

Statistical Field Theory

University of Cambridge Part III Mathematical Tripos

David Tong

*Department of Applied Mathematics and Theoretical Physics,
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge, CB3 0BA, UK*

<http://www.damtp.cam.ac.uk/user/tong/sft.html>
d.tong@damtp.cam.ac.uk

Recommended Books and Resources

There are a large number of books which cover the material in these lectures, although often from very different perspectives. They have titles like “Critical Phenomena”, “Phase Transitions”, “Renormalisation Group” or, less helpfully, “Advanced Statistical Mechanics”. Here are some that I particularly like

- Nigel Goldenfeld, *Phase Transitions and the Renormalization Group*

A great book, covering the basic material that we’ll need and delving deeper in places.

- Mehran Kardar, *Statistical Physics of Fields*

The second of two volumes on statistical mechanics. It cuts a concise path through the subject, at the expense of being a little telegraphic in places. It is based on lecture notes which you can find on the web; a link is given on the course website.

- John Cardy, *Scaling and Renormalisation in Statistical Physics*

A beautiful little book from one of the masters of conformal field theory. It covers the material from a slightly different perspective than these lectures, with more focus on renormalisation in real space.

- Chaikin and Lubensky, *Principles of Condensed Matter Physics*
- Shankar, *Quantum Field Theory and Condensed Matter*

Both of these are more all-round condensed matter books, but with substantial sections on critical phenomena and the renormalisation group. Chaikin and Lubensky is more traditional, and packed full of content. Shankar covers modern methods of QFT, with an easygoing style suitable for bedtime reading.

A number of excellent lecture notes are available on the web. Links can be found on the course webpage: <http://www.damtp.cam.ac.uk/user/tong/sft.html>.

Contents

0. Introduction	2
1. From Spins to Fields	5
1.1 The Ising Model	5
1.1.1 The Effective Free Energy	7
1.1.2 Mean Field Theory	9
1.2 Landau Approach to Phase Transitions	11
1.2.1 $B = 0$: A Continuous Phase Transitions	11
1.2.2 $B \neq 0$: First Order Phase Transitions	14
1.2.3 Validity of Mean Field Theory	17
1.2.4 A First Look at Universality	20
1.3 Landau-Ginzburg Theory	22
1.3.1 The Landau-Ginzburg Free Energy	24
1.3.2 The Saddle Point and Domain Walls	26
1.3.3 The Lower Critical Dimension	29
1.3.4 Lev Landau: 1908-1968	31
2. My First Path Integral	32
2.1 The Thermodynamic Free Energy Revisited	34
2.1.1 The Heat Capacity	37
2.2 Correlation Functions	39
2.2.1 The Gaussian Path Integral	41
2.2.2 The Correlation Function is a Green's Function	44
2.2.3 The Correlation Length	46
2.2.4 The Upper Critical Dimension	49
2.3 The Analogy with Quantum Field Theory	50
3. The Renormalisation Group	53
3.1 What's the Big Idea?	53
3.1.1 Universality Explained	59
3.2 Scaling	62
3.2.1 Critical Exponents Revisited	64
3.2.2 The Relevance of Scaling	66
3.3 The Gaussian Fixed Point	67

3.3.1	In the Vicinity of the Fixed Point	69
3.3.2	Dangerously Irrelevant	73
3.3.3	An Aside: The Emergence of Rotational Symmetry	74
3.4	RG with Interactions	75
3.4.1	Order g_0	76
3.4.2	Order g_0^2	78
3.4.3	Feynman Diagrams	83
3.4.4	Beta Functions	86
3.4.5	Again, the Analogy with Quantum Field Theory	87
3.5	The Epsilon Expansion	88
3.5.1	The Wilson-Fisher Fixed Point	90
3.5.2	What Happens in $d = 2$?	92
3.5.3	A History of Renormalisation	93
3.6	Looking Forwards: Conformal Symmetry	95
4.	Continuous Symmetries	97
4.1	The Importance of Symmetry	97
4.2	$O(N)$ Models	100
4.2.1	Goldstone Bosons	101
4.2.2	The $d = 4 - \epsilon$ Expansion	106
4.2.3	There Are No Goldstone Bosons in $d = 2$	111
4.3	Sigma Models	113
4.3.1	The Background Field Method	115
4.3.2	Asymptotic Freedom and the $d = 2 + \epsilon$ Expansion	119
4.3.3	Large N	121
4.4	The Kosterlitz-Thouless Transition	125
4.4.1	Vortices	127
4.4.2	From Coulomb Gas to Sine-Gordon	131
4.4.3	RG Flows in Sine-Gordon	133

Acknowledgements

These lectures are aimed at beginning graduate students. They assume a background in statistical mechanics and thermodynamics. No quantum field theory is required, although the techniques that we develop here are useful in that context.

The course is built on the foundation of previous courses given in Cambridge by Ron Horgan and Matt Wingate. I'm grateful to the students for their questions and to Carl Turner for comments. I'm supported by the Royal Society and Alex Considine Tong.

Conventions

In a previous course on [Statistical Physics](#), we diligently kept factors of the Boltzmann constant, k_B , in every equation. Now it is time to grow up. We set $k_B = 1$. This means that we measure temperature in units of energy.

0. Introduction

Superficially, this course is about *phase transitions*. This is the name given to the abrupt, discontinuous changes that occur when matter is deformed in some way, whether through heating or squeezing or something else. The familiar example is the violent shaking of a pot on a stove as water approaches its boiling point, and bubbles of steam erupt from within.

Despite their familiarity, phase transitions are striking, and even a little disconcerting. Usually in physics, things happens gradually. This fact is sewn into the heart of classical physics where the positions and momenta of particles are described by smooth, differentiable functions. Indeed, historically, the idea that change happens only infinitesimally resulted in the discovery of calculus. Yet, somehow, what holds on the micro level fails at the macro. Phase transitions tell us that a large number of particles can behave collectively in a way that any individual particle cannot, with the macroscopic properties of a system changing discontinuously.

A closer look at what happens at phase transitions – in particular at so-called critical points – reveals something startling. Many different substances, regardless of their microscopic composition, exhibit identical behaviour at a phase transition. This is not just a qualitative statement, but a quantitative one. For example, as a liquid changes into a gas at the critical temperature T_c , the heat capacity diverges as

$$c \sim \frac{1}{|T - T_c|^{0.11008\dots}}$$

The exponent is not known precisely. It is thought not to be a rational number, but should instead be viewed as a universal mathematical constant, similar to π or e , but more subtle. Remarkably, the same exponent occurs for all gases. It also occurs in other systems, including a certain class of magnets. It's as if all knowledge of the microscopic physics has been washed away, leaving us with something pure, that carries only a vague memory of what lies underneath. This phenomenon is known as *universality*.

All of this makes phase transitions interesting. They involve violence, universal truths and competition between rival states. The story of phase transitions is, quite literally, the song of fire and ice.

And yet these are not the only reasons to study phase transitions. In our attempt to understand what happens as water boils, we will need to develop new tools and a new way of thinking about the world. This leads us to a paradigm which now underlies huge swathes of physics, far removed from its humble origin of a pot on a stove. This

paradigm revolves around two deep facts about the Universe we inhabit: Nature is organised by symmetry. And Nature is organised by scale.

Nature is Organised by Symmetry

When I was a kid, I was told that there are three phases of matter: solid, liquid and gas. (Actually, this isn't quite true. Knowing that I was interested in this kind of stuff, the teacher conspiratorially let on that there was a fourth phase of matter, "plasma". To this day, I have no idea why. My best guess is that this fitted better with some old view of the basic elements as earth, water, air and fire.)

It won't be any surprise to learn that the real world is much more interesting than the one we're introduced to as kids. There are not three phases of matter, nor four: there are many. A key insight, due to Landau, is that these different phases are characterised by symmetry.

In this scheme a solid differs from a liquid because its crystal structure breaks the translational and rotational symmetries of space. Moreover, solids with different crystal structures should be viewed as different phases of matter because they break these symmetries in different ways. Perhaps more surprisingly, liquids and gases break no such symmetries and so should be viewed as the same phase. When you include further symmetries, such as rotations of spins in a magnet or more subtle quantum counterparts, this classification opens up a wide range of possibilities that allows us to understand almost all the known forms of matter.

This characterisation has its advantages. First, we can be sure that any attempt to change a material from one symmetry class to another will necessarily involve a violent phase transition. Second, it turns out that understanding the symmetries of a system will immediately determine many of its properties, especially at low temperature.

Moreover, the classification of matter in terms of symmetry has a power that goes far beyond its initial regime of application. The vacuum of space is, in many ways, like a complicated material, with quantum effects playing the role of thermal fluctuations. The vacuum can sit in different phases and is thought to have undergone several phase transitions as the Universe cooled after the big bang, each of which can be understood in terms of symmetries. All the ideas that we will develop here carry directly to theories of particle physics, cosmology and beyond.

Nature is Organised by Scale

There is an order to the Universe we live in. Roughly speaking, little things affect big things. Not the other way round.

