O}$ which have the nice scaling property (3.18). As we'll see in the examples below, these are typically complicated linear combinations of the operators ϕ^n and $\phi^n(\nabla\phi)^2$ and so on.

3.3 The Gaussian Fixed Point

It's now time to start calculating. We will start by sitting at a special point in theory space and enacting the renormalisation group. At this special point, only two quadratic terms are turned on:

$$F_0[\phi] = \int d^d x \; \left[\frac{1}{2}\nabla\phi\cdot\nabla\phi + \frac{1}{2}\mu_0^2\,\phi^2\right] = \int^\Lambda \frac{d^d k}{(2\pi)^d}\,\frac{1}{2}(k^2+\mu_0^2)\phi_{\mathbf{k}}\phi_{-\mathbf{k}} \tag{3.19}$$

where we've added a subscript to the coefficient μ_0^2 in anticipation the fact that this quantity will subsequently change under RG flow.

Because the free energy is quadratic in ϕ, it has the property that there is no mixing between the short and long wavelength modes, and so factorises as

$$F_0[\phi] = F_0[\phi^-] + F_0[\phi^+]$$

Integrating over the short wavelength modes is now easy, and results in an overall constant in the partition function

$$e^{-F'[\phi^-]} = \left[\int \mathcal{D}\phi^+ e^{-F_0[\phi^+]}\right] e^{-F_0[\phi^-]} = \mathcal{N} e^{-F_0[\phi^-]}$$

This constant $\mathcal{N}$ doesn't change any physics; it just drops out when we differentiate $\log Z$ to compute correlation functions. However, we're not yet done with the our RG; we still need to do the rescaling

$$\mathbf{k}' = \zeta \mathbf{k} \quad \text{and} \quad \phi'_{\mathbf{k}'} = \zeta^{-w}\, \phi^-_{\mathbf{k}} \tag{3.20}$$

where w is constant that we will determine. Written in terms of the rescaled momenta, we have

$$\begin{aligned} F_0[\phi^-] &= \int^{\Lambda/\zeta} \frac{d^dk}{(2\pi)^d}\, \frac{1}{2}(k^2 + \mu_0^2)\phi^-_{\mathbf{k}}\phi^-_{-\mathbf{k}} \\ &= \int^{\Lambda} \frac{d^dk'}{(2\pi)^d}\, \frac{1}{2\zeta^d}\left(\frac{k'^2}{\zeta^2} + \mu_0^2\right) \zeta^{2w} \phi'_{\mathbf{k}'}\phi'_{-\mathbf{k}'} \end{aligned}$$

We can put this back in the form we started with if we take

$$w = \frac{d+2}{2} \tag{3.21}$$

leaving us with

$$F'_0[\phi'] = \int^{\Lambda} \frac{d^dk}{(2\pi)^d}\, \frac{1}{2}(k^2 + \mu^2(\zeta))\phi'_{\mathbf{k}}\phi'_{-\mathbf{k}}$$

The only price that we've paid for this is that the coefficient of quadratic term has become

$$\mu^2(\zeta) = \zeta^2 \mu_0^2 \tag{3.22}$$

This illustrates how the length scales in the problem transform under RG. Recall that the correlation length (2.22) is $\xi^2 \sim 1/\mu^2$. We see that, under an RG procedure,

$$\xi \to \frac{\xi}{\zeta}$$

The fixed points obey $d\mu^2/d\zeta = 0$. As we anticipated previously, there are two of them. The first is $\mu^2 = \infty$. This corresponds to a state which has infinite temperature. It is not where our interest lies. The other fixed point is at $\mu^2 = 0$. This is known as the *Gaussian fixed point*.

3.3.1 In the Vicinity of the Fixed Point

As we mentioned previously, we would like to classify fixed points by thinking about what happens when you sit near them. Do you flow into the fixed point, or get pushed away?

We already have the answer to this question in one direction in coupling space. If we add the term $\mu^2\phi^2$, the scaling (3.22) tells us that μ^2 gets bigger as we flow towards the infra-red. This is an example of a relevant coupling: turning it on pushes us away from the fixed point.

Here is another example: it is simple to repeat the steps above including the term $\alpha_0(\nabla^2\phi)^2$ in the free energy. Upon RG, this coupling flows as $\alpha(\zeta) = \zeta^{-2}\alpha_0$. It is an example of an irrelevant coupling, one which becomes less important as we flow towards the infra-red.

More interesting are the slew of possible couplings of the form

$$F[\phi] = \int d^dx \, \left[\frac{1}{2}\nabla\phi\cdot\nabla\phi + \frac{1}{2}\mu_0^2\phi^2 + \sum_{n=4}^{\infty} g_{0,n}\phi^n\right] \tag{3.23}$$

where, to keep the $\mathbf{Z}_2$ symmetry, we restrict the sum to n even. Here things are a little more subtle because, once we turn these couplings, on the first step of the RG procedure is no longer so simple. Integrating out the short distance modes will shift each of these couplings,

$$g_{0,n} \to g'_n = g_{0,n} + \delta g_n$$

We will learn how to calculate the δg_n in section 3.4. But, for now, let's ignore this effect and concentrate on the second and third parts of the RG procedure, in which we rescale lengths and fields as in (3.20). In this approximation, the operators ϕ^n enjoy the nice scaling property (3.18),

$$\mathbf{x}' = \mathbf{x}/\zeta \quad \text{and} \quad \phi'(\mathbf{x}') = \zeta^{\Delta_\phi}\phi(\mathbf{x})$$

The free energy is then rescaled by

$$F[\phi'] = \int d^dx' \; \zeta^d \left[\frac{1}{2}\zeta^{-2-2\Delta_\phi}\nabla'\phi'\cdot\nabla'\phi' + \frac{1}{2}\mu_0^2\zeta^{-2\Delta_\phi}\phi'^2 + \sum_{n=4}^{\infty} g_{0,n}\zeta^{-n\Delta_\phi}\phi'^n\right]$$

To restore the coefficient of the gradient term, we pick the scaling dimension

$$\Delta_\phi = \frac{d-2}{2}$$

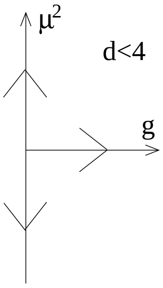


Figure 24: RG flows when $d < 4$

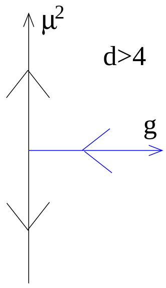


Figure 25: RG flows when $d > 4$

For once, the scaling dimension coincides with the engineering dimension (3.9): $\Delta_\phi = [\phi]$. This is because we're looking at a particularly simple fixed point. Note that this is related to our earlier result (3.21) by $\Delta_\phi = d - w$, with the extra factor of d coming from the $\int d^dk$ in the definition of the Fourier transform.

Our free energy now takes the same form as before,

$$F[\phi'] = \int d^dx' \left[\frac{1}{2}\nabla'\phi' \cdot \nabla'\phi' + \frac{1}{2}\mu^2(\zeta)\phi'^2 + \sum_{n=4}^{\infty} g_n(\zeta)\phi'^n\right]$$

where

$$g_n(\zeta) = \zeta^{d-n\Delta_\phi} g_{0,n} = \zeta^{(1-n/2)d+n} g_{0,n} \tag{3.24}$$

We see that the way these coupling scale depends on the dimension d. For example, the coefficient for ϕ^4 scales as

$$g_4(\zeta) = \zeta^{4-d} g_{0,4}$$

We learn that ϕ^4 is irrelevant for $d > 4$ and is relevant for $d < 4$. According to the analysis above, when $d = 4$, we have $g_4(\zeta) = g_{0,4}$ and the coupling is marginal. In this case, however, we need to work a little harder because the leading contribution to the scaling will come from the corrections δg_4 that we neglected. We'll look at this in the next section.

Restricting to the plane of couplings parameterised by μ^2 and g_4, we see that (if we neglect the interactions) the RG flow near the origin is very different when $d > 4$ and $d < 4$. These are shown in the two figures. In the former case, we need to tune only $\mu^2 \sim T - T_c$ if we want to hit the fixed point; the other couplings will take care of themselves. In contrast, when $d < 4$ both of these couplings are relevant. This means that we would need to tune both to zero if we want to hit the Gaussian fixed point.

We can tally this with our discussion in Section 3.2. The fact that the scaling dimension Δ_ϕ coincides with the naive engineering dimension $[\phi]$ immediately tells us that $\eta = 0$. Meanwhile, the scaling of $\mu^2 \sim t$ is given by $\Delta_t = [g_2] = 2$, which tells us that $\nu = 1/2$. From this we can use (3.13) - (3.16) to extract the remaining critical exponents. These agree with mean field for $d = 4$, but not for $d \leq 4$. (We will address the situation in $d > 4$ in Section 3.3.2.)

It is no coincidence that this behaviour switches at $d = 4$, which we previously identified as the upper critical dimension. In an experiment, one can always change μ^2 by varying the temperature. However, one may not have control over the ϕ^4 couplings which typically correspond to some complicated microscopic property of the system. If ϕ^4 is irrelevant, we don't care: the system will drive itself to the Gaussian fixed point. In contrast, if ϕ^4 is relevant the system will drive itself elsewhere. This is why we don't measure mean field values for the critical exponents: these are the critical exponents of the Gaussian fixed point.

The coupling for ϕ^6 scales as

$$g_6(\zeta) = \zeta^{6-2d} g_{0,6}$$

This is irrelevant in $d > 3$, relevant in $d < 3$ and, naively, marginal in $d = 3$.

Note that in dimension $d = 2$ all of the couplings $g_n \phi^n$ are relevant.