This is something you already know: particle physics underlies nuclear and atomic physics; atomic physics underlies condensed matter and chemistry; and so on up the chain. It's certainly true that it can be difficult to make the leap from one level to the next, and new creative ideas are needed at each step, but this doesn't change the fact that there is an ordering. Big things don't affect little things. This is the reason there are no astrology departments in universities.

But there is another aspect to this story, one which is not often stressed. Little things affect big things, but they rarely affect very big things. Instead, little things affect slightly bigger things. And these, in turn, affect slightly bigger things too. But as you go up the chain, you lose the information about what came long before.

This again is something that you know. A zoologist who is interested in the way that starlings flock has little reason to study the dynamics of the Higgs boson. It's also the reason that science is possible in the first place: neither Newton nor Einstein needed to understand how quantum gravity works on microscopic distance scales to write down theories that work extraordinarily well on larger scales.

In the 1970s a mathematical formalism was developed that makes these ideas concrete. This formalism is called the *renormalisation group* and provides a framework to describe physics at different scales. The renormalisation group gets little coverage in popular science articles, yet is arguably the single most important advance in theoretical physics in the past 50 years. While zoologists may have little need to talk to particle physicists, the right way to understand both the Higgs boson and the flocking of starlings is through the language of the renormalisation group.

These two ideas – symmetry and scale – now dominate the way we think about physics. Yet both have their origins in the simple question: what happens when you boil water? The purpose of this course is to find out.

1. From Spins to Fields

The ideas that we will introduce in these lectures have wide applicability across many areas of physics. However, we will spend much of these lectures studying just a single example. This example is known as the *Ising model* and it provides a particularly simple model of a magnet. Despite its simplicity, the Ising model already contains many of the concepts that will keep us occupied for the rest of the course. Having understood the Ising model in some detail, we see how these ideas play out in many other phases of matter in Section 4.

1.1 The Ising Model

The Ising model is easy to state, yet hard to solve. We have a lattice in d spatial dimensions, containing N sites. On each lattice site $i = 1, \dots, N$, there lives a discrete variable which can take one of two states: $s_i = +1$ or $s_i = -1$.

It is useful to refer to these states as *spins*, with $s_i = +1$ corresponding to spin up, and $s_i = -1$ to spin down. However, we won't be using any machinery from quantum mechanics: there are no non-commuting operators, nor quantum superpositions of states. In this sense, the Ising model, while discrete, is purely classical.

The collection of spins $\{s_i\}$ has energy

$$E = -B \sum_i s_i - J \sum_{\langle ij \rangle} s_i s_j \quad (1.1)$$

The first term arises due to an external magnetic field, B that we impose on the system. It has the effect that, for $B > 0$, the spins want to be up because that will lower their energy. (Properly this should be the magnetising field H , but we're using B to avoid confusion with the Hamiltonian. There is also a factor of the magnetic moment that we've absorbed into B .)

With the first term alone, the individual spins don't talk to each other and the model is easy to solve. It is the second term that makes life more interesting. This is an interaction between neighbouring spins. The notation $\langle ij \rangle$ means that we sum over all "nearest neighbour" pairs in the lattice. The number of such pairs depends both on the dimension d and the type of lattice.

If $J > 0$, neighbouring spins prefer to be aligned ($\uparrow\uparrow$ or $\downarrow\downarrow$). In the context of magnetism, such a system is called a *ferromagnet*. If $J < 0$, the spins want to anti-align ($\uparrow\downarrow$). This is an *anti-ferromagnet*. In the following, we choose $J > 0$ although, for our purposes, the differences are minor.

(For what it’s worth, the anti-ferromagnetic case, with $J < 0$, becomes more subtle on certain lattices where it’s not possible to arrange the spins so that they are opposite to all their neighbours. A 2d triangular lattice provides a simple example. The resulting physics is interesting and is known as (geometrical) *frustration*. We will not discuss it here.)

We are interested in the physics of the Ising model at a finite temperature T . We can already get some intuition for what will happen. The interaction energy encourages the spins to align in the same way. The magnetic field encourages the spins to align in a particular way. Meanwhile, the temperature encourages the spins to ignore both the interactions and magnetic field because, at finite temperature, energy is no longer at a premium. Instead, entropy becomes more important. Since there are many more random configurations than aligned configurations, the temperature will tend to mess up the nice ordered states that the interactions and magnetic field have so carefully prepared. Already we can see that, like any good story, we’re starting with a healthy dose of narrative tension between our main characters.

In the canonical ensemble, the probability of sitting in a configuration of spins $\{s_i\}$ is given by

$$p[s_i] = \frac{e^{-\beta E[s_i]}}{Z} \tag{1.2}$$

where $\beta = 1/T$ and the normalisation factor Z is called the *partition function*, and is given by

$$Z(T, J, B) = \sum_{\{s_i\}} e^{-\beta E[s_i]} \tag{1.3}$$

The beauty of statistical mechanics is that even though Z is first introduced as an innocuous normalisation factor, it actually contains everything we want to know about the system. If we’re able to perform the sum to compute Z , we can extract any information we want¹. For example, the interplay between energy and entropy S is captured by the thermodynamic free energy

$$F_{\text{thermo}}(T, B) = \langle E \rangle - TS = -T \log Z \tag{1.4}$$

As the notation shows, this is a function of thermodynamic variables, T and B . (We don’t usually add the subscript “thermo” to the free energy but for much of these lectures we’ll be working with a more refined version of the free energy which we’ll shortly introduce.)

¹The basic machinery of partition functions was described in the first course on [Statistical Physics](#).

Of particular interest to us will be the average spin of the configuration, which we refer to as the equilibrium *magnetization*.

$$m = \frac{1}{N} \langle \sum_i s_i \rangle \quad (1.5)$$

This takes values in the range $m \in [-1, +1]$. From our discussion above, we would expect that, for $B > 0$, $m \rightarrow +1$ at low temperatures where the spins are ordered, and $m \rightarrow 0$ at high temperatures where the spins are arranged randomly. We'd like to make this intuition more precise. Using the probability (1.2), it is straightforward to check that the magnetisation can be written as

$$m = \frac{1}{N} \sum_{\{s_i\}} \frac{e^{-\beta E[\{s_i\}]} }{Z} \sum_i s_i = \frac{1}{N\beta} \frac{\partial \log Z}{\partial B} \quad (1.6)$$

Taking further derivatives allows us to compute higher moments of the distribution, and so learn about fluctuations around the average. We will see this in action later in these lectures.

Our task is now clear: we should compute the sum (1.3). This is easier said than done. It turns out that the sum is straightforward in a $d = 1$ dimensional lattice, and you will do this on an example sheet. An exact solution also exists in $d = 2$ when $B = 0$, originally due to Onsager. It is not straightforward. In higher dimensions, no exact solutions are available, although various expansions and tricks have been invented to manipulate the sum (1.3) to extract some interesting information.

In these lectures, we will not attempt to directly sum the microscopic spins in the partition function. Instead, we will rephrase the question. We will massage the partition function into a somewhat different form, one that has much broader application.

1.1.1 The Effective Free Energy

The key idea was explained in the introduction: Nature is organised by scale. The energy of the Ising model (1.1) provides the microscopic description of our system in terms of individual spins. Computing the partition function exactly gives us a macroscopic description of the system in terms of thermodynamic variables like temperature T and magnetic field B . What we're looking for is something in between.

We're going to get to this "something in between" in two steps: we'll do something rather crude here, and then do a better job in Section 1.3. For our crude attempt, we rewrite the partition function (1.3) in the following manner:

$$Z = \sum_m \sum_{\{s_i\}|m} e^{-\beta E[\{s_i\}]} := \sum_m e^{-\beta F(m)} \quad (1.7)$$

where the notation $\{s_i\}|m$ means all configurations of spins such that $\frac{1}{N} \sum_i s_i = m$. In other words, we first sum over all configurations with fixed average magnetisation $m = \frac{1}{N} \sum s_i$, and subsequently sum over all possible m .

Note that we're using m here in a subtly different way to the original definition (1.5). In the sum (1.7), the magnetisation refers to the average magnetisation of a given configuration and can take any value. In contrast, in (1.5) we are talking about the equilibrium value of magnetisation, averaged over all configurations in the canonical ensemble. We will see shortly how to find this equilibrium value.

The average magnetisation lies in the range $-1 \leq m \leq 1$. Strictly speaking, this takes only discrete values, quantised in units of $2/N$. However, we are ultimately interested in the limit of large N , so we don't lose anything by writing this as an integral,

$$Z = \frac{N}{2} \int_{-1}^{+1} dm e^{-\beta F(m)} \quad (1.8)$$

where the factor of $N/2$ is the (inverse) width between the allowed m values. Such overall coefficients in the partition function are unimportant for the physics, and we will not be careful in keeping them below.

This way of writing things has allowed us to define something new: an effective free energy, $F(m)$, which depends on the magnetisation m of the system, in addition to both T and B . This goes beyond the usual, thermodynamic idea of free energy F_{thermo} given in (1.4) which is defined only in equilibrium, where the magnetisation m takes a specific value, determined by (1.6).