So far, this all looks rather trivial. However, things become much more interesting at other fixed points. For example, around most fixed points $\Delta_{\phi^n} \neq n\Delta\phi$. Indeed, around most fixed points neither ϕ nor ϕ^n will have well defined scaling dimension; instead those operators to which one can assign a scaling dimension consist of some complicated linear combination of the ϕ^n. We will start to understand this better in Section 3.4.

The Meaning of Mean Field

The meaning of the phrase "mean field theory" has evolved as these lectures have progressed. We started in Section 1.1.2 by introducing mean field as a somewhat dodgy approximation to the partition function. Subsequently, we used the expression "mean field theory" to mean writing down a free energy $F[\phi]$ and focussing on the saddle point equations. This saddle point is a good approximation to the partition function only when the couplings are small; this is true only in the vicinity of the Gaussian fixed point. For this reason, using mean field theory is usually synonymous with working at the Gaussian fixed point, and ignoring any relevant operators like ϕ^4.

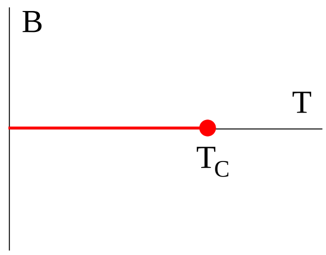


Figure 26: The phase diagram of the Ising model.

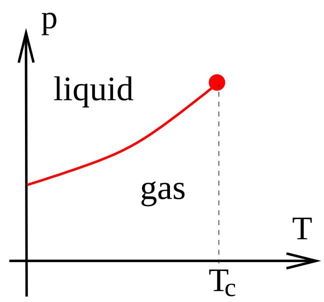


Figure 27: The phase diagram of the liquid-gas system.

Interactions that Break $\mathbf{Z}_2$ Symmetry

Until now, we have have restricted ourselves to interactions ϕ^n with n even, to zealously safeguard the $\mathbf{Z}_2$ symmetry $\phi \to -\phi$. One particularly nice aspect of RG is that if we restrict ourselves to a class of free energies that obey a certain symmetry, then we will remain in that class under RG. We'll see examples of this in Section 3.4.

However, suppose that we sit outside of this class and turn on interactions ϕ^n with n odd. The leading order effect is the magnetic field $B\phi$ that we included in our original Ising model. This is always a relevant interaction. This means that if we want to hit the critical point, we must tune this to zero.

It may be more natural to tune $B = 0$ in some systems that others. For example, a magnet in the Ising class automatically has $B = 0$ unless you choose to submit it to a background magnetic field. This means that it's easy to hit the critical point: just heat up a magnetic and it will exhibit a second order phase transition.

In contrast, in the liquid gas system, setting "$B = 0$" is less natural. Unlike in the Ising model, there is no $\mathbf{Z}_2$ symmetry manifest in the microscopic physics of gases. Instead, it is an emergent symmetry which relates the density of liquid and gas states at the phase transition. Correspondingly, if we simply take a liquid and heat it up then we're most likely to encounter a first order transition, or no transition at all. If we want to hit the critical point, we must now tune the two relevant operators: temperature μ^2 and pressure, which corresponds to the linear term with coefficient B.

In both situations above we really need to tune two relevant couplings to zero to hit the critical point. Of these, one is even under $\mathbf{Z}_2$ and one is odd under $\mathbf{Z}_2$. Doing this will allow us to hit a fixed point with two relevant deformations, one even one odd. This is the Gaussian fixed point in $d > 4$ and is something else (to be described below) in $d < 4$.

What about higher order interactions ϕ^n with n odd. If we have to tune ϕ, do we not also need to tune ϕ^3? It turns out that that the ϕ^3 interaction is redundant. If you have a free energy with no $\mathbf{Z}_2$ symmetry, and all powers of ϕ^n, then you can always redefine your field as $\phi \to \phi + c$ for some constant c. This freedom allows you to eliminate the ϕ^3 term. Note that if your free energy enjoys the $\mathbf{Z}_2$ symmetry $\phi \to -\phi$ then it prohibits you from making this shift.

3.3.2 Dangerously Irrelevant

We've learned that the ϕ^4 interaction is irrelevant for $d \geq 4$, and so one can hit the Gaussian fixed point by tuning just one parameter: $\mu^2 = 0$.

However, there's one tricky issue that we haven't yet explained: the mean field exponents agree with the scaling analysis of Section 3.2 only when $d = 4$. Comparing the two results, we have

	α	β	γ	δ	η	ν
MF	$\frac{4-d}{2}$	$\frac{1}{2}$	1	3	0	$\frac{1}{2}$
Scaling	$\frac{4-d}{2}$	$\frac{d-2}{4}$	1	$\frac{d+2}{d-2}$	0	$\frac{1}{2}$

where we've used the result (2.14), including quadratic fluctuations, for the mean field value of α. This agrees with the scaling analysis. However, for $d > 4$, the exponents β and δ differ. It turns out that the results from Landau mean field are correct, and those from the scaling analysis are wrong. Why?

To understand this, let's recall our scaling argument from Section 3.2. We set $B = 0$ and focus on the critical exponent β. The magnetisation scales with the temperature $t = |T - T_c|/T_c$ as

$$m \sim t^\beta$$

Here m is identified with the scalar field ϕ. Scaling analysis gives $\Delta_\phi = \beta \Delta_t$. But both mean field and scaling analysis agree that $\Delta_t = 1/\nu = 2$ and $\Delta_\phi = (d-2)/2$, and this gives $\beta = (d-2)/4$, rather than the mean field result $\beta = 1/2$.

However, we were a little quick in the scaling analysis because we neglected the quartic coupling g_4. Mean field really told us (1.31),

$$m_0 \sim \left(\frac{t}{g_4}\right)^{1/2}$$

But both t and g_4 scale under RG flow. The scaling dimension of g_4 is $\Delta_{g_4} = 4 - d$ and now the mean field result, with $\beta = 1/2$ is fully compatible with scaling.

There's a more general lesson to take from this. It is tempting, when doing RG, to think that we can just neglect the irrelevant operators because their coefficients flow to zero as we approach the infra-red. However, sometimes we will be interesting in quantities – such as the magnetisation above – which have the irrelevant coupling constants sitting in the denominator. In this case, one cannot just blindly ignore these irrelevant couplings as they affect the scaling analysis. When this happens, the irrelevant coupling is referred to as *dangerously irrelevant*.

3.3.3 An Aside: The Emergence of Rotational Symmetry

This is a good point to revisit an issue that we previously swept under the rug. We started our discussion with a lattice model, but very quickly moved to the continuum, field theory. Along the way we stated, without proof, that we expect the long distance physics to enjoy rotational invariance and we restricted our attention to field theories with this property. Why are we allowed to do this?

To make the discussion concrete, consider a square lattice in $d = 2$ dimensions. This has a discrete $\mathbf{Z}_4$ rotational symmetry, together with a $\mathbf{Z}_2$ reflection symmetry. These combine together into the dihedral group D_8.

Our field theory description will respect the D_8 symmetry of the underlying lattice model, together with the $\mathbf{Z}_2$ symmetry $\phi \to -\phi$ which ensures that fields come in pairs. But this would appear to be much less powerful than the full $O(2)$ continuous rotation and reflection symmetry. Have we cheated?

Let's see what kind of terms we might expect. First, there are some simple terms that are prohibited by the dihedral symmetry. For example a lone term $(\partial_1\phi)^2$ would break the $x_1 \to x_2$ discrete rotational symmetry and so would not appear in the free energy. Similarly, a term $\phi\partial_1\phi$ breaks the $x_1 \to -x_1$ symmetry. (On top of this, it is also a total derivative and so doesn't contribute to the free energy.) The lowest dimension term that includes derivatives and is compatible with the discrete symmetry is

$$\mathcal{O}_2 \sim (\partial_1\phi)^2 + (\partial_2\phi)^2$$

But this term happens to be invariant under the full, continuous $O(2)$ rotational symmetry. We should keep going. The first term which preserves D_8, but not $SO(2)$, is

$$\mathcal{O}_4 \sim \phi\partial_1^4\phi + \phi\partial_2^4\phi$$

There is no reason not to add such terms to the free energy and, in general, we expect that these will be present in any field theoretic description that accurately describes the microscopic physics. However, this operator has dimension $\Delta_{\mathcal{O}_4} = d+2$ and so is irrelevant. This means that it gets washed away by the renormalisation group, and the long wavelength physics is invariant under the full $O(2)$ symmetry. We say that the continuous rotational symmetry is *emergent*.

3.4 RG with Interactions

The previous section left two questions hanging. What happens to the renormalisation of the coupling g_4 in $d=4$ dimensions? And where does the flow of g_4 take us in $d<4$ dimensions? In this section we will answer the first of these. In Section 3.5 we will see that our analysis also contains the answer to the second.

We now repeat our RG procedure, but with a different starting point in theory space,

$$F[\phi] = \int d^dx \ \left[\frac{1}{2}\nabla\phi\cdot\nabla\phi + \frac{1}{2}\mu_0^2\phi^2 + g_0\phi^4\right]$$

The renormalisation group procedure tells us to split the Fourier modes of the field into long and short wavelengths,

$$\phi_{\mathbf{k}} = \phi^-_{\mathbf{k}} + \phi^+_{\mathbf{k}} \tag{3.25}$$

and write the free energy as

$$F[\phi] = F_0[\phi^-] + F_0[\phi^+] + F_I[\phi^-,\phi^+]$$

where we take $F_0[\phi]$ to coincide with the quadratic terms (3.19), and the interaction terms are

$$F_I[\phi] = \int d^dx \ g\phi^4$$

Note that we've chosen to include, for example, $(\phi^-)^4$ in the interaction terms rather than F_0. This is a matter of convention.

The effective free energy for $\phi^-_{\mathbf{k}}$, defined in (3.4), is given by

$$e^{-F'[\phi^-]} = e^{-F_0[\phi^-_{\mathbf{k}}]}\int \mathcal{D}\phi^+_{\mathbf{k}} \ e^{-F_0[\phi^+_{\mathbf{k}}]}e^{-F_I[\phi^-_{\mathbf{k}},\phi^+_{\mathbf{k}}]}$$

There is a nice interpretation of this functional integral $\int\mathcal{D}\phi^+_{\mathbf{k}}$. We can think of it as computing the expectation value of $e^{-F_I[\phi^-_{\mathbf{k}},\phi^+_{\mathbf{k}}]}$, treating $\phi^+_{\mathbf{k}}$ as the random variable with Gaussian distribution $e^{-F_0[\phi^+_{\mathbf{k}}]}$. In other words, we can write this as

$$e^{-F'[\phi^-]} = e^{-F_0[\phi^-_{\mathbf{k}}]}\left\langle e^{-F_I[\phi^-_{\mathbf{k}},\phi^+_{\mathbf{k}}]}\right\rangle_+$$

where the subscript on $\langle\cdot\rangle_+$ is there to remind us that we are averaging over the $\phi^+_{\mathbf{k}}$ modes only. We take the definition of the path integral to be suitably normalised so that $\langle 1\rangle_+ = 1$. Taking the log of both sides,

$$F'[\phi^-] = F_0[\phi^-_{\mathbf{k}}] - \log\left\langle e^{-F_I[\phi^-_{\mathbf{k}},\phi^+_{\mathbf{k}}]}\right\rangle_+ \tag{3.26}$$

Our task is to compute this.

We can't do this functional integral exactly. Instead, we resort to perturbation theory. We assume that g_0 is suitably small, and expand. (The dimensionless small parameter is $g_0\mu_0^{d-4}$.) We first Taylor expand the exponential,

$$\log\left\langle e^{-F_I[\phi^-_{\mathbf{k}},\phi^+_{\mathbf{k}}]}\right\rangle_+ = \log\left\langle 1 - F_I + \frac{1}{2}F_I^2 + \ldots\right\rangle_+$$

and we then Taylor expand $\log(1+x)$, to get

$$\log\left\langle e^{-F_I[\phi^-_{\mathbf{k}},\phi^+_{\mathbf{k}}]}\right\rangle_+ = -\left\langle F_I\right\rangle_+ + \frac{1}{2}\left[\left\langle F_I^2\right\rangle_+ - \left\langle F_I\right\rangle_+^2\right] + \ldots \tag{3.27}$$

where, in general, the n^{th} term is $(-1)^n \times n^{\text{th}}$ cumulant of F_I. This also follows from the same kind of manipulations that we did at the beginning of Section 2.2. We will deal with each of terms above in turn.

3.4.1 Order g_0

At leading order in g_0, we need to compute $\left\langle F_I[\phi^-_{\mathbf{k}},\phi^+_{\mathbf{k}}]\right\rangle_+$. The first order of business is to expand the interaction terms (3.26) in Fourier modes. We have

$$F_I[\phi^-_{\mathbf{k}},\phi^+_{\mathbf{k}}] = g_0 \int \prod_{i=1}^{4} \frac{d^dk_i}{(2\pi)^d} \times (\text{terms with } \phi) \times (2\pi)^d\delta^d(\sum_i \mathbf{k}_i)$$

There are five different "terms with ϕ", most of which do not give anything interesting. These five terms are:

i) $\phi^-_{\mathbf{k}_1}\phi^-_{\mathbf{k}_2}\phi^-_{\mathbf{k}_3}\phi^-_{\mathbf{k}_4}$: This term doesn't include any $\phi^+_{\mathbf{k}}$, the average is trivial. It carries over to give the term $g_0\int d^dx\ \phi^{-\,4}$ in the effective free energy.

ii) $4\phi^-_{\mathbf{k}_1}\phi^-_{\mathbf{k}_2}\phi^-_{\mathbf{k}_3}\phi^+_{\mathbf{k}_4}$: This term has just a single $\phi^+_{\mathbf{k}}$ and so vanishes when averaged over the Gaussian ensemble.

iii) $6\phi^-_{\mathbf{k}_1}\phi^-_{\mathbf{k}_2}\phi^+_{\mathbf{k}_3}\phi^+_{\mathbf{k}_4}$: This term is interesting. We will look at it more closely below. For now, note that the factor of 6 comes from the different choices of momentum labels.

iv) $4\phi^-_{\mathbf{k}_1}\phi^+_{\mathbf{k}_2}\phi^+_{\mathbf{k}_3}\phi^+_{\mathbf{k}_4}$: This term is cubic in $\phi^+_{\mathbf{k}}$ and, like any term with an odd number of insertions, vanishes when averaged over the Gaussian ensemble.

v) $\phi^+_{\mathbf{k}_1}\phi^+_{\mathbf{k}_2}\phi^+_{\mathbf{k}_3}\phi^+_{\mathbf{k}_4}$: This term doesn't include any $\phi^-_{\mathbf{k}}$, it simply gives constant to the free energy. It will not be important here.

We learn that we need to compute just a single term,

$$\langle F_I\rangle_+ = 6g_0 \int \prod_{i=1}^{4} \frac{d^dk_i}{(2\pi)^d}\, \phi^-_{\mathbf{k}_1}\phi^-_{\mathbf{k}_2} \times \left\langle \phi^+_{\mathbf{k}_3}\phi^+_{\mathbf{k}_4}\right\rangle_+ \times (2\pi)^d\delta^d(\sum_i \mathbf{k}_i) \tag{3.28}$$

But this is the same kind of correlation function that we computed in Section 2.2: it is given by

$$\left\langle \phi^+_{\mathbf{k}}\phi^+_{\mathbf{k}'}\right\rangle_+ = (2\pi)^d\delta^d(\mathbf{k}+\mathbf{k}')\, G_0(k) \quad \text{with} \quad G_0(k) = \frac{1}{k^2+\mu_0^2} \tag{3.29}$$

After playing around with the delta-functions, and relabelling momentum variables, we're left with our first correction to the free energy,

$$\langle F_I\rangle_+ = 6g_0 \int_0^{\Lambda/\zeta} \frac{d^dk}{(2\pi)^d}\, \phi^-_{\mathbf{k}}\phi^-_{-\mathbf{k}} \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d}\frac{1}{q^2+\mu_0^2}$$

where the limits on the $\int dq$ integral reflect the fact that we've only integrated out the short wavelength modes, whose momenta lie within this band. We see that, at order g_0, we get a correction only to the quadratic term whose coefficient becomes

$$\mu_0^2 \to \mu'^2 = \mu_0^2 + 12g_0 \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d}\frac{1}{q^2+\mu_0^2} \tag{3.30}$$

Finally, we should enact the rescaling steps of the renormalisation group. This takes the same form as before (3.20),

$$\mathbf{k}' = \zeta\mathbf{k} \quad \text{and} \quad \phi'_{\mathbf{k}} = \zeta^{-w}\phi^-_{\mathbf{k}/\zeta} \quad \text{with} \quad w = \frac{d+2}{2}$$

This gives the same scaling of the parameters that we saw in Section 3.3. We have

$$\mu^2(\zeta) = \zeta^2\left(\mu_0^2 + 12g_0 \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d}\frac{1}{q^2+\mu_0^2}\right) \quad \text{and} \quad g(\zeta) = \zeta^{4-d}g_0 \tag{3.31}$$

The upshot of this calculation is that turning on a ϕ^4 coupling will give rise to a quadratic ϕ^2 coupling under RG flow. This is typical of these kinds of calculations: couplings of one type will induce others.

The coefficient of the ϕ^2 term is particularly important for our story, since the critical point is defined to be the place where this vanishes. We see that it's not so easy to make this happen. You can't simply set $\mu_0^2 = 0$ at some high scale and expect to hit criticality. Indeed, the result (3.31) tells us that, at long wavelengths, the "natural" value is $\mu^2 \sim g_0 \Lambda^{d-2}$, which is typically large. If you want to hit the critical point, you must "fine tune" the original coefficient μ_0^2 to cancel the new terms that are generated by RG flow.

You might think that this calculation answers the question of what happens to the theory in $d = 4$ when we turn on $g\phi^4$. It certainly tells us that turning on this coupling will induce the relevant coupling $\mu^2\phi^2$ and so take us away from the Gaussian fixed point. However, a closer look tells us at (3.31) reveals that it's possible to turn on a combination of g_0 and μ_0^2, so that $\mu^2(\zeta)$ remains zero. This combination is a marginal coupling. We learn that, at this order, there remains one relevant and one marginal deformation.

3.4.2 Order g_0^2

The corrections to the ϕ^4 terms first arise at order g_0^2. Here we have the contribution

$$F'[\phi^-] \sim -\frac{1}{2}\left(\langle F_I^2\rangle - \langle F_I\rangle^2\right) \tag{3.32}$$

Expanding out $\langle F_I^2\rangle$, we find 256 different terms. We will see how to organise them shortly, but for now we make a few comments before focussing on the term of interest.

Some of the terms in $\langle F_I^2\rangle$ will result in corrections that cannot be written as a local free energy, but are instead of the form $\left(\int d^dx\, f(\phi^-)\right)^2$ for some $f(\phi)$. These terms will be cancelled by the $\langle F_I\rangle^2$ terms. This is a general phenomena which you can learn more about in the lectures on Quantum Field Theory. In terms of Feynman diagrams, which we will introduce below, these kind of terms correspond to disconnected diagrams.