Note that (1.8) looks very much like a standard path integral, with $F(m)$ playing the role of the energy for m . But there is a difference because, unlike in the microscopic theory, $F(m)$ can depend on temperature. This means that the β dependence in the exponent can be more complicated than we're used to. We'll see this explicitly below.

The effective free energy $F(m)$ contains more information than the thermodynamic free energy F_{thermo} . In particular, $F(m)$ can tell us the correct, equilibrium value of the magnetisation m . To see this, we define the free energy per unit spin,

$$f(m) = \frac{F(m)}{N}$$

Our partition function becomes

$$Z = \int dm e^{-\beta N f(m)}$$

Here N is a very large number (think $N \sim 10^{23}$) while $\beta f(m) \sim 1$. Integrals of this kind are very well approximated by the value of m which minimises $f(m)$, an approximation known as the *saddle point* or *steepest descent*,

$$\left. \frac{\partial f}{\partial m} \right|_{m=m_{\min}} = 0$$

The minimum m_{\min} is the equilibrium value of the magnetisation that we previously computed in (1.6). Substituting this into the the partition function, we have

$$Z \approx e^{-\beta N f(m_{\min})} \Rightarrow F_{\text{thermo}} \approx F(m_{\min}) \quad (1.9)$$

In this way, we can reconstruct F_{thermo} from knowledge of $F(m)$.

1.1.2 Mean Field Theory

To proceed, we want to find a way to compute $F(m) = N f(m)$, defined in (1.7). But this is tantamount to performing the sum in the path integral and, as we stressed above, this is not easy. We need to find another way.

We will use a method called the *mean field approximation*. Here the word “approximation” is somewhat generous; a better name would be the mean field “guess” since there is little justification for what we’re about to do. Instead, the purpose is to provide a starting point for our discussion. Not all the consequences that we derive from this guess will be accurate, but as the lectures progress we’ll get a better understanding about what we can trust, what we can’t, and how to do better.

We wish to sum over configurations $\{s_i\}$ with $\sum_i s_i = Nm$. We can get an estimate for the energy of such configurations by replacing each spin s in (1.1) with its expectation (i.e. *mean*) value $\langle s \rangle = m$,

$$E = -B \sum_i m - J \sum_{\langle ij \rangle} m^2 \Rightarrow \frac{E}{N} = -Bm - \frac{1}{2} J q m^2 \quad (1.10)$$

Here q denotes the number of nearest neighbours of each spin. For example, in $d = 1$ a lattice has $q = 2$; in $d = 2$, a square lattice has $q = 4$. A square lattice in d dimensions has $q = 2d$. The factor of $1/2$ is there because $\sum_{\langle ij \rangle}$ is a sum over pairs rather than a sum of individual sites.

Among other things, this means that we’re assuming (falsely!) that the energy depends only on the magnetisation m of a configuration, rather than any more finely grained details. The sole reason for doing this is that it makes the resulting sum over

configurations $\{s_i\}|m$ very easy to do: we simply need to count the number of configurations with magnetisation m . A configuration with N_\uparrow up spins and $N_\downarrow = N - N_\uparrow$ down spins has magnetisation

$$m = \frac{N_\uparrow - N_\downarrow}{N} = \frac{2N_\uparrow - N}{N}$$

The number of such configurations is

$$\Omega = \frac{N!}{N_\uparrow!(N - N_\uparrow)!}$$

and we can use Stirlings formula to evaluate

$$\begin{aligned} \log \Omega &\approx N \log N - N_\uparrow \log N_\uparrow - (N - N_\uparrow) \log(N - N_\uparrow) \\ \Rightarrow \frac{\log \Omega}{N} &\approx \log 2 - \frac{1}{2}(1 + m) \log(1 + m) - \frac{1}{2}(1 - m) \log(1 - m) \end{aligned}$$

In our mean field approximation, the effective free energy defined in (1.7) is then given by

$$e^{-\beta N f(m)} \approx \Omega(m) e^{-\beta E(m)}$$

Substituting the energy (1.10), and taking logs of both sides, we find ourselves with the following expression:

$$f(m) \approx -Bm - \frac{1}{2}Jqm^2 - T \left(\log 2 - \frac{1}{2}(1 + m) \log(1 + m) - \frac{1}{2}(1 - m) \log(1 - m) \right)$$

From this, we can compute the equilibrium value for m . As explained above, this occurs at the minimum

$$\frac{\partial f}{\partial m} = 0 \quad \Rightarrow \quad \beta(B + Jqm) = \frac{1}{2} \log \left(\frac{1 + m}{1 - m} \right)$$

A little algebra shows us that the equilibrium magnetisation obeys the self-consistency condition

$$m = \tanh(\beta B + \beta Jqm) \tag{1.11}$$

There's a nice intuition behind this equation. It can be derived by assuming that each spin experiences an effective magnetic field given by $B_{\text{eff}} = B + Jqm$, which includes an extra contribution from the spins around it. In this way, the tricky interactions in the Ising model have been replaced by an averaged effective field B_{eff} . This is sometimes called the *mean field* and gives its name to this technique.

There are a number of ways forward at this point. We could, for example, analyse the properties of the Ising model by looking at solutions to the self-consistency condition (1.11); this was the strategy taken in the [Statistical Physics](#) lectures. However, instead we're going to focus on the free energy $f(m)$, since this will prove to be the more useful tool moving forward.

1.2 Landau Approach to Phase Transitions

A phase transition occurs when some quantity changes in a discontinuous fashion. For us, this quantity is m which, as we will see, will exhibit non-analytic behaviour as we vary different parameters.

Landau theory is a simple, yet effective way to understand the qualitative manner in which these phase transitions occur. It is based on two ideas: the free energy, and symmetry. Here we will apply Landau's theory to the Ising model, focussing first on the free energy. We will comment on the role of symmetry at the end of this section, although the full power of this idea will only become apparent in Section 4 where we describe a whole raft of further examples.

In the previous section, we introduced the effective free energy $F(m) = Nf(m)$, which is a function of the magnetisation m . This kind of object is the starting point for Landau's theory. In this context, the magnetisation is known as an *order parameter*. Using the mean field approximation, we computed this to be

$$f(m) \approx -Bm - \frac{1}{2}Jqm^2 - T \left(\log 2 - \frac{1}{2}(1+m) \log(1+m) - \frac{1}{2}(1-m) \log(1-m) \right)$$

Landau's approach focuses on situations where the order parameter is small. Here we can Taylor expand,

$$f(m) \approx -T \log 2 - Bm + \frac{1}{2}(T - Jq)m^2 + \frac{1}{12}Tm^4 + \dots \quad (1.12)$$

As we've seen above, the preferred value of m is simply the minimum of $f(m)$. The idea of Landau theory is simply to observe how the function $f(m)$, and in particular its minima, change as we vary different parameters. Crucially, the story is rather different depending on whether $B \neq 0$ or $B = 0$. We will deal with each of these in turn.

1.2.1 $B = 0$: A Continuous Phase Transitions

Let's first consider the situation with vanishing magnetic field, $B = 0$, so that there is no term in (1.12) that is linear in m . Since our interest lies in the m dependence of

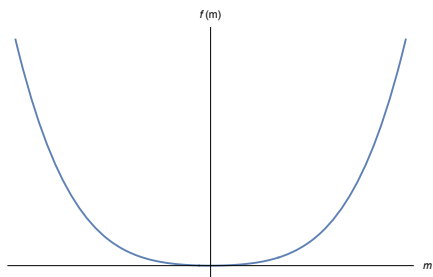


Figure 1: Free energy when $T > T_c$

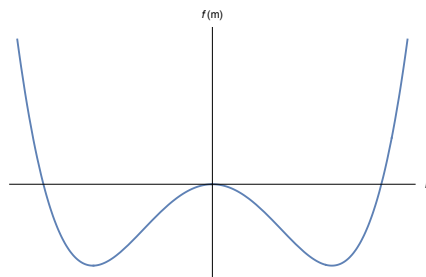


Figure 2: Free energy when $T < T_c$

$f(m)$, we won't lose anything if we drop the constant $-T \log 2$ term. We're left with the free energy

$$f(m) \approx \frac{1}{2}(T - Jq)m^2 + \frac{1}{12}Tm^4 + \dots \quad (1.13)$$

The behaviour of the free energy depends on whether the quadratic term is positive or negative. To distinguish between these two, we define the *critical temperature*

$$T_c = Jq$$

The free energy is sketched in the case where $T > T_c$ (on the left) and $T < T_c$ (on the right). We see that at high temperatures, the magnetisation vanishes at the minimum: $m = 0$. This is in agreement with our expectation that temperature will randomise the spins. However, as the temperature is reduced below T_c , the point $m = 0$ becomes a maximum of the free energy and the minima now lie at $m = \pm m_0$ which, if we chose to truncate the free energy (1.13) at order m^4 , is given by

$$m_0 = \sqrt{\frac{3(T_c - T)}{T}} \quad (1.14)$$

This form is valid only when T is close to T_c , so that m is small and higher order terms can be ignored. As the temperature is lowered further, the minimum m_0 grows. We're then obliged to turn to the full form of the free energy $f(m)$ which, among other things, knows that the magnetisation lies in the range $m \in [-1, +1]$.