The terms that we care about in $\langle F_I^2\rangle$ are those which can be written as local corrections to the free energy. Of immediate interest for us are a subset of terms in $F_I^2 \sim \int d^dx\, d^dy\ \phi^4(\mathbf{x})\,\phi^4(\mathbf{y})$, given by

$$\begin{aligned}\frac{1}{2}\langle F_I^2\rangle_+ \sim \frac{1}{2}\binom{4}{2}^2 g_0^2 \int_0^{\Lambda/\zeta} \prod_{i=1}^{4}\left[\frac{d^dk_i}{(2\pi)^d}\,\phi^-_{\mathbf{k}_i}\right] \int_{\Lambda/\zeta}^{\Lambda} \prod_{j=1}^{4}\left[\frac{d^dq_j}{(2\pi)^d}\right]\ \langle \phi^+_{\mathbf{q}_1}\phi^+_{\mathbf{q}_2}\phi^+_{\mathbf{q}_3}\phi^+_{\mathbf{q}_4}\rangle_+ \\ \times\ (2\pi)^{2d}\delta^d(\mathbf{k}_1+\mathbf{k}_2+\mathbf{q}_1+\mathbf{q}_2)\delta^d(\mathbf{k}_3+\mathbf{k}_4+\mathbf{q}_3+\mathbf{q}_4)\end{aligned} \tag{3.33}$$

Let's explain what's going on here. Each $\phi(\mathbf{x})$ is decomposed into Fourier modes ϕ^- and ϕ^+. The same is true for each $\phi(\mathbf{y})$. In the above term, we have chosen two ϕ^- out of the $\phi^4(\mathbf{x})$ and two ϕ^- out of the $\phi^4(\mathbf{y})$; the remaining terms are ϕ^+. Each combinatoric factor $\binom{4}{2} = 6$ out front reflects the choice of picking two ϕ^- from ϕ^4. Meanwhile, the two delta functions come from doing the $\int d^dx$ and $\int d^dy$ integrals respectively. Matching the momenta in the arguments of the delta functions to the $\phi^\pm$ tells us that we've picked two ϕ^- from the $\phi^4(\mathbf{x})$ and two ϕ^- from $\phi^4(\mathbf{y})$ (as opposed to, say, all four from $\phi^4(\mathbf{y})$).

To proceed, we need to compute the four-point function $\langle \phi^+_{\mathbf{q}_1}\phi^+_{\mathbf{q}_2}\phi^+_{\mathbf{q}_3}\phi^+_{\mathbf{q}_4}\rangle_+$. To do this we need a result known as *Wick's theorem.*

Wick's Theorem

As we proceed in our perturbative expansion, the integrals start to blossom. From the form of the expansion (3.27), we can see that the integrand will involve expectation values of the form $\langle \phi^+_{\mathbf{k}_1}\ldots\phi^+_{\mathbf{k}_l}\rangle_+$. There is a simple way to compute expectation values of this type in Gaussian ensembles. This follows from:

Lemma: Consider n variables ϕ_a drawn from a Gaussian ensemble. This means that, for any function $f(\phi)$, the expectation value is

$$\langle f(\phi)\rangle = \frac{1}{\mathcal{N}}\int_{-\infty}^{\infty} d^n\phi\; f(\phi)\, e^{-\frac{1}{2}\phi\cdot G^{-1}\phi}$$

for some invertible $n\times n$ matrix G. The normalisation factor is $\mathcal{N} = \det^{1/2}(2\pi G)$ and ensures that $\langle 1\rangle = 1$. The following identity then holds:

$$\left\langle e^{B_a\phi_a}\right\rangle = e^{\frac{1}{2}B_a\langle\phi_a\phi_b\rangle B_b} \tag{3.34}$$

for any constant B^a.

Proof: This is straightforward to show since we can just evaluate both sides

$$\begin{aligned}\left\langle e^{B_a\phi_a}\right\rangle &= \frac{1}{\mathcal{N}}\int_{-\infty}^{\infty} d^ny\; e^{-\frac{1}{2}\phi\cdot G^{-1}\phi + B\cdot\phi}\\ &= \frac{1}{\mathcal{N}}\int_{-\infty}^{\infty} d^ny\; e^{-\frac{1}{2}(\phi-GB)\cdot G^{-1}(\phi-GB)}e^{\frac{1}{2}B\cdot GB} = e^{\frac{1}{2}B\cdot GB} = e^{\frac{1}{2}B_a\langle\phi_a\phi_b\rangle B_b}\end{aligned}$$

where, in the last step, we used the fact that $\langle\phi_a\phi_b\rangle = G_{ab}$. $\square$

The Taylor expansion of the identity (3.34) gives us the expressions that we want. The left-hand-side is

$$\left\langle e^{B_a\phi_a}\right\rangle = 1 + B_a\langle\phi_a\rangle + \frac{1}{2}B_aB_b\langle\phi_a\phi_b\rangle + \frac{1}{3!}B_aB_bB_c\langle\phi_a\phi_b\phi_c\rangle + \frac{1}{4!}B_aB_bB_cB_d\langle\phi_a\phi_b\phi_c\phi_d\rangle + \ldots$$

Meanwhile, the right-hand-side is

$$e^{\frac{1}{2}B_a\langle\phi_a\phi_b\rangle B_b} = 1 + \frac{1}{2}B_aB_b\langle\phi_a\phi_b\rangle + \frac{1}{8}B_aB_bB_cB_d\langle\phi_a\phi_b\rangle\langle\phi_c\phi_d\rangle + \ldots$$

Now we just match powers of B_a on both sides. We immediately learn that

$$\langle\phi_{a_1}\ldots\phi_{a_l}\rangle = 0 \quad \text{for } l \text{ odd}$$

Our real interest is in l even. Here we have to be a little careful because multiplying by the string of B's automatically symmetrises the products of $\langle\phi_a\phi_b\rangle$ over the $a = 1, \ldots, n$ indices. So, for example, comparing the B^4 terms gives

$$\langle\phi_a\phi_b\phi_c\phi_d\rangle = \langle\phi_a\phi_b\rangle\langle\phi_c\phi_d\rangle + \langle\phi_a\phi_c\rangle\langle\phi_b\phi_d\rangle + \langle\phi_a\phi_d\rangle\langle\phi_b\phi_c\rangle \tag{3.35}$$

It's not hard to convince yourself that

$$\langle\phi_{a_1}\ldots\phi_{a_{2l}}\rangle = \langle\phi_{a_1}\phi_{a_2}\rangle\ldots\langle\phi_{a_{2l-1}}\phi_{a_{2l}}\rangle + \text{ symmetrised}$$

This leaves us with a sum over all pairwise contractions. This result is known as *Wick's theorem.*

Back to the Free Energy

We can now apply Wick's theorem to our free energy (3.33),

$$\langle\phi^+_{\mathbf{q}_1}\phi^+_{\mathbf{q}_2}\phi^+_{\mathbf{q}_3}\phi^+_{\mathbf{q}_4}\rangle_+ = \langle\phi^+_{\mathbf{q}_1}\phi^+_{\mathbf{q}_2}\rangle_+\langle\phi^+_{\mathbf{q}_3}\phi^+_{\mathbf{q}_4}\rangle_+ + \langle\phi^+_{\mathbf{q}_1}\phi^+_{\mathbf{q}_3}\rangle_+\langle\phi^+_{\mathbf{q}_2}\phi^+_{\mathbf{q}_4}\rangle_+ + \langle\phi^+_{\mathbf{q}_1}\phi^+_{\mathbf{q}_4}\rangle_+\langle\phi^+_{\mathbf{q}_2}\phi^+_{\mathbf{q}_3}\rangle_+$$

Recall that each of these propagators comes with a delta function,

$$\langle\phi^+_{\mathbf{q}}\phi^+_{\mathbf{q}'}\rangle_+ = (2\pi)^d\delta^d(\mathbf{q}+\mathbf{q}')\,G_0(q)$$

The trick is to see how these new delta functions combine with the original delta functions in (3.33). There are two different structures that emerge. The first term, gives (ignoring factors of 2π for now)

$$\begin{aligned}\int\prod_{j=1}^4 d^dq_j\;\; &\langle\phi^+_{\mathbf{q}_1}\phi^+_{\mathbf{q}_2}\rangle_+\langle\phi^+_{\mathbf{q}_3}\phi^+_{\mathbf{q}_4}\rangle_+\delta^d(\mathbf{k}_1+\mathbf{k}_2+\mathbf{q}_1+\mathbf{q}_2)\,\delta^d(\mathbf{k}_3+\mathbf{k}_4+\mathbf{q}_3+\mathbf{q}_4)\\ &\sim\int d^dq_2d^dq_4\;G_0(q_2)G_0(q_4)\,\delta^d(\mathbf{k}_1+\mathbf{k}_2)\,\delta^d(\mathbf{k}_3+\mathbf{k}_4)\end{aligned} \tag{3.36}$$

We're still left with two delta functions over the $\mathbf{k}$ variables. This means that when we go back to real space, this term does not become a local integral. Instead, if you follow it through, it becomes a double integral of the form $\left(\int d^dx\;\phi^-(\mathbf{x})^2\right)^2$. As we explained after (3.32), these terms are ultimately cancelled by corresponding terms in $\langle F_I\rangle^2$. They are not what interests us.

Instead, we care about the second and third terms in the Wick expansion of $\langle \phi^+_{\mathbf{q}_1}\phi^+_{\mathbf{q}_2}\phi^+_{\mathbf{q}_3}\phi^+_{\mathbf{q}_4}\rangle_+$. Each of them gives a contribution of the form

$$\int \prod_{j=1}^{4} d^dq_j \ \langle \phi^+_{\mathbf{q}_1}\phi^+_{\mathbf{q}_3}\rangle_+ \langle \phi^+_{\mathbf{q}_2}\phi^+_{\mathbf{q}_4}\rangle_+ \delta^d(\mathbf{k}_1+\mathbf{k}_2+\mathbf{q}_1+\mathbf{q}_2)\,\delta^d(\mathbf{k}_3+\mathbf{k}_4+\mathbf{q}_3+\mathbf{q}_4)$$
$$\sim \int d^dq_1 d^dq_2 \ G_0(q_1)\,G_0(q_2)\,\delta^d(\mathbf{k}_1+\mathbf{k}_2+\mathbf{q}_1+\mathbf{q}_2)\,\delta^d(\mathbf{k}_3+\mathbf{k}_4-\mathbf{q}_1-\mathbf{q}_2)$$
$$\sim \int d^dq \ G_0(q)\,G_0(|\mathbf{k}_1+\mathbf{k}_2+\mathbf{q}|)\,\delta^d(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3+\mathbf{k}_4) \tag{3.37}$$

where, in going to the last line, we have done the integral $\int d^dq_2$ and relabelled $q_1 = q$. Now we have just a single delta function over $\mathbf{k}_i$ and, correspondingly, when we go back to real space this will give a local contribution to the free energy. Indeed, the terms (3.33) now become

$$\frac{1}{2}\langle F_I^2\rangle_+ \sim \binom{4}{2}^2 g_0^2 \int_0^{\Lambda/\zeta} \prod_{i=1}^{4}\left[\frac{d^dk_i}{(2\pi)^d}\ \phi^-_{\mathbf{k}_i}\right] f(\mathbf{k}_1+\mathbf{k}_2)\ (2\pi)^d\delta^d(\sum_i \mathbf{k}_i) \tag{3.38}$$

where the factor of $\frac{1}{2}$ in (3.33) has disappeared because we get two contributions from the Wick expansion, each of which gives the same contribution (3.37). The remaining integral over $\int d^dq$ is hidden in the function $f(\mathbf{k})$, which is given by

$$f(\mathbf{k}_1+\mathbf{k}_2) = \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d}\ \frac{1}{q^2+\mu_0^2}\,\frac{1}{(\mathbf{k}_1+\mathbf{k}_2+\mathbf{q})^2+\mu_0^2} \tag{3.39}$$

This is not as complicated as it looks. We can write it as

$$f(\mathbf{k}_1+\mathbf{k}_2) = \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d}\frac{1}{(q^2+\mu_0^2)^2}\,(1+\mathcal{O}(\mathbf{k}_1,\mathbf{k}_2))$$

All the terms that depend on the external momenta $\mathbf{k}_1$ and $\mathbf{k}_2$ will generate terms in the free energy of the form $k^2(\phi^-)^4 \sim (\phi^-)^2\nabla^2(\phi^-)^2$. These are irrelevant terms that will not be interesting for us other than to note that once we let loose the dogs of RG, we will no longer sit comfortably within some finite dimensional subspace of the coupling constants. Integrating out degrees of freedom generates all possible terms consistent with symmetries; flowing to the IR allows us to focus on the handful of relevant ones.

The contribution (3.38) to the free energy is what we want. Translating back to real space, we learn that the quadratic term gets corrections at this order. We have

$$g_0 \to g_0' = g_0 - 36g_0^2 \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d}\frac{1}{(q^2+\mu_0^2)^2} \tag{3.40}$$

The minus sign in (3.40) is important. It can be traced to the minus sign in (3.32), and it determines the fate of the would-be marginal coupling $g\phi^4$ in $d = 4$ dimensions. Recall that, in $d = 4$, there is no contribution to the running of $g(\zeta)$ from the second and third steps of RG. Here, the leading contribution comes from the first step and, as we see above, this causes $g(\zeta)$ to get smaller as ζ increases. This means that the theory in $d = 4$ is similar in spirit to those in $d > 4$, with ϕ^4 an irrelevant coupling.

However, in $d = 4$, the RG flow for g happens much more slowly than other couplings. For this reason it is sometimes called *marginally irrelevant*, to highlight the fact that it only failed to be marginal when the perturbative corrections were taken into account. This is a general phenomenon: most couplings which naively appear marginal will end up becoming either marginally relevant or marginally irrelevant due to such corrections. In the vast majority of cases, the coupling turns out to be marginally irrelevant. However, there are a number of very important examples – the Kondo effect and non-Abelian gauge theories prominent among them – where a marginal coupling turns relevant. We'll see such an example in Section 4.3.

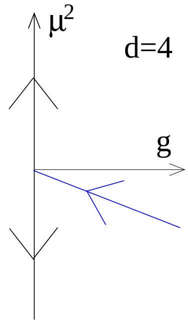


Figure 28:

Finally, just because g is marginally irrelevant in $d = 4$ does not mean that you can turn it on and expect to flow back to the Gaussian fixed point. As depicted in the diagram, the coupling mixes with μ^2. If you want to flow back to the Gaussian fixed point, you need to turn on a particular combination of μ_0^2 and g_0.

3.4.3 Feynman Diagrams

The calculation above needed some care. As we go to higher order terms in the expansion, the number of possibilities starts to blossom. Fortunately, there is a simple graphical formalism to keep track of what's going on.

Suppose that we're interested in a term in the expansion (3.27) of the form $g_0^p(\phi^-)^n(\phi^+)^l$. This can be represented by one or more *Feynman diagrams*. Here is the game:

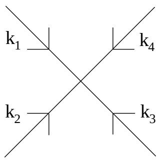
- Each $\phi^-_{\mathbf{k}}$ is represented by an external, solid line.
- Each $\phi^+_{\mathbf{k}}$ is represented by a dotted line.
- The dotted lines are connected with each other to form internal loops. They are paired up in all possible ways, reflecting the pair contraction of Wick's theorem. No dotted lines can be left hanging which means that the diagrams only make sense for l even.

- Each factor of g_0 is represented as a vertex at which four lines meet.
- Each line has an attached momentum $\mathbf{k}$ which is conserved as we move around the diagram.

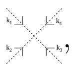
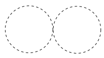
Each of these diagrams is really shorthand for an integral. The dictionary is as follows:

- Each internal line corresponds to an insertion of the propagator $\left\langle \phi^+_{\mathbf{k}} \phi^+_{\mathbf{k}'} \right\rangle_+$ defined in (3.29)
- For each internal loop, there is an integral $\int d^dq/(2\pi)^d$.
- Each vertex comes with a power of $g_0 (2\pi)^d \delta^d(\sum_i \mathbf{k}_i)$ where the delta function imposes momentum conservation, with the convention that all momenta are incoming.
- There are a bunch of numerical coefficients known as symmetry factors.

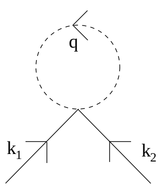
This means that a term in the effective action of the form $g_0^p (\phi^-)^n$ will correspond to a diagram with n external lines and p vertices. We can see what these diagrams look like for some of the terms that we've met so far. At order g_0, the rules don't allow us to draw diagrams with an odd number of $\phi^-_{\mathbf{k}}$. The term $\phi^-_{\mathbf{k}_1} \phi^-_{\mathbf{k}_2} \phi^-_{\mathbf{k}_3} \phi^-_{\mathbf{k}_4}$ in the expansion gives the trivial contribution to ϕ^4. In diagrams, it is



$$= g \int d^dx \; (\phi^-)^4 \tag{3.41}$$

Similarly, the $(\phi^+)^4$ terms are $\phi^+_{\mathbf{k}_1} \phi^+_{\mathbf{k}_2} \phi^+_{\mathbf{k}_3} \phi^+_{\mathbf{k}_4} \;\sim$ , but because these are internal lines we should join them up to get a diagram that looks like .

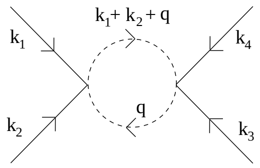
The interesting term at order g_0 is $\phi^-_{\mathbf{k}_1} \phi^-_{\mathbf{k}_2} \phi^+_{\mathbf{q}} \phi^+_{-\mathbf{q}}$, which includes both $\phi^-_{\mathbf{k}}$ and $\phi^+_{\mathbf{k}}$. This was evaluated in (3.28). It is represented by the diagram



$$= 6 g_0 \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d} \frac{1}{q^2 + \mu_0^2} \int d^dx \; (\phi^-)^2$$

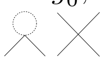
where the integral arises because the momentum $\mathbf{q}$ of the ϕ^+ excitation running in the loop is not determined by momentum conservation of the external legs. Such corrections are said to arise from *loop* diagrams, as opposed to the *tree* diagrams of (3.41).

Finally, at order g_0^2, the correction (3.38) to the ϕ^4 coupling comes from the



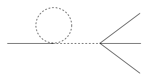
$$= 36 g_0^2 \int_0^{\Lambda/\zeta} \prod_{i=1}^{4} \left[\frac{d^d k_i}{(2\pi)^d} \, \phi^-_{\mathbf{k}_i} \right] f(\mathbf{k}_1 + \mathbf{k}_2) \; (2\pi)^d \delta^d (\sum_i \mathbf{k}_i)$$

where the function $f(\mathbf{k}_1 + \mathbf{k}_2)$ was given in (3.39) and contains two propagators associated to the ϕ^+ fields running in the loop. Evaluating this leads to the correction (3.40), as well as a slew of higher derivative couplings.