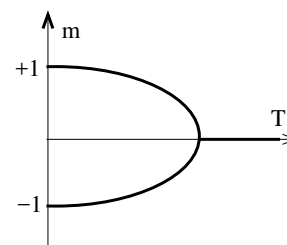


Figure 3:

The upshot is that as we vary the temperature, the magnetisation takes the form shown on the right. This is perhaps somewhat surprising. Usually in physics, things turn on gradually. But here the magnetisation turns off

abruptly at $T = T_c$, and remains zero for all higher temperatures. This kind of sharp change is characteristic of a phase transition. When $m = 0$, we say that the system sits in the *disordered phase*; when $m \neq 0$, it is in the *ordered phase*.

The magnetisation itself is continuous and, for this reason, it is referred to as a *continuous phase transition* or, sometimes, a *second order phase transition*.

As an aside: phase transitions can be classified by looking at the thermodynamic free energy $F_{\text{thermo}} = Nf(m_{\text{min}})$ and taking derivatives with respect to some thermodynamic variable, like T . If the discontinuity first shows up in the n^{th} derivative, it is said to be an n^{th} order phase transition. However, in practice we very rarely have to deal with anything other than first order transitions (which we will see below) and second order transitions. In the present case, we'll see shortly that the heat capacity is discontinuous, confirming that it is indeed a second order transition.

The essence of a phase transition is that some quantity is discontinuous. Yet, this should make us nervous. In principle, everything is determined by the partition function Z , defined in (1.3), which is a sum of smooth, analytic functions. How is it possible, then, to get the kind of non-analytic behaviour characteristic of a phase transition? The loophole is that Z is only necessarily analytic if the sum is finite. But there is no such guarantee that it remains analytic when the number of lattice sites $N \rightarrow \infty$. This means that genuine, discontinuous phase transitions only occur in infinite systems. In reality, we have around $N \approx 10^{23}$ atoms. This gives rise to functions which are, strictly speaking, smooth, but which change so rapidly that they act, to all intents and purposes, as if they were discontinuous.

These kind of subtleties are brushed under the carpet in the mean field approach that we're taking here. However, it's worth pointing out that the free energy, $f(m)$ is an analytic function which, when Taylor expanded, gives terms with integer powers of m . Nonetheless, the minima of f behave in a non-analytic fashion.

For future purposes, it will be useful to see how the heat capacity $C = \partial\langle E\rangle/\partial T$ changes as we approach T_c . In the canonical ensemble, the average energy is given by $\langle E\rangle = -\partial\log Z/\partial\beta$. From this, we find that we can write the heat capacity as

$$C = \beta^2 \frac{\partial^2}{\partial\beta^2} \log Z \tag{1.15}$$

To proceed, we need to compute the partition function Z , by evaluating $f(m)$ at the minimum value m_{min} as in (1.9). When $T > T_c$, this is simple: we have $m_{\text{min}} = 0$ and $f(m_{\text{min}}) = 0$ (still neglecting the constant $T \log 2$ term which doesn't contribute to the

heat capacity). In contrast, when $T < T_c$ the minimum lies at $m = m_0$ given in (1.14), and the free energy is $f(m_0) = -\frac{3}{4}(T_c - T)^2/T$.

Now we simply need to differentiate to get the heat capacity, C . The leading contribution as $T \rightarrow T_c$ comes from differentiating the $(T_c - T)$ piece, rather than the $1/T$ piece. We have

$$c = \frac{C}{N} \rightarrow \begin{cases} 0 & \text{as } T \rightarrow T_c^+ \\ 3/2 & \text{as } T \rightarrow T_c^- \end{cases} \quad (1.16)$$

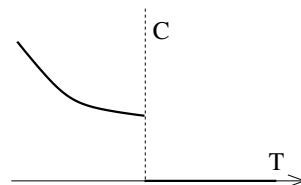


Figure 4:

We learn that the heat capacity jumps discontinuously. The part of the heat capacity that comes from differentiating $(T - T_c)$ terms is often called the *singular piece*. We'll be seeing more of this down the line.

Spontaneous Symmetry Breaking

There is one further aspect of the continuous phase transition that is worth highlighting. The free energy (1.13) is invariant under the \mathbf{Z}_2 symmetry $m \rightarrow -m$. This is no coincidence: it follows because our microscopic definition of the Ising model (1.1) also enjoys this symmetry when $B = 0$.

However, below T_c , the system must pick one of the two ground states $m = +m_0$ or $m = -m_0$. Whichever choice it makes breaks the \mathbf{Z}_2 symmetry. When a symmetry of a system is not respected by the ground state we say that the symmetry is *spontaneously broken*. This will become an important theme for us as we move through the course. It is also an idea which plays an important role in particle physics .

Strictly speaking, spontaneous symmetry breaking can only occur in infinite systems, with the magnetisation defined by taking the limit

$$m = \lim_{B \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i \langle s_i \rangle$$

It's important that we take the limit $N \rightarrow \infty$, before we take the limit $B \rightarrow 0$. If we do it the other way round, we find $\sum \langle s_i \rangle \rightarrow 0$ as $B \rightarrow 0$ for any finite N .

1.2.2 $B \neq 0$: First Order Phase Transitions

Let's now study what happens when $B \neq 0$. Once again, we ignore the constant $-T \log 2$ term in the free energy (1.12). We're left with the free energy

$$f(m) \approx -Bm + \frac{1}{2}(T - Jq)m^2 + \frac{1}{12}Tm^4 + \dots \quad (1.17)$$

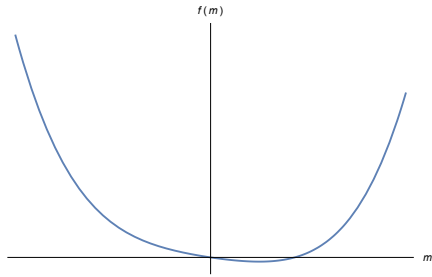


Figure 5: Free energy at high temperatures

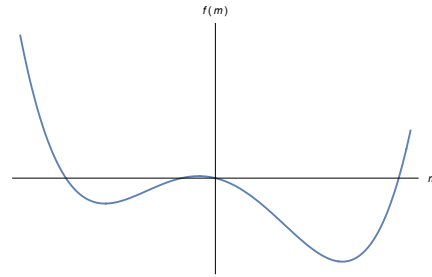


Figure 6: Free energy at low temperatures.

There are again two, qualitatively different forms of this function at low and high temperatures, shown in the figures above for $B > 0$.

At low temperatures there are two minima, but one is always lower than the other. The global minima is the true ground state of the system. The other minima is a *meta-stable state*. The system can exit the meta-stable state by fluctuating up, and over the energy barrier separating it from the ground state, and so has a finite lifetime. As we increase the temperature, there is a temperature (which depends on B) beyond which the meta-stable state disappears. This temperature is referred to as the *spinodal point*. It will not play any further role in these lectures.

For us, the most important issue is that the ground state of the system – the global minimum of $f(m)$ – does not qualitatively change as we vary the temperature. At high temperatures, the magnetisation asymptotes smoothly to zero as

$$m \rightarrow \frac{B}{T} \quad \text{as } T \rightarrow \infty$$

At low temperatures, the magnetization again asymptotes to the state $m \rightarrow \pm 1$ which minimises the energy. Except this time, there is no ambiguity as to whether the system chooses $m = +1$ or $m = -1$. This is entirely determined by the sign of the magnetic field B . A sketch of the magnetisation as a function of temperature is shown on the right. The upshot is that, for $B \neq 0$, there is no phase transition as a function of the temperature.

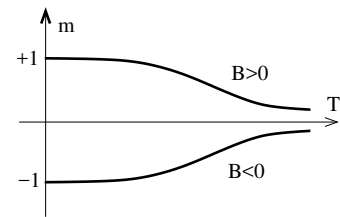


Figure 7:

However, we do not have to look too hard to find a phase transition: we just need to move along a different path in the phase diagram. Suppose that we keep T fixed at a

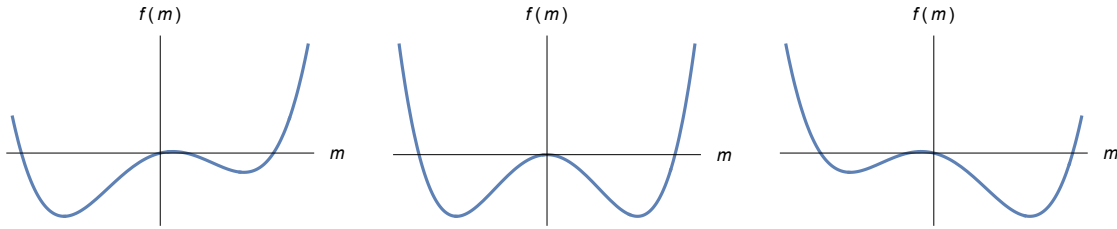


Figure 8: $T < T_c$: The free energy of the Ising model for $B < 0$ on the left, $B = 0$ in the middle, and $B > 0$ on the right.

value below T_c . We then vary the magnetic field from $B < 0$ to $B > 0$. The resulting free energy is shown in Figure 8.