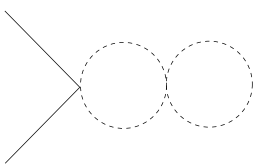
Any graph that we can draw will appear somewhere in the expansion of $\log \langle e^{-F_I[\phi^-_{\mathbf{k}}, \phi^+_{\mathbf{k}}]} \rangle_+$ given in (3.27). As we noted above, this is a cumulant expansion which has the rather nice graphical interpretation that only *connected* graphs appear in the expansion. For example, at order g_0^2, the disconnected graph that appears in the expansion of $\langle F_I^2 \rangle_+$ that looks like  will be cancelled by the same disconnected graph appearing in $-\langle F_I \rangle_+^2$.

In the language of quantum field theory, it's tempting to view the lines in the Feynman diagrams as the worldlines of particles. There is no such interpretation in the present case: they're simply useful.

More Diagrams

We can now also look at other diagrams and see what role they play. For example, you might be worried about the diagram . This is strictly vanishing, because the incoming momentum for lone ϕ^- leg is forced to be equal to the intermediate momentum for the ϕ^+ propagator. Yet the momentum for ϕ^- and ϕ^+ can never be equal .

There are, however, two further diagrams that we neglected which do have an interesting role to play. Both of these are two loop diagrams. They will not affect what we're going to do later but, nonetheless, are worth highlighting. The first diagram is

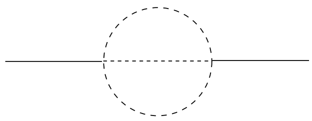


$$= g_0^2 \int \frac{d^d k}{(2\pi)^d} \, \frac{1}{2} C(\Lambda) \phi^-_{\mathbf{k}} \phi^-_{-\mathbf{k}}$$

for some $C(\Lambda)$ whose exact form will not be important. This gives rise to a shift in the quadratic term, so that (3.30) is replaced by

$$\mu_0^2 \to \mu'^2 = \mu_0^2 + 12g_0 \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d} \frac{1}{q^2+\mu_0^2} + g_0^2 C(\Lambda)$$

The second diagram is



$$= g_0^2 \int \frac{d^dk}{(2\pi)^d}\ \frac{1}{2} A(k,\Lambda)\phi^-_{\mathbf{k}}\phi^-_{-\mathbf{k}} \tag{3.42}$$

We've called the result of this diagram $A(k,\Lambda)$; again we won't need its detailed form. Importantly, it is now a function of the external momentum k. This means that it gives rise to two effects that are (in a technical sense) relevant. The first is that there is yet another renormalisation of μ^2, this time depending on $A(0,\Lambda)$. The second is novel: upon Taylor expanding $A(k) = A(0) + \frac{1}{2}k^2 A''(0) + \ldots$, we get a contribution to the the gradient term, which is now

$$F'[\phi] = \int d^dx\ \frac{1}{2}\gamma' \nabla\phi\cdot\nabla\phi + \ldots \quad \text{ with } \gamma' = 1 - 2g_0^2 A''(0,\Lambda)$$

This, in turn, means that we need a new rescaling of the field. To order g_0^2, we can write this as

$$\mathbf{k}' = \zeta\mathbf{k} \quad \text{and} \quad \phi'_{\mathbf{k}} = \frac{\zeta^{-(d+2)/2}}{1 - g_0^2 A''(0,\Lambda)}\,\phi^-_{\mathbf{k}/\zeta}$$

This last, additional step is known as *field renormalisation*. (Actually, that's not completely true. It should be known as "field renormalisation", but instead is known as "wavefunction renormalisation". This a terrible name, one that betrays the long and deep confusion that permeated the origins of this subject. Even in the context of quantum field theory, this rescaling has nothing to do with wavefunctions. It is a rescaling of fields!)

Although we won't compute this field renormalisation exactly, it is nonetheless important for this is what gives rise to the anomalous dimension of ϕ, and this was underlying the whole scaling analysis of Section 3.2.

3.4.4 Beta Functions

It is useful to write down equations which describe the flow of the coupling constants. These are first order differential equations which, for historic reasons are known as *beta functions*. It turns out to be convenient to parameterise the change in cut-off as

$$\Lambda' = \frac{\Lambda}{\zeta} = \Lambda e^{-s}$$

The renormalisation group transformation described above tells us that each coupling changes with scale, $g_n = g_n(s)$. The beta function is defined as

$$\frac{dg_n}{ds} = \beta_n(g_n)$$

Note that s increases as we flow towards the IR. This means that a positive beta function tells us that g_n gets stronger in the IR, while a negative beta function means that g_n gets weaker in the IR. (As an aside: this is the opposite to how beta functions are sometimes defined in quantum field theory, where one parameterises the flow in terms of energy rather than length.)

Before we jump straight in, it's useful to take a step backwards and build up the beta functions. Let's go back to our original scaling analysis around the Gaussian fixed point (3.24), where the running of the couplings is given by $g_n(s) = e^{(d-nd/2+n)s} g_{0,n}$. The beta functions are

$$\frac{dg_n}{ds} = \left(d - \frac{1}{2}nd + n\right) g_n \tag{3.43}$$

Notice that, at this leading order, there's no mixing between different couplings: turning on one coupling g_n does not induce another to flow. As we saw above, this state of affairs no longer holds when we include interactions.

We now focus on the two most important couplings, μ_0^2 and g_0. At order g_0, the RG equations are given by (3.31); the additional correction at order g_0^2, given in (3.40), means that these get replaced by

$$\mu^2(\zeta) = \zeta^2 \left(\mu_0^2 + a g_0\right) \quad \text{and} \quad g(\zeta) = \zeta^{4-d}(g_0 - b g_0^2) \tag{3.44}$$

where

$$a = 12 \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d} \frac{1}{q^2 + \mu_0^2} \quad \text{and} \quad b = 36 \int_{\Lambda/\zeta}^{\Lambda} \frac{d^dq}{(2\pi)^d} \frac{1}{(q^2 + \mu_0^2)^2}$$

Note that we have kept our original scaling dimensions in (3.44); the corrections in scaling due to the diagram (3.42) will be subleading and not needed in what follows.

When we differentiate $\mu^2(\zeta)$ and $g(\zeta)$ to derive the beta functions, we will get two terms: the first is (3.43) and comes from the scaling; the second comes from the corrections, given by the integrals a and b. Differentiating these integrals is particularly easy. For small s, we we write

$$\int_{\Lambda e^{-s}}^{\Lambda} dq\ f(q) \approx [\Lambda - \Lambda e^{-s}] f(\Lambda) \approx \Lambda f(\Lambda) s \quad \Rightarrow \quad \frac{d}{ds} \int_{\Lambda e^{-s}}^{\Lambda} dq\ f(q) = \Lambda f(\Lambda)$$

Let's restrict to $d = 4$ dimensions. The beta function equations are

$$\frac{d\mu^2}{ds} = 2\mu^2 + \frac{3g}{2\pi^2}\frac{\Lambda^4}{\Lambda^2+\mu^2} \quad \text{and} \quad \frac{dg}{ds} = -\frac{9}{2\pi^2}\frac{\Lambda^4}{(\Lambda^2+\mu^2)^2}g^2 \tag{3.45}$$

These don't (yet) contain any new physics, but it's worth reiterating what information we can extract from these equations.

First, the beta function for μ^2 has two terms; the first term comes from the second and third steps in RG (scaling), while the second comes from the first step in RG (integrating out). Meanwhile, the beta function for g has only a single term. There is no term linear in g because it was marginal under scaling, but it does receive a contribution when we integrate out the high momentum modes at order g^2. This contribution is negative, which tells us the coupling is marginally irrelevant. (A repeat of the warning above: this is the opposite convention to quantum field theory where one flows in decreasing energy, rather than increasing length, which means that a marginally irrelevant interaction is usually said to have a positive beta function.)

3.4.5 Again, the Analogy with Quantum Field Theory

The calculations above are very similar to the kind of loop integrals that you do in quantum field theory in $d = 3 + 1$ dimensions. There are, however, some philosophical differences between the approaches.

In statistical mechanics, the field $\phi(\mathbf{x})$ is, by construction, a coarse grained object: at the microscopic level, it dissolves into constituent parts, whether spins or atoms or something else. This has the practical advantage that we have no expectation that the statistical field theory will describe physics on arbitrarily short distance scales. In contrast, when we talk about quantum field theory in the context of high energy physics, it is tempting to think of the fields as "fundamental", a basic building block of our Universe. We may then wish for the theory to make sense down to arbitrarily small distance scales.

This ambition leads to a subtly different viewpoint on renormalisation. In quantum field theory one must introduce a cut-off, as we have above, to render integrals finite. However, this cut-off is very often viewed as an artefact, one which we would ultimately like to get rid of and make sense of the theory as $\Lambda \to \infty$. The trouble is that the renormalised quantities – things that we've called μ^2 and g – typically depend on this cut-off. We saw this, for example, in (3.44). Often this makes it tricky to take the limit $\Lambda \to \infty$.

To avoid this problem, one makes the so-called bare couplings – things we've called μ_0^2 and g_0 – depend on Λ. This is not such a dumb thing to do; after all, these quantities were defined at the cut-off scale Λ. The original game of renormalisation was to find a way to pick $\mu_0^2(\Lambda)$ and $g_0(\Lambda)$ such that all physical quantities remain finite as $\Lambda \to \infty$. It is by no means obvious that this is possible. Theories which can be rendered finite in this way are said to be *renormalisable*.