We see that the magnetisation jumps discontinuously from $-m_0$ to $+m_0$ as B flips from negative to positive. This is an example of a *first order phase transition*.

Our analysis above has left us with the following picture of the phase diagram for the Ising model: if we vary B from positive to negative then we cross the red line in the figure and the system suffers a first order phase transition. Note, however, that if we first raise the temperature then it's always possible to move from any point $B \neq 0$ to any other point without suffering a phase transition.

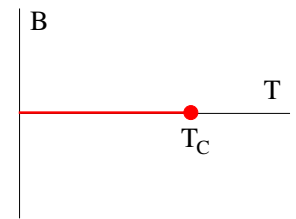


Figure 9:

This line of first order phase transitions ends at a second order phase transition at $T = T_c$. This is referred to as the *critical point*.

Close to the Critical Point

It will prove interesting to explore what happens when we sit close to the critical temperature $T = T_c$. There are a bunch of different questions that we can ask. Suppose, for example, that we sit at $T = T_c$ and vary the magnetic field: how does the magnetisation change? Here the free energy (1.17) becomes

$$f(m) \approx -Bm + \frac{1}{12}T_c m^4 + \dots$$

For m small, where we can neglect the higher order terms, minimising the free energy gives $m^3 \sim B$. So, when $B > 0$ we have

$$m \sim B^{1/3} \tag{1.18}$$

and when $B < 0$ we have $m \sim -|B|^{1/3}$.

Here is another question: the *magnetic susceptibility* χ is defined as

$$\chi = \left. \frac{\partial m}{\partial B} \right|_T \quad (1.19)$$

We will compute this at $B = 0$, and as $T \rightarrow T_c$ from both above and below. First, from above: when $T \gtrsim T_c$ we can keep just the linear and quadratic terms in the free energy

$$f(m) \approx -Bm + \frac{1}{2}(T - T_c)m^2 \quad \Rightarrow \quad m \approx \frac{B}{T - T_c} \quad \Rightarrow \quad \chi \sim \frac{1}{T - T_c}$$

When $T < T_c$, we need to work a little harder. For small B , we can write the minimum of (1.17) as $m = m_0 + \delta m$ where m_0 is given by (1.14). Working to leading order in B/T , we find

$$m \approx m_0 + \frac{B}{2(T_c - T)} \quad \Rightarrow \quad \chi \sim \frac{1}{T_c - T}$$

We can combine these two results by writing

$$\chi \sim \frac{1}{|T - T_c|} \quad (1.20)$$

We'll see the relevance of this shortly.

1.2.3 Validity of Mean Field Theory

The first thing that we should ask ourselves is: are the results above right?! We have reason to be nervous because they were all derived using the mean field approximation, for which we offered no justification. On the other hand, there is reason to be optimistic because, at the end of the day, the structure of the phase diagram followed from some fairly basic properties of the Taylor expansion of the free energy.

In this section and the next, we will give some spoilers. What follows is a list of facts. In large part, the rest of these lectures will be devoted to explaining where these facts come from. It turns out that the validity of mean field theory depends strongly on the spatial dimension d of the theory. We will explain this in detail shortly but here is the take-home message:

- In $d = 1$ mean field theory fails completely. There are no phase transitions.
- In $d = 2$ and $d = 3$ the basic structure of the phase diagram is correct, but detailed predictions at $T \approx T_c$ are wrong.

- In $d \geq 4$, mean field theory gives the right answers.

This basic pattern holds for all other models that we will look at too. Mean field theory always fails completely for $d \leq d_l$, where d_l known as the *lower critical dimension*. For the Ising model, $d_l = 1$, but we will later meet examples where this approach fails in $d_l = 2$.

In contrast, mean field theory always works for $d \geq d_c$, where d_c is known as the *upper critical dimension*. For the Ising model, mean field theory works because, as d increases, each spin has a larger number of neighbours and so indeed experiences something close to the average spin.

Critical Exponents

What about the intermediate dimensions, $d_l < d < d_c$? These are very often the dimensions of interest: for the Ising model it is $d = 2$ and $d = 3$. Here the crude structure of the phase diagram predicted by mean field theory is correct, but it gives misleading results near the critical point $T = T_c$.

To explain this, recall that we computed the behaviour of four different quantities as we approach the critical point. For three of these, we fixed $B = 0$ and dialled the temperature towards the critical point. We found that, for $T < T_c$, the magnetisation (1.14) varies as

$$m \sim (T_c - T)^\beta \quad \text{with } \beta = \frac{1}{2} \quad (1.21)$$

The heat capacity (1.16) varies as

$$c \sim c_\pm |T - T_c|^{-\alpha} \quad \text{with } \alpha = 0 \quad (1.22)$$

where the c_\pm is there to remind us that there is a discontinuity as we approach T_c from above or below. The magnetic susceptibility (1.20) varies as

$$\chi \sim \frac{1}{|T - T_c|^\gamma} \quad \text{with } \gamma = 1$$

The fourth quantity requires us to take a different path in the phase diagram. This time we fix T and dial the magnetic field B towards zero, in which case the magnetisation (1.18) varies as

$$m \sim B^{1/\delta} \quad \text{with } \delta = 3 \quad (1.23)$$

The coefficients α , β , γ and δ are known as *critical exponents*. (The Greek letters are standard notation; more generally, one can define a whole slew of these kind of objects.)

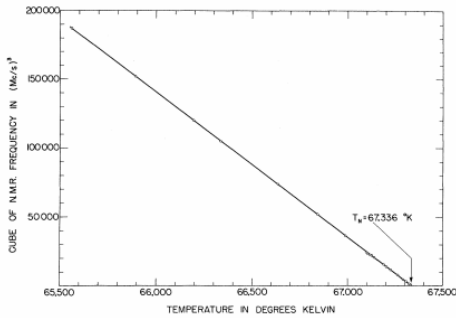


Figure 10: Magnetisation m^3 , plotted against temperature for an Ising magnet, suggests that $1/\beta \approx 3$.

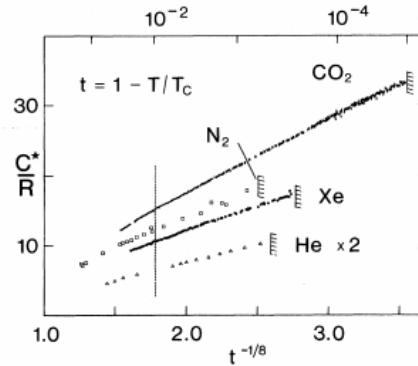


Figure 11: Heat capacity $c \sim |T - T_c|^{-\alpha}$ for various gases shows good agreement with $\alpha \approx 1/8$.

When the Ising model is treated correctly, one finds that these quantities do indeed scale to zero or infinity near T_c , but with different exponents. Here is a table of our mean field (MF) values, together with the true results for $d = 2$ and $d = 3$,

	MF	$d = 2$	$d = 3$
α	0 (disc.)	0 (log)	0.1101
β	$\frac{1}{2}$	$\frac{1}{8}$	0.3264
γ	1	$\frac{7}{4}$	1.2371
δ	3	15	4.7898

Note that the heat capacity has critical exponent $\alpha = 0$ in both mean field and in $d = 2$, but the discontinuity seen in mean field is replaced by a log divergence in $d = 2$.

The $d = 2$ results are known analytically, while the $d = 3$ results are known only numerically (to about 5 or 6 significant figures; I truncated early in the table above). Both the $d = 2$ and $d = 3$ results are also in fairly good agreement with experiment, which confirm that the observed exponents do not take their mean field values. For example, the left-hand figure above shows the magnetisation $m^{1/\beta} \sim (T - T_c)$ taken from MnF_2 , a magnet with uniaxial anisotropy which is thought to be described by the Ising model². The data shows a good fit to $1/\beta = 3$; as shown in the table, it is now thought that the exponent is not a rational number, but $1/\beta \approx 3.064$.

²This data is taken from P. Heller and G. Benedek, *Nuclear Magnetic Resonance in MnF_2 Near the Critical Point*, Phys. Rev. Lett. 8, 428 (1962).

This kind of behaviour is very surprising. It's rare that we see any kind of non-analytic behaviour in physics, but rarer still to find exponents that are not integers or simple fractions. What is going on here? This is one of the questions we will answer as these lectures progress.

1.2.4 A First Look at Universality

Before we go digging further into the Ising model, there is one other important aspect that deserves a mention at this point. The phase diagram for the Ising model is rather similar to the phase diagram for the transition between a liquid and gas, now drawn in the pressure-temperature plane. This is shown in the figure. In both cases, there is a line of first order transitions, ending at a critical point. For water, the critical point lies at a temperature $T_c \approx 374 \text{ }^\circ\text{C}$ and a pressure $p_c \approx 218 \text{ atm}$.

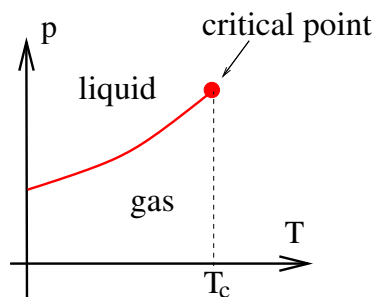


Figure 12:

The similarities are not just qualitative. One can use an appropriate equation of state for an interacting gas – say, the van der Waals equation – to compute how various quantities behave as we approach the critical point. (See the lectures on [Statistical Physics](#) for details of these calculations.) As we cross the first order phase transition, keeping the pressure constant, the volume V of the liquid/gas jumps discontinuously. This suggests that the rescaled volume $v = V/N$, where N is the number of atoms in the gas, is analogous to m in the Ising model. We can then ask how the jump in v changes as we approach the critical point. One finds,

$$v_{\text{gas}} - v_{\text{liquid}} \sim (T_c - T)^\beta \quad \text{where } \beta = \frac{1}{2}$$

From the phase diagram, we see that the pressure p is analogous to the magnetic field B . We could then ask how the volume changes with pressure as we approach the critical point keeping $T = T_c$ fixed. We find

$$v_{\text{gas}} - v_{\text{liquid}} \sim (p - p_c)^{1/\delta} \quad \text{where } \delta = 3$$

Finally, we want the analog of the magnetic susceptibility. For a gas, this is the compressibility, $\kappa = -\frac{1}{v} \frac{\partial v}{\partial p} \Big|_T$. As we approach the critical point, we find

$$\kappa \sim \frac{1}{|T - T_c|^\gamma} \quad \text{where } \gamma = 1$$

It has probably not escaped your attention that these critical exponents are exactly the same as we saw for the Ising model when treated in mean field. The same is also

true for the heat capacity C_V which approach different constant values as the critical point is approached from opposite sides.

However, the coincidence doesn't stop there. Because, it turns out, that the critical exponents above are also wrong! The true critical exponents for the liquid-gas transitions in $d = 2$ and $d = 3$ dimensions are the same as those of the Ising model, listed previously in the table. For example, experimental data for the critical exponent α of a number of different gases was plotted two pages back³ showing that it is approximately $\alpha \approx 0.1$, and certainly not consistent with mean field expectations.

This is an astonishing fact. It's telling us that at a second order phase transition, all memory of the underlying microscopic physics is washed away. Instead, there is a single theory which describes the physics at the critical point of the liquid-gas transition, the Ising model, and many other systems. This is a theoretical physicist's dream! We spend a great deal of time trying to throw away the messy details of a system to focus on the elegant essentials. But, at a critical point, Nature does this for us. Whatever drastic "spherical cow" approximation you make doesn't matter: if you capture the correct physics, you will get the exact answer! The fact that many different systems are described by the same critical point is called *universality*.

We might ask: does every second order phase transition have the same critical exponents as the Ising model? The answer is no! Instead, in each dimension d there is a set of critical points. Any system that undergoes a second order phase transition is governed by one member of this set. If two systems are governed by the same critical point, we say that they lie in the same *universality class*. The choice of universality class is, in large part, dictated by the symmetries of the problem; we will see some examples in Section 4.

The Ising Model as a Lattice Gas

It is, at first sight, surprising that a magnet and gas lie in the same universality class. However, there is a different interpretation of the Ising model that makes it look a little more gassy.

³The data is taken from J. Lipa, C. Edwards, and M. Buckingham *Precision Measurement of the Specific Heat of CO₂ Near the Critical Point*, Phys. Rev. Lett. 25, 1086 (1970). This was before the theory of critical phenomena was well understood. The data shows a good fit with $\alpha \approx 1/8$, not too far from the now accepted value $\alpha \approx 0.11$. Notice that the data stops around $t = 1 - T/T_c \approx 10^{-4}$. This is apparently because the effect of gravity becomes important as the critical point is approached, making experiments increasingly difficult.

To see this, consider the same d -dimensional lattice as before, but now with particles hopping between lattice sites. These particles have hard cores, so no more than one can sit on a single lattice site. We introduce the variable $n_i \in \{0, 1\}$ to specify whether a given lattice site, labelled by i , is empty ($n_i = 0$) or filled ($n_i = 1$). We can also introduce an attractive force between atoms by offering them an energetic reward if they sit on neighbouring sites. The Hamiltonian of such a lattice gas is given by

$$E = -4J \sum_{\langle ij \rangle} n_i n_j - \mu \sum_i n_i$$

where μ is the chemical potential which determines the overall particle number. But this Hamiltonian is trivially the same as the Ising model (1.1) if we make the identification

$$s_i = 2n_i - 1 \in \{-1, 1\}$$

The chemical potential μ in the lattice gas plays the role of magnetic field in the spin system while the magnetization of the system (1.6) measures the average density of particles away from half-filling.

There's no a priori reason to think that the Ising model is a particular good description of a gas. Nonetheless, this interpretation may make it a little less surprising that the Ising model and a gas share the same critical point.

1.3 Landau-Ginzburg Theory

The idea of universality – that many different systems enjoy the same critical point – is a powerful one. It means that if we want to accurately describe the critical point, we needn't concern ourselves with the messy details of any specific system. Instead, we should just search for the simplest model which gives the correct physics and work with that.

What is the simplest model? The Landau approach – in which the configuration of the Ising model is reduced to a single number m – is too coarse. This is because it misses any spatial variation in the system. And, as we will see shortly, the critical point is all about spatial variations. Here we describe a simple generalisation of Landau's ideas which allows the system to move away from a homogeneous state. This generalisation is known as *Landau-Ginzburg theory*⁴.

⁴It is also known as *Ginzburg-Landau* theory. The original paper, dating from 1950, has the authors in alphabetical order and constructs the free energy of a superconductor. Here we use the term Landau-Ginzburg theory to reflect the fact that, following Landau's earlier work, these ideas apply more broadly to any system.

The key idea is simple: we take the order parameter from Landau theory – which, for the Ising model, is the magnetisation m – and promote it to a field which can vary in space, $m(\mathbf{x})$. This is referred to as a *local order parameter*.

How do we do this? We started with a lattice. We divide this lattice up into many into boxes, with sides of length a . Each of these boxes contains many lattice sites – call it N' – but the box size a is smaller than any other length scale in the game. (In particular, a should be smaller than something called the *correlation length* that we will encounter below.) For each box, we define the average magnetisation $m(\mathbf{x}) = \frac{1}{N'} \sum_i s_i$, where \mathbf{x} denotes the centre of the box. This kind of procedure is known as *coarse-graining*.

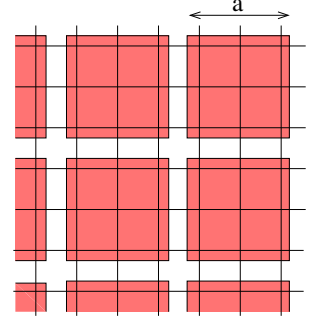


Figure 13:

There are all sorts of subtleties involved in this construction, and we will do our best to sweep them firmly under the rug. First, $m(\mathbf{x})$ takes values only at certain values of \mathbf{x} which label the positions of boxes. To mitigate this, we take the number of boxes N/N' to be big enough so that \mathbf{x} can be treated as a continuous variable. Second, at each \mathbf{x} , $m(\mathbf{x})$ is quantised in units of $1/N'$. We will take N' big enough so that $m(\mathbf{x})$ can take any value in $[-1, +1]$. The upshot of this is that we will treat $m(\mathbf{x})$ as a continuous function. However, we will at some point need to remember that it makes no sense for $m(\mathbf{x})$ to vary on very short distance scales – shorter than the separation between boxes.

You may have noticed that we’ve not been particularly careful about defining $m(\mathbf{x})$ and this cavalier attitude will continue below. You might, reasonably ask: does the physics depend on the details of how we coarse grain? The answer is no. This is the beauty of universality. We’ll see this more clearly as we proceed.

Our next step is to repeat the ideas that we saw in Section 1.1.1. Following (1.7), we write the partition function as

$$Z = \sum_{m(\mathbf{x})} \sum_{\{s_i\}|m(\mathbf{x})} e^{-\beta E[s_i]} := \sum_{m(\mathbf{x})} e^{-\beta F[m(\mathbf{x})]} \quad (1.24)$$

Here, the notation $\{s_i\}|m(\mathbf{x})$ means that we sum over all configurations of spins such that the coarse graining yields $m(\mathbf{x})$. In general, there will be many spin configurations for each $m(\mathbf{x})$. We then sum over all possible values of $m(\mathbf{x})$.

This procedure has allowed us to define a free energy $F[m(\mathbf{x})]$. This is a *functional*, meaning that you give it a function $m(\mathbf{x})$ and it spits back a number, $F[m(\mathbf{x})]$. This is known as the *Landau-Ginzburg free energy*

We will invoke one last notational flourish. We're left in (1.24) with a sum over all possible configurations $m(\mathbf{x})$. With the assumption that $m(\mathbf{x})$ is a continuous function, this is usually written as

$$Z = \int \mathcal{D}m(\mathbf{x}) e^{-\beta F[m(\mathbf{x})]} \quad (1.25)$$

This is a *functional integral*, also known as a *path integral*. The notation $\mathcal{D}m(\mathbf{x})$ – which is usually shortened to simply $\mathcal{D}m$ – means that we should sum over all field configurations $m(\mathbf{x})$.