The high-energy approach to renormalisation predates the statistical physics approach and is now considered rather old-fashioned. The idea that a theory needs to make sense up to arbitrarily high energy scales smacks of hubris. The right way to view renormalisation – whether in statistical mechanics or in high energy physics – is through the renormalisation group procedure that has been our main focus in this chapter, in which one integrates out short wavelength modes to leave an effective long-distance theory.

Nonetheless, the high-energy approach to renormalisation has its advantages. Once one goes beyond the calculations described above, things are substantially easier with a high-energy viewpoint. You will learn more about these issues in the lectures on *Advanced Quantum Field Theory*.

3.5 The Epsilon Expansion

We have learned that ϕ^4 interaction is irrelevant for $d \geq 4$ and relevant for $d < 4$, sweeping us away from the Gaussian fixed point. But we seem to be no closer to figuring out where we end up. All we know is that we're not in Kansas anymore.

The difficulty is that we're limited in what we can calculate. We can't do the path integral exactly in the presence of ϕ^4 interactions, and are forced to work perturbatively in the coupling g. Yet, as we have seen, in dimension $d < 4$ the RG flow increases g, taking us to a regime where perturbation theory is no longer valid.

Nonetheless, the calculations that we did above do contain information about where we might expect to end up. But to see it, we have to do something rather dramatic. We will consider our theories not in $d = 4$ dimensions, but in

$$d = 4 - \epsilon$$

dimensions where ϵ is a small number, much less than 1. Clearly, this is an odd thing to do. You could view it as an act of wild creativity or, one of utter desperation. Probably it is a little bit of both. But, as we shall see, it will give us the insight we need to understand critical phenomena.

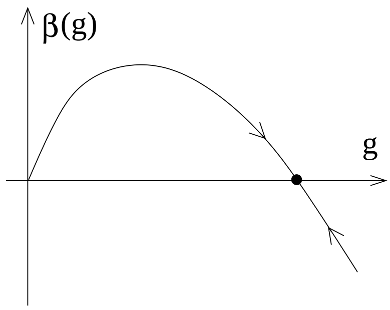


Figure 29: The beta function for g when $d < 4$. The arrows show the flow of g as we move towards the infra-red. The fixed point lies within the regime of perturbation theory when $d = 4 - \epsilon$.

First, we should ask whether it makes sense to work in a non-integer dimension. The lattice models that we started with surely need to be defined in dimension $d \in \mathbf{Z}^+$. Similarly, it was important for us that the free energy is local, meaning that it is written as an integral over space, and this too requires $d \in \mathbf{Z}^+$. However, by the time we get to the beta function equations, it makes mathematical, if not physical, sense to relax this and work in arbitrary $d \in \mathbf{R}$. We can read off these beta functions from the RG equations (3.44): they are

$$\begin{aligned}\frac{d\mu^2}{ds} &= 2\mu^2 + \frac{12\Omega_{d-1}}{(2\pi)^d}\frac{\Lambda^4}{\Lambda^2+\mu^2}\tilde{g} + \ldots \\ \frac{d\tilde{g}}{ds} &= \epsilon\tilde{g} - \frac{36\Omega_{d-1}}{(2\pi)^d}\frac{\Lambda^4}{(\Lambda^2+\mu^2)^2}\tilde{g}^2 + \ldots\end{aligned}$$

where Ω_{d-1} is the area of the sphere $\mathbf{S}^{d-1}$. We've introduced the dimensionless coupling $\tilde{g} = \Lambda^{-\epsilon} g$. Note that the beta function for $\tilde{g}$ now includes a term linear in $\tilde{g}$ arising from the scaling.

It can be checked that the two-loop diagrams we neglected contribute only at order ϵ^2. This means that it's consistent to truncate to the beta functions above. We could use the general formula for the area of a sphere, $\Omega_{d-1} = 2\pi^{d/2}/\Gamma(d/2)$, but this will give corrections of order ϵ^2, so instead we simply use $\Omega_3 = 2\pi^2$. Similar comments apply to $(2\pi)^d$. We're left with

$$\begin{aligned}\frac{d\mu^2}{ds} &\approx 2\mu^2 + \frac{3}{2\pi^2}\frac{\Lambda^4}{\Lambda^2+\mu^2}\tilde{g} \\ \frac{d\tilde{g}}{ds} &\approx \epsilon\tilde{g} - \frac{9}{2\pi^2}\frac{\Lambda^4}{(\Lambda^2+\mu^2)^2}\tilde{g}^2\end{aligned}$$

The novelty of these beta functions is that they have two fixed points. There is the Gaussian fixed point $\mu^2 = \tilde{g} = 0$ that we discussed before. And there is a new fixed point,

$$\mu_\star^2 = -\frac{3}{4\pi^2}\frac{\Lambda^4}{\Lambda^2+\mu_\star^2}\tilde{g}_\star \quad \text{and} \quad \tilde{g}_\star = \frac{2\pi^2}{9}\frac{(\Lambda^2+\mu_\star^2)^2}{\Lambda^4}\epsilon$$

Since we're working to leading order in ϵ, the solution is

$$\mu_\star^2 = -\frac{3}{4\pi^2}\Lambda^2\tilde{g}_\star = -\frac{1}{6}\Lambda^2\epsilon \quad \text{and} \quad \tilde{g}_\star = \frac{2\pi^2}{9}\epsilon$$

This is the *Wilson-Fisher fixed point*. Importantly, when ϵ is small then the fixed point $\tilde{g}_\star$ is also small, so our calculation is self-consistent (although, since we are in a fractional dimension, arguably unphysical!).

3.5.1 The Wilson-Fisher Fixed Point

To understand the flows in the vicinity of the new fixed point, we write $\mu^2 = \mu_\star^2 + \delta\mu^2$ and $\tilde{g} = \tilde{g}_\star + \delta\tilde{g}$. Linearising the beta functions, we find

$$\frac{d}{ds}\begin{pmatrix}\delta\mu^2\\ \delta\tilde{g}\end{pmatrix} = \begin{pmatrix} 2-\epsilon/3 & \frac{3}{2\pi^2}\Lambda^2\left(1+\frac{\epsilon}{6}\right)\\ 0 & -\epsilon\end{pmatrix}\begin{pmatrix}\delta\mu^2\\ \delta\tilde{g}\end{pmatrix}$$

where, as with all our other calculations, the entries in the matrix hold only up to $\mathcal{O}(\epsilon^2)$.

The eigenvalues of a triangular matrix coincide with the diagonal entries. We see that this fixed point has one positive and one negative eigenvalue,

$$\Delta_t = 2 - \frac{\epsilon}{3} + \mathcal{O}(\epsilon^2) \quad \text{and} \quad \Delta_g = -\epsilon + \mathcal{O}(\epsilon^2)$$

In other words, the Wilson-Fisher fixed point has one relevant direction and one irrelevant. The flows are shown in the figure.

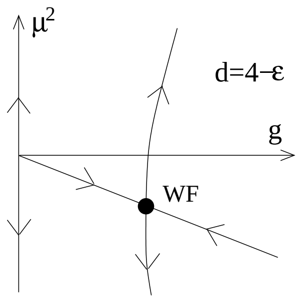


Figure 30:

We see that the epsilon expansion provides us with a global picture of the RG flows in $d < 4$ dimensions. One can check that all other couplings are also irrelevant at the Wilson-Fisher fixed point. Crucially, the fixed point sits at small g where our perturbative analysis is valid.

Now suppose that we increase ϵ. The Wilson-Fisher fixed point moves to higher g, and our perturbative approach breaks down. Nonetheless, it is not unreasonable to suppose that the qualitative picture of the flows remains the same. Indeed, this is thought to happen. Because the Wilson-Fisher fixed point has just a single relevant operator, it means that if we will generically end up there if we we're willing to tune just a single parameter, namely $T \to T_c$.

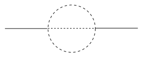
It is now a simple matter to compute the critical exponents in the epsilon expansion. Recall from Section 3.2 that these are related to the scaling dimensions of various terms. The relevant direction away from the Wilson-Fisher fixed point is temperature, $t = |T - T_c|/T_c$. Its dimension is determined by the way it scales as we approach the critical point, $t \to \zeta^{\Delta_t} t = e^{s\Delta_t} t$. But this is precisely the eigenvalue Δ_t that we just computed.

The critical exponent ν, defined by $\xi \sim t^{-\nu}$, is then given by (3.12)

$$\nu = \frac{1}{\Delta_t} = \frac{1}{2} + \frac{\epsilon}{12} + \mathcal{O}(\epsilon^2)$$

We can then use the hyperscaling relation $\alpha = 2 - d\nu$, given in (3.13), to compute the critical exponent for the heat capacity

$$c \sim t^{-\alpha} \quad \text{with} \quad \alpha = \frac{\epsilon}{6} + \mathcal{O}(\epsilon^2)$$

To compute the other critical exponents, we need to evaluate the anomalous dimension η. As we mentioned briefly above, this is related to the diagram  and turns out to be $\eta = \epsilon^2/6$, which is higher order in the expansion. We then have, from (3.10),

$$\Delta_\phi \approx \frac{d-2}{2} = 1 - \frac{\epsilon}{2}$$

Equations (3.14), (3.15) and (3.16) then give

$$\beta \approx \frac{1}{2} - \frac{\epsilon}{6} \quad , \quad \gamma \approx 1 + \frac{\epsilon}{6} \quad , \quad \delta \approx 3 + \epsilon$$

where all expressions hold up to corrections of order $\mathcal{O}(\epsilon^2)$.