Path integrals may be somewhat daunting at first sight. But it's worth remembering where it comes from: an integration over $m(\mathbf{x})$ at each point \mathbf{x} labelling a box. In other words, it's nothing more than an infinite number of normal integrals. We will start to learn how to play with these objects in Section 2.

The path integral looks very much like a usual partition function, but with the Landau-Ginzburg free energy $F[m(\mathbf{x})]$ playing the role of an effective Hamiltonian for the continuous variable $m(\mathbf{x})$. There is some nice intuition behind this. In the thermal ensemble, a given field configuration $m(\mathbf{x})$ arises with probability

$$p[m(\mathbf{x})] = \frac{e^{-\beta F[m(\mathbf{x})]}}{Z} \quad (1.26)$$

The path integral (1.25), is nothing but the usual partition function, now for a field $m(\mathbf{x})$ rather than the more familiar variables of classical physics like position and momentum. In other words, we're doing the statistical mechanics of fields, rather than particles. The clue was in the title of these lectures.

1.3.1 The Landau-Ginzburg Free Energy

The next step is to ask: how do we calculate $F[m(\mathbf{x})]$? This seems tricky: already in our earlier discussion of Landau theory, we had to resort to an unjustified mean field approximation. What are we going to do here?

The answer to this question is wonderfully simple, as becomes clear if we express the question in a slightly different way: what could the free energy $F[m(\mathbf{x})]$ possibly be? There are a number of constraints on $F[m(\mathbf{x})]$ that arise from its microscopic origin

- **Locality:** The nearest neighbour interactions of the Ising model mean that a spin on one site does not directly affect a spin on a far flung site. It only does so

through the intermediate spins. The same should be true of the magnetisation field $m(\mathbf{x})$. The result is that the free energy should take the form

$$F[m(\mathbf{x})] = \int d^d x f[m(\mathbf{x})]$$

where $f[m(\mathbf{x})]$ is a local function. It can depend on $m(\mathbf{x})$, but also on $\nabla m(\mathbf{x})$ and higher derivatives. These gradient terms control how the field $m(\mathbf{x})$ at one point affects the field at neighbouring points.

- **Translation and Rotation Invariance:** The original lattice has a discrete translation symmetry. For certain lattices (e.g. a square lattice) there can be discrete rotation symmetries. At distances much larger than the lattice scale, we expect that the continuum version of both these symmetries emerges, and our free energy should be invariant under them.
- **\mathbf{Z}_2 symmetry.** When $B = 0$, the original Ising model (1.1) is invariant under the symmetry $s_i \rightarrow -s_i$, acting simultaneously on all sites. This symmetry is inherited in our coarse-grained description which should be invariant under

$$m(\mathbf{x}) \rightarrow -m(\mathbf{x}) \tag{1.27}$$

When $B \neq 0$, the Ising model is invariant under $m(\mathbf{x}) \rightarrow -m(\mathbf{x})$, together with $B \rightarrow -B$. Again, our free energy should inherit this symmetry.

- **Analyticity:** We will make one last assumption: that the free energy density is an analytic function of $m(\mathbf{x})$ and its derivatives. Our primary interest lies in the critical point where m first becomes non-zero, although we will also use this formalism to describe first order phase transitions where $m(\mathbf{x})$ is small. In both cases, we can Taylor expand the free energy in m and restrict attention to low powers of m .

Furthermore, we will restrict attention to situations where $m(\mathbf{x})$ varies rather slowly in space. In particular, we will assume that $m(\mathbf{x})$ varies appreciably only over distances that are much larger than the distance a between boxes. This means that we can also consider a gradient expansion of $f[m(\mathbf{x})]$, in the dimensionless combination $a\nabla$. This means that ∇m terms are more important than $a\nabla^2 m$ terms and so on.

With these considerations in place, we can now simply write down the general form of the free energy. When $B = 0$, the symmetry (1.27) means that the free energy can

depend only on even powers of m . The first few terms in the expansion are

$$F[m(\mathbf{x})] = \int d^d x \left[\frac{1}{2} \alpha_2(T) m^2 + \frac{1}{4} \alpha_4(T) m^4 + \frac{1}{2} \gamma(T) (\nabla m)^2 + \dots \right] \quad (1.28)$$

There can also be a $F_0(T)$ piece – like the $T \log 2$ that appeared in the Landau free energy – which doesn’t depend on the order parameter and so, for the most part, will play no role in the story. Notice that we start with terms quadratic in the gradient: a term linear in the gradient would violate the rotational symmetry of the system.

When $B \neq 0$, we can have further terms in the free energy that are odd in m , but also odd in B , such as Bm and Bm^3 . Each of these comes with a coefficient which is, in general, a function of T .

The arguments that led us to (1.28) are very general and powerful; we will see many similar arguments in Section 4. The downside is that we are left with a bunch of unknown coefficients $\alpha_2(T)$, $\alpha_4(T)$ and $\gamma(T)$. These are typically hard to compute from first principles. One way forward is to import our results from Landau mean field approach. Indeed, for constant $m(\mathbf{x}) = m$, the free energy (1.28) coincides with our earlier result (1.13) in Landau theory, with

$$\alpha_2(T) \sim (T - T_c) \quad \text{and} \quad \alpha_4(T) \sim \frac{1}{3} T \quad (1.29)$$

Happily, however, the exact form of these functions will not be important. All we will need is that these functions are analytic in T , and that $\alpha_4(T) > 0$ and $\gamma(T) > 0$, while $\alpha_2(T)$ flips sign at the second order phase transition.

Looking ahead, there is both good news and bad. The good news is that the path integral (1.25), with Landau-Ginzburg free energy (1.28), does give a correct description of the critical point, with the surprising d -dependent critical exponents described in Section 1.2.3. The bad news is that this path integral is hard to do! Here “hard” means that many of the unsolved problems in theoretical physics can be phrased in terms of these kinds of path integrals. Do not fear. We will tread lightly.

1.3.2 The Saddle Point and Domain Walls

We are going to build up slowly to understand how we can perform the functional integral (1.25). As a first guess, we’ll resort to the saddle point method and assume that the path integral is dominated by the configurations which minimise $F[m(\mathbf{x})]$. In subsequent sections, we’ll treat the integral more seriously and do a better job.

To find the minima of functionals like $F[m(\mathbf{x})]$, we use the same kind of variational methods that we met when working with Lagrangians in [Classical Dynamics](#). We take some fixed configuration $m(\mathbf{x})$ and consider a nearby configuration $m(\mathbf{x}) + \delta m(\mathbf{x})$. The change in the free energy is then

$$\begin{aligned}\delta F &= \int d^d x \left[\alpha_2 m \delta m + \alpha_4 m^3 \delta m + \gamma \nabla m \cdot \nabla \delta m \right] \\ &= \int d^d x \left[\alpha_2 m + \alpha_4 m^3 - \gamma \nabla^2 m \right] \delta m\end{aligned}$$

where, to go from the first line to the second, we have integrated by parts. This encourages us to introduce the *functional derivative*,

$$\frac{\delta F}{\delta m(\mathbf{x})} = \alpha_2 m(\mathbf{x}) + \alpha_4 m^3(\mathbf{x}) - \gamma \nabla^2 m(\mathbf{x})$$

Note that I've put back the \mathbf{x} dependence to stress that, in contrast to F , $\delta F/\delta m(\mathbf{x})$ is evaluated at some specific position \mathbf{x} .

If the original field configuration $m(\mathbf{x})$ was a minimum of the free energy it satisfies the Euler-Lagrange equations,

$$\left. \frac{\delta F}{\delta m} \right|_{m(\mathbf{x})} = 0 \quad \Rightarrow \quad \gamma \nabla^2 m = \alpha_2 m + \alpha_4 m^3 \tag{1.30}$$

The simplest solutions to this equation have m constant. This recovers our earlier results from Landau theory. When $T > T_c$, we have $\alpha_2 > 0$ and the ground state has $m = 0$. In contrast, when $T < T_c$, $\alpha_2 < 0$ and there is a degenerate ground state $m = \pm m_0$ with

$$m_0 = \sqrt{-\frac{\alpha_2}{\alpha_4}} \tag{1.31}$$

This is the same as our previous expression (1.14), where we replaced α_2 and α_4 with the specific functions (1.29). We see that what we previously called the mean field approximation, is simply the saddle point approximation in Landau-Ginzburg theory. For this reason, the term “mean field theory” is given to any situation where we write down the free energy, typically focussing on a Taylor expansion around $m = 0$, and then work with the resulting Euler-Lagrange equations.

Domain Walls

The Landau-Ginzburg theory contains more information than our earlier Landau approach. It also tells us how the magnetisation changes in space.