Of course, our real interest lies in the system at $d = 3$, or $\epsilon = 1$. It would be in poor taste to simply plug in $\epsilon = 1$. But I know that you're curious. Here's what we get:

	α	β	γ	δ	η	ν
MF	0	$\frac{1}{2}$	1	3	0	$\frac{1}{2}$
$\epsilon = 1$	0.17	0.33	1.17	4	0	0.58
$d = 3$	0.1101	0.3264	1.2371	4.7898	0.0363	0.6300

Our answers are embarrassingly close to the correct values given the dishonest method we used to get there. One can, however, make this approach more respectable. The ϵ expansion has been carried out to order $\mathcal{O}(\epsilon^5)$. It is not a convergent series. Nonetheless, sophisticated resummation techniques can be used to make sense of it, and the resulting expressions give a fairly accurate account of the critical exponents.

The real power of the epsilon expansion, however, is more qualitative than quantitative; it usually – but not always – gives a reliable view of the structure of RG flows.

3.5.2 What Happens in $d = 2$?

We have not yet discussed much about what happens in $d = 2$ dimensions. Here the story is somewhat richer. The first hint of this can be seen in a simple analysis of the Gaussian fixed point, which shows that

$$\Delta_\phi = [\phi] = 0$$

This means that the Gaussian fixed point has an infinite number of relevant deformations since ϕ^n, for each n, is relevant.

It turns out that, in contrast to $d = 3$, there are actually an infinite number of fixed points in $d = 2$. Roughly speaking, the n^{th} fixed point can be reached from the Gaussian fixed point by turning on

$$F[\phi] = \int d^2x \; \frac{1}{2}(\nabla\phi)^2 + g_{2(n+1)}\phi^{2(n+1)}$$

Of course, as we've seen above, the RG flow is not quite so simple. When we turn on the coupling $g_{2(n+1)}\phi^{2(n+1)}$ we will generate all other terms, including ϕ^2 and ϕ^4 and so on. To reach the n^{th} fixed point, we should tune all such terms to zero as we flow towards the infra-red.

One can show that the n^{th} fixed point has n relevant operators: schematically, these can be thought of as $\phi^2, \phi^4, \ldots, \phi^{2n}$ although, as above, there will be mixing between these. By turning on the least relevant operator, one can flow from the n^{th} fixed point to the $(n-1)^{\text{th}}$ fixed point.

The results above are not derived using the $\epsilon = 4-d$ expansion which, unsurprisingly, is not much use in $d = 2$. Instead, they rely on something new which we will briefly describe in Section 3.6.

3.5.3 A History of Renormalisation

> "After about a month of work [at General Atomic Corp] I was ordered to write up my results, as a result of which I swore to myself that I would choose a subject for research where it would take at least five years before I had anything worth writing about. Elementary particle theory seemed to offer the best prospects of meeting this criterion."
>
> *Kenneth Wilson*

Renormalisation first entered physics in the context of quantum field theory, with the need to make sense of the UV divergences that arise in quantum electrodynamics. The theory, developed by Schwinger, Feynman, Tomonaga, Dyson and others, amounts to finding a consistent way to subtract one infinity from another, leaving behind a finite answer. This method yields excellent agreement with experiment but is, in the words of Feynman, a "dippy process", in which the infinities are not so much understood as swept under a very large rug.

The first hint of something deeper – and the first hint of a (semi)-group action – was seen in the work of Gell-Mann and Low in 1954. They realised that one could define an effective charge of the electron, $e(\mu)$ where μ denotes the energy scale at which the experiment takes place. This interpolates between the physical charge, as $\mu \to 0$, and the so called bare charge at high energies.

Meanwhile, throughout the 50s and 60s, a rather different community of physicists were struggling to understand second order phase transitions. It had long been known that Landau theory fails at critical points, but it was far from clear how to make progress, and the results that we've described in this course took several decades to uncover. In Kings College London, a group led by Cyril Domb stressed the importance of focussing on critical exponents; at Cornell University, Benjamin Widom showed that the relationships between critical exponents could be derived by invoking a scale invariance, albeit with little understanding of where such scale invariance came from; and at the University of Illinois, Leo Kadanoff introduced the idea of "blocking" in lattice models, a real-space renormalisation group in which one worked with successively coarser lattice models.

While many people contributed to these developments, the big picture, linking ideas from particle physics, statistical physics and condensed matter physics, is due mostly to. . .

Kenneth Wilson: 1936-2013

Ken Wilson received his PhD in 1961, working with Murray Gell-Mann on an assortment of topics in particle physics that fed his interest in the renormalisation group. He went on to spend much of the 1960s confused, scrabbling to understand the physics of scale, first in quantum field theory and later in the context of critical phenomena. He wrote very few papers in this time, but his reputation was strong enough to land him postdocs at Harvard and CERN and later even tenure at Cornell. (Some career advice for students: the strategy of being a genius and not writing anything rarely leads to such success.)

The floodgates opened in 1971 when Wilson set out his grand vision of the renormalisation group and, with his colleague Michael Fisher, suggested the epsilon expansion as a perturbative method to compute critical exponents in a paper charmingly titled "Critical Exponents in 3.99 Dimensions". Wilson used these methods to solve the "Kondo problem" in which an isolated spin, sitting in a bath of mobile electrons, exhibits asymptotic freedom, and he was among the first to understand the importance of numerical approaches to solve statistical and quantum field theories, pioneering the subject of lattice gauge theory. In 1982 he was awarded the Nobel prize for his contributions to critical phenomena.

3.6 Looking Forwards: Conformal Symmetry

There are many questions that we have not yet answered? How do we know the critical exponents in $d = 2$ exactly? How do we know that there are an infinite number of fixed points in $d = 2$? Why are the critical exponents in $d = 2$ rational numbers while, in $d = 3$ they have no known closed form? How are we able to compute the $d = 3$ critical exponents to 5 significant figures?

The answers to all these questions can be found in the emergence of a rich mathematical structure at the critical point. As we've seen throughout these lectures, the basic story of RG ensures that physics at the critical point is invariant under scale transformations

$$\mathbf{x} \to \lambda \mathbf{x} \tag{3.46}$$

More surprising is the fact that the physics is invariant under a larger class of symmetries, known as conformal transformations. These transformation consist of any map

$$\mathbf{x} \to \tilde{\mathbf{x}}(\mathbf{x}) \quad \text{such that} \quad \frac{\partial \tilde{x}^i}{\partial x^k}\frac{\partial \tilde{x}^j}{\partial x^l}\,\delta_{ij} = \phi(\mathbf{x})\delta_{kl} \tag{3.47}$$

for some function $\phi(\mathbf{x})$. Such conformal transformations have the property that they preserve angles between lines.

The equation (3.47) has obvious solutions, such as translations and rotations, for which $\phi(\mathbf{x}) = 1$. Furthermore, it is simple to see that the scaling (3.46) falls into the class of conformal transformations, with $\phi(\mathbf{x}) = \lambda^2$. However, it turns out that there is one further, less intuitive transformation that obeys this condition. This is known as the *special conformal transformation* and is given by

$$x'^i = \frac{x^i - (\mathbf{x}\cdot\mathbf{x})a^i}{1 - 2(\mathbf{x}\cdot\mathbf{a}) + (\mathbf{a}\cdot\mathbf{a})(\mathbf{x}\cdot\mathbf{x})} \tag{3.48}$$

parameterised by an arbitrary vector $\mathbf{a}$.

The first question that we should ask is: why are theories at the fixed points invariant under the larger group of conformal transformations, rather than just scale transformations? The answer to this, which goes somewhat beyond this course, involves a deeper understanding of the nature of the RG flows and hinges, crucially, on being able to construct a quantity which decreases monotonically along the flow. This quantity is, unhelpfully, called c in $d = 2$ dimensions, f in $d = 3$ dimensions and a in $d = 4$ dimensions, and the fact that it decreases monotonically is referred to as the c-theorem, f-theorem and a-theorem respectively. Using this machinery, it is then possible to prove that scale invariance implies conformal invariance. (The proof is more clearcut in $d = 2$; it relies on some extra assumptions in higher d, and there is a general feeling that there is more to understand here.)

The existence of an extra symmetry (3.48) brings a newfound power to the study of fixed points. The d translational symmetries, $\frac{1}{2}d(d-1)$ rotational symmetries, d special conformal symmetries and single scale transformation combine to form the conformal group, which can be shown to be isomorphic to $SO(d+1,1)$. All fields and, in particular, all correlation functions must fall into representations of this group, a fact which restricts their form. In recent years, our understanding of these representations has allowed new precision in the computation of critical exponents in $d = 3$ dimensions. This programme goes by the name of the *conformal bootstrap*.

Conformal Symmetry in $d = 2$

In $d = 2$ dimensions, conformal symmetry turns out to be particularly powerful. The group of finite conformal transformations follows the pattern in higher dimensions, and is $SO(3,1) \cong SL(2,\mathbf{C})$. However, something rather special happens if you look at infinitesimal transformations where one finds that many many more are allowed. In fact, there are an infinite number. This means that there is a powerful, infinite dimensional algebra, known as the *Virasoro algebra*, underlying conformal theories in $d = 2$ dimensions. This places much stronger constraints on these fixed points, ultimately rendering many of them solvable without resorting to perturbation theory. This is the reason why the critical exponents are rational numbers which can be computed exactly. This is also what allows us to understand the structure of the infinite number of multi-critical fixed points described in Section 3.5.2.

Conformal field theory in $d = 2$ dimensions is a vast subject which arises in many different areas of physics. Although originally developed to understand critical phenomena, it also plays an important role in the lectures on the Quantum Hall Effect and the lectures on String Theory, where you can find an introduction to the basics of the subject.