Suppose that we have $T < T_c$ so there exist two degenerate ground states, $m = \pm m_0$. We could cook up a situation in which one half of space, say $x < 0$, lives in the ground state $m = -m_0$ while the other half of space, $x > 0$ lives in $m = +m_0$. The two regions in which the spins point up or down are called domains. The place where these regions meet is called the *domain wall*.

We would like to understand the structure of the domain wall. How does the system interpolate between these two states? The transition can't happen instantaneously because that would result in the gradient term $(\nabla m)^2$ giving an infinite contribution to the free energy. But neither can the transition linger too much because any point at which $m(\mathbf{x})$ differs significantly from the value m_0 costs free energy from the m^2 and m^4 terms. There must be a happy medium between these two.

To describe the system with two domains, $m(\mathbf{x})$ must vary but it need only change in one direction: $m = m(x)$. Equation (1.30) then becomes an ordinary differential equation,

$$\gamma \frac{d^2 m}{dx^2} = \alpha_2 m + \alpha_4 m^3$$

This equation is easily solved. If we insist that the field interpolate between the two different ground states, $m \rightarrow \mp m_0$ as $x \rightarrow \mp \infty$, then the solution is given by

$$m = m_0 \tanh\left(\frac{x - X}{W}\right) \quad (1.32)$$

This is plotted in the figure. Here X is the position of the domain wall and

$$W = \sqrt{-\frac{2\gamma}{\alpha_2}}$$

is its width. For $|x - X| > W$, the magnetisation relaxes exponentially quickly back to the ground state values $\pm m_0$.

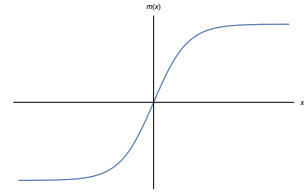


Figure 14:

We can also compute the cost in free energy due to the presence of the domain wall. To do this, we substitute the solution back into the expression for the free energy (1.28). The cost is not proportional to the volume of the system, but instead proportional to the area of the domain wall. This means that if the system has linear size L then the free energy of the ground state scales as L^d while the additional free energy required by the wall scales only as L^{d-1} . It is simple to see that the excess cost in free energy of the domain wall has parametric dependence

$$F_{\text{wall}} \sim L^{d-1} \sqrt{-\frac{\gamma \alpha_2^3}{\alpha_4^2}} \quad (1.33)$$

Here's the quick way to see this: from the solution (1.32), the field m changes by $\Delta m \sim m_0$ over a distance $\Delta x \sim W$. This means that $dm/dx \sim \Delta m/\Delta x \sim m_0/W$ and the gradient term in the free energy contributes $F \sim L^{d-1} \int dx \gamma (dm/dx)^2 \sim L^{d-1} W \gamma (m_0/W)^2$, where the second expression comes because the support of the integral is only non-zero over the width W where the domain wall is changing. This gives the parametric dependence (1.33). There are further contributions from the m^2 and m^4 terms in the potential, but our domain wall solves the equation of motion whose purpose is to balance the gradient and potential terms. This means that the potential terms contribute with the same parametric dependence as the gradient terms, a fact you can check by hand by plugging in the solution.

Notice that as we approach the critical point, and $\alpha_2 \rightarrow 0$, the two vacua are closer, the width of the domain wall increases and its energy decreases.

1.3.3 The Lower Critical Dimension

We stated above that the Ising model has no phase transition in $d = 1$ dimensions. The reason behind this can be traced to the existence of domain walls, which destroy any attempt to sit in the ordered phase.

Let's set $\alpha_2(T) < 0$ where we would expect two, ordered ground states $m = \pm m_0$. To set up the problem, we start by considering the system on a finite interval of length L . At one end of this interval – say the left-hand edge $x = -L/2$ – we'll fix the magnetisation to be in its ground state $m = +m_0$. One might think that the preferred state of the system is then to remain in $m = +m_0$ everywhere. But what actually happens?

There is always a probability that a domain wall will appear in the thermal ensemble and push us over to the other ground state $m = -m_0$. The probability for a wall to appear at some point $x = X$ is given by (1.26)

$$p[\text{wall at } x = X] = \frac{e^{-\beta F_{\text{wall}}}}{Z}$$

This looks like it's exponentially suppressed compared to the probability of staying put. However, we get an enhancement because the domain wall can sit anywhere on the line $-L/2 \leq x \leq L/2$. This means that the probability becomes

$$p[\text{wall anywhere}] = \frac{e^{-\beta F_{\text{wall}}}}{Z} \frac{L}{W}$$

For a large enough system, the factor of L/W will overwhelm any exponential suppression. This is an example of the entropy of a configuration – which, in this context is $\log(L/W)$ – outweighing the energetic cost.

Of course, once we've got one domain wall there's nothing to stop us having another, flipping us back to $m = +m_0$. If we have an even number of domain walls along the line, we will be back at $m = m_0$ by the time we get to the right-hand edge at $x = +L/2$; an odd number and we'll sit in the other vacuum $m = -m_0$. Which of these is most likely?

The probability to have n walls, placed anywhere, is

$$p[n \text{ walls}] = \frac{e^{-n\beta F_{\text{wall}}}}{Z} \frac{1}{W^n} \int_{-L/2}^{L/2} dx_1 \int_{x_1}^{L/2} dx_2 \dots \int_{x_{n-1}}^{L/2} dx_n = \frac{1}{Z n!} \left(\frac{L e^{-\beta F_{\text{wall}}}}{W} \right)^n$$

This means that the probability that we start at $m = m_0$ and end up at $m = m_0$ is

$$p[m_0 \rightarrow m_0] = \frac{1}{Z} \sum_{n \text{ even}} \frac{1}{n!} \left(\frac{L e^{-\beta F_{\text{wall}}}}{W} \right)^n = \frac{1}{Z} \cosh \left(\frac{L e^{-\beta F_{\text{wall}}}}{W} \right)$$

Meanwhile, the probability that the spin flips is

$$p[m_0 \rightarrow -m_0] = \frac{1}{Z} \sum_{n \text{ odd}} \frac{1}{n!} \left(\frac{L e^{-\beta F_{\text{wall}}}}{W} \right)^n = \frac{1}{Z} \sinh \left(\frac{L e^{-\beta F_{\text{wall}}}}{W} \right)$$

We see that, at finite L , there is some residual memory of the boundary condition that we imposed on the left-hand edge. However, as $L \rightarrow \infty$, this gets washed away. You're just as likely to find the spins up as down on the right-hand edge.

Although we phrased this calculation in terms of pinning a choice of ground state on a boundary, the same basic physics holds on an infinite line. Indeed, this is a general principle: whenever we have a Landau-Ginzburg theory characterised by a discrete symmetry – like the \mathbf{Z}_2 of the Ising model – then the ordered phase will have a number of degenerate, disconnected ground states which spontaneously break the symmetry. In all such cases, the lower critical dimension is $d_l = 1$ and in all cases the underlying reason is the same: fluctuations of domain walls will take us from one ground state to another and destroy the ordered phase. It is said that the domain walls *proliferate*.

We could try to run the arguments above in dimensions $d \geq 2$. The first obvious place that it fails is that the free energy cost of the domain wall (1.33) now scales with the system size, L . This means that as L increases, we pay an exponentially large cost from $e^{-F_{\text{wall}}}$. Nonetheless, one could envisage a curved domain wall, which encloses some finite region of the other phase. It turns out that this is not sufficient to disorder the system. However, in $d = 2$, the fluctuations of the domain wall become important as one approaches the critical point.

1.3.4 Lev Landau: 1908-1968

Lev Landau was one of the great physicists of the 20th century. He made important contributions to nearly all areas of physics, including the theory of magnetism, phase transitions, superfluids and superconductors, Fermi liquids and quantum field theory. He founded a school of Soviet physics whose impact lasts to this day.

Landau was, by all accounts, boorish and did not suffer fools gladly. This did not sit well with the authorities and, in 1938, he was arrested in one of the great Soviet purges and sentenced to 10 years in prison. He was rescued by Kapitza who wrote a personal letter to Stalin arguing, correctly as it turned out, that Landau should be released as he was the only person who could understand superfluid helium. Thus, for his work on superfluids, Landau was awarded both his freedom and, later, the Nobel prize. His friend, the physicist Yuri Rumer, was not so lucky, spending 10 years in prison before exile to Siberia.

Legend has it that Landau hated to write. The extraordinarily ambitious, multi-volume “Course on Theoretical Physics” was entirely written by his co-author Evgeny Lifshitz, usually dictated by Landau. As Landau himself put it: “Evgeny is a great writer: he cannot write what he does not understand”.

Here is a story about a visit by Niels Bohr to Landau’s Moscow Institute. Bohr was asked by a journalist how he succeeded in creating such a vibrant atmosphere in Copenhagen. He replied “I guess the thing is, I’ve never been embarrassed to admit to my students that I’m a fool”. This was translated by Lifshitz as: “I guess the thing is, I’ve never been embarrassed to admit to my students that they’re fools”. According to Kapitza, the translation was no mistake: it simply reflected the difference between Landau’s school and Bohr’s.

In 1962, Landau suffered a car accident. He survived but never did physics again.



Figure 15: