May 3, 2024

Euclidean Field Theory

Kasper Peeters & Marija Zamaklar

Lecture notes for the M.Sc. in Elementary Particle Theory at Durham University.

Department of Mathematical Sciences
University of Durham
South Road
Durham DH1 3LE
United Kingdom

kasper.peeters@durham.ac.uk

marija.zamaklar@durham.ac.uk

1

Introducing Euclidean Field Theory

1.1. Executive summary and recommended literature

This course is all about the close relation between two subjects which at first sight have very little to do with each other: *quantum field theory* and the *theory of critical statistical phenomena*. On the one hand, the methods and insight from quantum field theory have helped tremendously to understand the concept of "universality" in statistical mechanics. This is the fact that the effective theory relevant for long distance behaviour of statistical systems depends only on a small number of parameters. On the other hand, the theory of critical phenomena has shed new light on the role of "renormalisation" in particle physics. Meant here is the way in which subtle tuning of parameters of the model is required to remove, or make sense of, spurious infinities in loop computations. It is the goal of these lectures to explain how these two fields are related.

Given that this course is about the common ground of two quite different disciplines, it is no surprise that there is a vast literature to help you with many technical and conceptual issues. The list below is highly incomplete, but you are strongly encouraged to consult some of these books to get a wider view on the topics discussed here.

- M. Le Bellac, "Quantum and statistical field theory", Oxford University Press, 1991.

  Good for a number of explicit computations, though not terribly transparent in its connection between the two fields, at least on first reading.

- J. L. Cardy, "Scaling and renormalization in statistical physics", Cambridge University Press, 1996.

  Standard book on the topic, gives a good overview, but is somewhat thin on explicit computations.

- J. M. Yeomans, "Statistical mechanics of phase transitions", Oxford University Press, 1992.

  Very good, compact and explicit book on the statistical aspects of phase transitions.

- J. B. Kogut and M. A. Stephanov, "The phases of quantum chromodynamics: From confinement to extreme environments", 2004.

  Mainly focussed on applications to QCD, but contains some good introductory chapters as well.

- A. M. Polyakov, "Gauge fields and strings", Harwood, 1987. Contemporary concepts in physics, volume 3.

  An original take on many of the topics discussed here, though sometimes a bit too compact for a first introduction.

- M. H. Thoma, "New developments and applications of thermal field theory", `hep-ph/0010164`.

  Very clear text on quantum field theory at finite temperature, in particular the connection between the imaginary time and real time formalisms. Also contains a number of explicit applications to QCD.

- G. P. Lepage, "What is Renormalization?", `hep-ph/0506330`.

  Lectures on the meaning of renormalisation from the point of view of a particle physics theorist.

- M. Creutz, "Quarks, gluons and lattices", Cambridge University Press, 1983.

  Another very accessible intro to lattice field theory.

- I. Montvay and G. Münster, "Quantum fields on a lattice", Cambridge University Press, 1994.

  More advanced book on lattice field theory. Good examples on the strong-coupling expansion.

- J. Smit, "Introduction to quantum fields on a lattice: A robust mate", Cambridge Lect. Notes Phys., 2002.

  Nice and compact book about lattice field theory. Used heavily in the preparation of these lectures.

- J. Zinn-Justin, "Quantum field theory and critical phenomena", *Int. Ser. Monogr. Phys.* **113** (2002) 1–1054.

  A thousand page reference book, not for the faint of heart, and not always organised in the most pedagogical way. Nevertheless, a very good source if you want to look up details.

- J. Zinn-Justin, "Quantum field theory at finite temperature: An introduction", `hep-ph/0005272`.

  A much smaller set of lectures exposing the relation between finite temperature quantum field theory and statistical mechanics. The first few sections are worth reading as an introduction to the topic.

1.2. Path integrals and partition sums

The root of the connection between quantum field theory and statistical systems is the functional integral (or path integral). This functional integral is closely related to an object in classical statistical mechanics: the partition sum. In order to illustrate this, let us briefly recall the quantum mechanics of a non-relativistic particle in the path integral language.

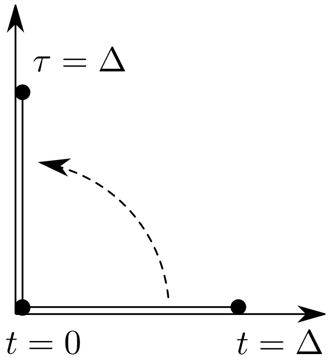
Let us start with the transition amplitude for such a particle to go from q to q' in time Δ. It is given by

$$\mathcal{A}(q',\Delta;q,0) = \int_{q(t=0)=q}^{q(t=\Delta)=q'} \mathcal{D}q(t)\;\exp\Big[\frac{i}{\hbar}\int_0^{\Delta}\Big(\frac{m}{2}\dot{q}^i(t)\dot{q}^i(t) - V(q(t))\Big)\,\mathrm{d}t\Big]\,. \tag{1.1}$$

As you have seen in the course on quantum field theory, this path integral expression can be related to the usual expression for a transition amplitude in the operator formalism. For a suitable normalisation of the path integral measure, we have

$$\mathcal{A}(q',\Delta;q,0) = \langle q'|\exp\Big[-\frac{i}{\hbar}\hat{H}\Delta\Big]|q\rangle\,,\quad \text{with}\quad \hat{H} = \frac{\hat{p}^2}{2m} + V(q)\,. \tag{1.2}$$

Here $|q\rangle$ is a time-independent state in the Heisenberg picture, and $\hat{H}$ the Hamiltonian corresponding to the particle action used in (1.1). For the details of this relation, see your favourite quantum mechanics book.



The Wick rotation in the complex time plane.

Consider now what happens to the path integral expression if we make a substitution $t \rightarrow -i\tau$. This is called a *Wick rotation* (it rotates the integration in the complex time plane by 90°, as in the figure). Provided the Lagrangian does not depend explicitly on t (which, in our case, it does not), this changes the path integral weight into a nice real expression,

$$\mathcal{A}(q',\Delta;q,0) = \int_{q(t=0)=q}^{q(\tau=\Delta)=q'} \mathcal{D}q(\tau)\;\exp\Big[-\frac{1}{\hbar}\int_0^{\Delta}\Big(\frac{m}{2}\dot{q}^i(\tau)\dot{q}^i(\tau) + V(q(\tau))\Big)\,\mathrm{d}\tau\Big]\,, \tag{1.3}$$

where the dot now indicates a derivative with respect to τ. Similarly, if we substitute $\Delta \rightarrow -i\Delta$ in the operator expression, we get

$$\mathcal{A}(q',\Delta;q,0) = \langle q'|\exp\Big[-\frac{1}{\hbar}\hat{H}\Delta\Big]|q\rangle\,. \tag{1.4}$$

This expression should make some bells ring. If we would set $|q\rangle = |q'\rangle$, and sum over all position eigenstates, then (1.4) looks exactly like the expression for the partition function $Z[\beta]$ of a statistical system at inverse temperature $\beta := 1/(kT)$ given by $\beta = \Delta/\hbar$,

$$Z[\beta] = \sum_n \langle n|\exp\Big[-\beta\hat{H}\Big]|n\rangle\,, \tag{1.5}$$

where we sum over a complete basis of states. This gives us our *first* relation between a quantum mechanical transition amplitude with periodic boundary conditions and a quantum statistical system at non-zero temperature.

The quantum mechanical transition amplitude for a system with periodic boundary conditions is, after summing over all boundary conditions and Wick rotation, equal to a quantum statistical partition function, with $\beta \leftrightarrow \Delta/\hbar$.

Let us now take a somewhat different look at the exponent in (1.3). We can write it as

$$V_{\text{string}} = -\frac{1}{\hbar}\int_0^{\Delta}\Big(\frac{m}{2}\dot{q}^i(\tau,\tilde{\tau})\dot{q}^i(\tau,\tilde{\tau}) + V(q(\tau))\Big)\,\mathrm{d}\tau\,, \tag{1.6}$$

i.e. as the potential energy of a field $q^i(\tau,\tilde{\tau})$ in two dimensions, where we now view $\tilde{\tau}$ as the time (the total energy would have an additional kinetic term involving $\tilde{\tau}$ derivatives). The path integral then simply sums over time-independent configurations of this field. We thus see that we can also interpret (1.3) as computing the partition function of a *classical string* of length Δ, at an inverse temperature $\beta = 1/\hbar$. This gives us our *second* relation, between a quantum mechanical transition function for a particle and a classical statistical system for a string.

The transition amplitude for a quantum particle for a time $-i\Delta$ is equal to the classical partition function of a string of length Δ, with $\beta \leftrightarrow 1/\hbar$.

The second of these relations is easily generalised to quantum field theory. In a Lorentzian field theory, the "generating functional in the presence of a source" takes the form

$$Z[J] = \int \mathcal{D}\phi \, \exp\left[\frac{i}{\hbar}S[\phi] + \int \mathrm{d}^{d+1}x\, J(x)\phi(x) - \frac{\epsilon}{\hbar}\int \mathrm{d}^{d+1}x\, \phi(x)^2\right] . \tag{1.7}$$

(end of lecture 1)

where $S[\phi]$ is the action corresponding to the field ϕ. We have included a damping factor dependent on ϵ, which ensures that field configurations which do not fall off at space-like and time-like infinity are suppressed in the path integral. This ensures that all correlation functions which we compute from $Z[J]$ are correlation functions in the vacuum state, e.g.

$$\langle 0|\, T\Big(\hat{\phi}(x)\hat{\phi}(y)\Big)\, |0\rangle = \frac{1}{Z[J]}\frac{\delta}{\delta J(x)}\frac{\delta}{\delta J(y)} Z[J]\Big|_{J=0} . \tag{1.8}$$

We can now do a Wick rotation, and then interpret the exponent in (1.7) as the potential of a field ϕ in one extra dimension. The path integral then computes for us the classical partition sum of a model at inverse temperature $\beta = 1/\hbar$.

The first relation is a bit more tricky to generalise to field theory, because of the explicit dependence of the statistical model on the time interval. In field theory we usually take the time interval to $\pm\infty$. Moreover, we never consider arbitrary field configurations at the endpoint of this time interval, but rather force the field to sit in the vacuum state. So summing over boundary conditions does not produce a sensible object in quantum field theory. However, what we can do is turn the first relation around, and use it to describe a quantum field theory at finite temperature in terms of a Euclidean path integral. This is the topic of the next section.

▸ ***Summary:*** The generating functional of a Euclidean quantum field theory in d space dimensions can alternatively be interpreted as a partition sum in a classical statistical model in d dimensions. The corresponding inverse temperature is then related to Planck's constant by $\beta \sim 1/\hbar$.

▸ ***See also:*** Discussions like these can be found in many places, e.g. chapter 2 of the book by Smit [10] or chapter V.2 of the book by Zee [13].

1.3. Finite temperature quantum field theory

In standard quantum field theory courses one almost always works with systems at vanishing temperature. That is to say, the state which is used to define all correlation functions is taken to be the vacuum state $|0\rangle$ of the free theory, with no particles. The propagator, for instance, is defined as

$$G_F(x-y) = \langle 0|\, T\Big(\hat{\phi}(x)\hat{\phi}(y)\Big)\, |0\rangle\, . \tag{1.9}$$

However, we can certainly also construct a quantum theory of fields in which we use a thermal state instead of the no-particle vacuum. Instead of computing a vacuum expectation value, we then have to compute

$$G_F^{T>0}(x-y) = \frac{\sum_n \langle n|\, T\Big(\hat{\phi}(x)\hat{\phi}(y)\Big)\, |n\rangle\, e^{-\beta E_n}}{\sum_n e^{-\beta E_n}} . \tag{1.10}$$

Here $|n\rangle$ is an eigenstate of the Hamiltonian with eigenvalue E_n. Such states are generically complicated multi-particle states. As particles can be created with any three-momentum, and in any number, they take the form

$$\big|\{n_1,\ldots,n_M;\vec{k}_1,\ldots,\vec{k}_M\}\big\rangle = \prod_{i=1}^{M} \frac{\left(\hat{a}^\dagger(\vec{k}_i)\right)^{n_i}}{\sqrt{n_i!}}|0\rangle\,. \tag{1.11}$$

This state has $\sum_i n_i$ particles, and the generic label n that appears in (1.10) is seen to describe both the set of momenta that appear, as well as the number of particles of each momentum.

We will leave this expression for the time being, and instead ask the question whether we can write (1.10) back in a path integral language. It is not immediately clear how to do this, because the Gaussian factor in the path integral (1.7) enforces the appearances of the $|0\rangle$ state in the correlator. The way to continue is to observe that (1.10) is an expression similar to (1.4). We can thus take it from there, and go backwards, i.e. rewrite the thermal sum over energy eigenstates in terms of a path integral, as we did for the non-relativistic particle. We view the right-hand side of (1.10) as a propagator for a field theory over a compact Euclidean time with period β, with periodic boundary conditions in this direction.

In fact, this additional Euclidean direction is simply the imaginary part of the time coordinate which is already present in x^μ and y^μ. That is to say, the Green function (1.10) is periodic under a shift of the time difference by $-i\hbar\beta$:

$$\begin{aligned}
G_F^{T>0}(x^0,\vec{x};0,\vec{y}) &= \frac{1}{Z}\,\mathrm{Tr}\left[e^{-\beta\hat{H}}T\Big(\hat{\phi}(x^0,\vec{x})\hat{\phi}(0,\vec{y})\Big)\right]\\
&= \frac{1}{Z}\,\mathrm{Tr}\left[e^{-\beta\hat{H}}\hat{\phi}(x^0,\vec{x})\hat{\phi}(0,\vec{y})\right] \qquad (\text{assuming } x^0>0)\\
&= \frac{1}{Z}\,\mathrm{Tr}\left[\hat{\phi}(0,\vec{y})e^{-\beta\hat{H}}\hat{\phi}(x^0,\vec{x})\right] \qquad\qquad (1.12)\\
&= \frac{1}{Z}\,\mathrm{Tr}\left[e^{-\beta\hat{H}}e^{\beta\hat{H}}\hat{\phi}(0,\vec{y})e^{-\beta\hat{H}}\hat{\phi}(x^0,\vec{x})\right]\\
&= \frac{1}{Z}\,\mathrm{Tr}\left[e^{-\beta\hat{H}}T\Big(\hat{\phi}(x^0,\vec{x})\hat{\phi}(-i\hbar\beta,\vec{y})\Big)\right] = G_F^{T>0}(x^0,\vec{x};-i\hbar\beta,\vec{y})\,.
\end{aligned}$$

The last step, in which we re-introduce time-ordering, requires that we can compare x^0 with $-i\hbar\beta$. That is possible if $x^0 = -i\tau$, i.e. if the time coordinate was imaginary to start with. The derivation thus goes through only for correlation functions which are independent of the real time. For time-dependent situations, we need what is called the *real-time formalism* for thermal quantum field theory, which goes beyond the Euclidean Field Theory concepts which we focus on in these lectures.

▸ ***Summary:*** The *static* properties of a quantum field theory at finite temperature can be obtained by doing a Wick rotation and imposing periodic boundary conditions in imaginary time.[1] The periodic boundary conditions of the path integral correspond to the trace of the thermal sum, and the period corresponds to the inverse temperature $\Delta\tau \leftrightarrow \beta$.

▸ ***See also:*** The lectures by Thoma [6].

---

[1] You may have heard about Hawking radiation of black holes, and the fact that it has a certain temperature. The simplest way to see that black holes have a temperature is by looking at the analytic continuation of a black hole metric in the complex plane. It turns out that the imaginary part of time needs to be compact, and hence that any field theory on top of the black hole background is in a thermal state.

1.4. Phase transitions and critical phenomena

Although we have quickly stepped over the relation between the path integral and the quantum mechanical transition amplitude, a crucial ingredient in the derivation of that relation is a proper definition of the path integral measure. The infinite-dimensional integral is only well-defined if we regularise the model somehow, for instance by putting it on a lattice (which, in the quantum-mechanical model, means discretising time). As a result, *all* the systems that we are dealing with in these lectures are in a sense *discrete* systems, with a countable number N of degrees of freedom. The main question we are then interested in is how such systems behave when we change their parameters, when we change the lattice spacing, and when we take the thermodynamic limit $N \to \infty$.

From statistical systems, we know that many interesting things can happen. To illustrate some of these, let us focus on an example which will return several more times during these lectures: the Ising model. This model has its roots in the search for a simple model which describes a ferromagnet. The main property of a ferromagnet is that it retains its magnetisation in the absence of an external magnetic field. However, for sufficiently large temperature, this magnetisation disappears.

(end of lecture 2)

The key ingredient is the interaction between the electron spins at different sites of the atomic lattice. A simple model of this type is the *Heisenberg model*, with Hamiltonian operator

$$\hat{H}_{\text{Heisenberg}} = -J \sum_{\langle i,j \rangle} \vec{\sigma}_i \cdot \vec{\sigma}_j \,, \quad J > 0 \,. \tag{1.13}$$

Here the $\vec{\sigma}_i$ are the Pauli matrices acting on the spin degrees of freedom at site i. The notation $\langle i,j \rangle$ indicates a pair of spins at site i and j. Unfortunately, this model can only be solved on a one-dimensional lattice (Bethe 1931). In that case, for $J > 0$ the ground state is one in which all spins are aligned (while for $J < 0$ the ground state is the so-called Neel state, with alternating spins). Even in the classical limit, in which we replace the spin operator $\vec{\sigma}_i$ at each site by a classical three-vector $\vec{S}_i$, the model is not solvable except in one dimension.

The Ising model incorporates one last simplification, namely the reduction of the three-vector at each site to a simple scalar S_i, which can take the values ± 1. Its classical Hamiltonian, in the presence of an external magnetic field, reads

$$H_{\text{Ising}} = -J \sum_{\langle i,j \rangle} S_i S_j - h \sum_i S_i \,. \tag{1.14}$$

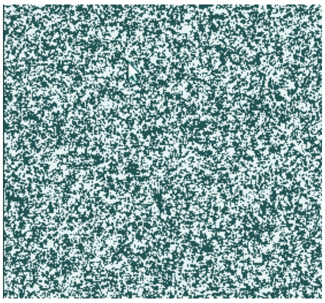
This model was introduced by Lenz in the 1920's, and solved by Ising in the one-dimensional case. In the absence of a magnetic field h, the model was solved by Onsager in two dimensions. No other exact solutions are known.

The parameters of this model are thus the coupling J, the external field h and the number of sites N. All distances between spins are determined by the lattice spacing a (which in this simple model does not actually appear in the Hamiltonian). The partition sum, finally, adds the temperature T to the set of parameters. What we want to learn is how we can do measurements on this system, how we can express those measurements in terms of J, h, N, a and T, and how these behave when we vary the parameters and take the $N \to \infty$ limit. This same story holds for all the other discrete models that appear in these lectures.

What are the observables which we can measure in a statistical model of the type (1.14)? In general, these are correlation functions, just like in a quantum field theory. The simplest correlation function is the expectation value of a single spin, which is called the *local magnetisation*. The expectation value of the sum of all spins is the *total magnetisation*,[2]

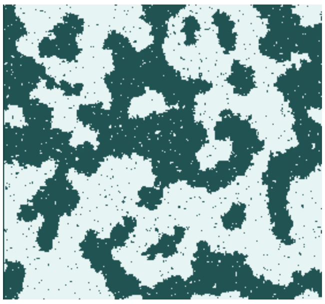
$$M := \frac{\sum\limits_{\{S_i\}} \left[\big(\sum\limits_i S_i\big) e^{-\beta H_{\text{Ising}}[\{S_i\}]} \right]}{\sum\limits_{\{S_i\}} e^{-\beta H_{\text{Ising}}[\{S_i\}]}} \,. \tag{1.16}$$

Naively, one would perhaps be tempted to say that $M = 0$ when the background magnetic field h vanishes: for every configuration $\{S_i\}$ there exists another configuration in which all the spins are flipped, and these two configurations contribute with opposite sign to the partition sum.



The 2d Ising model at high temperature, in the absence of a magnetic field. The spins are randomly fluctuating, there is no magnetisation, and the model is thus paramagnetic (dark and light denote spin up and down, respectively).

While this is mathematically true, we have to ask the question what happens physically to a configuration of aligned spins, i.e. with net magnetisation. When the temperature is very high, the thermal fluctuations will indeed wash out such magnetisation, as in the figure at the top. However, for sufficiently low temperatures, it becomes very hard for the thermal fluctuations to sample the entire configuration space, and we *do* see large connected regions with a net magnetisation. In this regime, the spin up/down symmetry of the Hamiltonian is *spontaneously broken* by the state of the system.



The 2d Ising model at low temperature. Large connected regions have formed, in which the spin is predominantly in one direction and the magnetisation is non-zero. The model is ferromagnetic.

In the continuum limit $N \to \infty$, this transition becomes sharp. Mathematically, the absence of a phase transition for a finite system is expressed by the fact that the two limits $N \to \infty$ and $h \to 0$ do not commute,

$$\begin{aligned} &\lim_{N\to\infty} \lim_{h\to 0} M(h) = 0 \,, \\ &\lim_{h\to 0} \lim_{N\to\infty} M(h) \neq 0 \,. \end{aligned} \tag{1.17}$$

From a practical point of view these expressions are, however, of limited value.

We will more generically see that discrete models exhibit remnants of phase transitions, which are in some sense 'smoothed out' versions of the continuum transition. The more degrees of freedom we include (the larger N) the closer we get to the proper phase transition of the continuum model.

Moving on to higher order correlation functions, the next one is the correlator of spins at two different sites,

$$\langle S_i S_j \rangle := \frac{\sum\limits_{\{S_i\}} \left[S_i S_j \; e^{-\beta H_{\text{Ising}}[\{S_i\}]} \right]}{\sum\limits_{\{S_i\}} e^{-\beta H_{\text{Ising}}[\{S_i\}]}} \,. \tag{1.18}$$

This correlator will be a non-trivial function of the distance between the two sites. If there are n sites between i and j, so that the distance is $r = n\,a$, the correlator typically behaves at large distances as

$$\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle = \langle S_i S_j \rangle - \left(\frac{M}{N} \right)^2 \sim \frac{1}{r^{D+\eta-2}} \exp\left[-\frac{r}{\xi} \right] \,. \tag{1.19}$$

The length ξ is called the *correlation length*, D is the dimension of the system, and η is called the *anomalous dimension* (to which we will return in much greater detail later).

---

[2] The total magnetisation can be obtained from the partition sum in the presence of a magnetic field by simply taking a derivative,

$$M = \frac{1}{\beta} \frac{\partial \ln Z}{\partial h} \,. \tag{1.15}$$

For more complicated correlators, a similar trick requires that we know the partition function coupled to a non-homogeneous field.

As is clear from the figures, the correlation length becomes large when we approach the phase transition, and we will see that it in fact diverges in a very particular way as we approach the critical temperature.

We will often use a thermodynamical language to describe the large-distance behaviour of statistical systems. Without trying to include a full course on thermodynamics, let us briefly recall some essential concepts. The basic quantity we need is the *entropy*. At the microscopic level, it is defined in terms of the number of microstates Ω given the macroscopic parameters (like magnetisation),

$$S := k_B \ln \Omega\,. \tag{1.20}$$

It can also be expressed in terms of the probability of a single state $p(\{S_i\})$ to occur, using *Shannon's formula*,

$$S = -k_B \sum_{\{S_i\}} p(\{S_i\}) \ln \Big[p(\{S_i\})\Big]\,. \tag{1.21}$$

This allows us to relate the entropy to the energy and eventually to the partition function, as we have

$$p(\{S_i\}) = \frac{1}{Z} \exp\Big[-\beta \mathrm{H}(\{S_i\})\Big]\,, \qquad Z = \sum_{\{S_i\}} \exp\Big[-\beta \mathrm{H}(\{S_i\})\Big]\,. \tag{1.22}$$

All of thermodynamics follows from this microscopic starting point, together with the laws of thermodynamics.

The *second law* of thermodynamics states that a system out of equilibrium will eventually evolve to an equilibrium in which the entropy is at least as large as before. If we have two connected systems at energy E_A and E_B respectively, equilibrium will be reached when the total entropy reaches an extremum, i.e. when

$$\left.\frac{\mathrm{d}S_{\text{tot}}}{\mathrm{d}E_A}\right|_{\text{eq.}} = 0\,. \tag{1.23}$$

This leads to the fundamental definition of temperature in terms of entropy,

$$T := \left(\frac{\mathrm{d}S}{\mathrm{d}E}\right)^{-1}\,. \tag{1.24}$$

From here, all the other usual thermodynamical potentials follow by making use of the *first law* of thermodynamics, which essentially expresses that energy is conserved. The change in internal energy of a system is related to the heat absorbtion and the performed work as

$$\mathrm{d}U = \delta Q - \delta W\,. \tag{1.25}$$

In terms of state functions, this gives

$$\mathrm{d}U = T\mathrm{d}S - P\mathrm{d}V + h\mathrm{d}M\,. \tag{1.26}$$

This expression holds true for a closed system. If the external parameters S and V are fixed, the system in thermodynamical equilibrium will be determined by a minimum of the internal energy U.

However, we often deal with open systems, which are coupled to a heat bath which can provide an arbitrary amount of energy to keep the temperature T constant. In this case, the energy flow means that S is no longer constant. The second law now imposes that a minimum is found for

$$F = U - TS - hM\,. \tag{1.27}$$

This quantity F is the Helmholtz free energy, and we can relate it to the partition function by making use of (1.21),

$$S = -k_B \sum_{\{S_i\}} \frac{1}{Z} \exp\Big[-\beta H\Big]\big(-\beta H - \ln Z\big) = k_B\Big[\beta(U - hM) + \ln Z\Big]\,, \tag{1.28}$$

so that we find

(end of lecture 3)

$$F = -k_B T \ln Z\,. \tag{1.29}$$

There are also thermodynamic potentials which are relevant when the pressure is kept fixed (and thus the volume is allowed to change), but these are not very relevant for systems like the Ising model. For completeness, they are all listed in the table below.

Internal energy	$U(S,V,M)$	$\mathrm{d}U = T\mathrm{d}S - P\mathrm{d}V + h\mathrm{d}M$
Helmholtz free energy	$F(T,V,h) = U - TS - hM$	$\mathrm{d}F = -S\mathrm{d}T - P\mathrm{d}V - M\mathrm{d}h$
Gibbs free energy	$G(T,P) = F + PV$	
Enthalpy	$H(S,P) = U + PV$	

In the last column of this table, you can see how the potentials change when the parameters are modified.

▸ ***See also:*** A more extensive explanation of entropy and free energy can be found in [14].

2

Discrete models

As announced, our main interest will be in systems which have a finite number of degrees of freedom, either because they are defined on a lattice or because they are obtained by cutting off high-momentum modes of a continuum theory. In chapter 4 we will be interested in analysing how such systems behave when we send the lattice spacing to zero, or remove the cutoff. For the time being, however, we are interested in analysing the discrete models by themselves.

2.1. Ising models

2.1.1 One-dimensional Ising model

We have already seen the Ising model Hamiltonian,

$$H_{\text{Ising}} = -J\sum_{\langle i,j\rangle} S_i S_j - h\sum_i S_i\,. \tag{2.1}$$

The partition function $Z[\beta, J, h]$ for the one-dimensional model is then given by

$$Z = \sum_{\{S_i=\pm 1\}} \prod_{i=0}^{N-1} \exp\left[\beta\big(JS_iS_{i+1} + \frac{1}{2}h(S_i + S_{i+1})\big)\right]. \tag{2.2}$$

It can be computed by a method known as the *transfer matrix* method. For this purpose, we make the spin chain periodic, i.e. the spin to the right of S_{N-1} is S_0 again. The partition can then be expressed as the trace of a product of 2 × 2 matrices,

$$\begin{aligned} Z = \sum_{\{S_i=\pm 1\}} T_{S_0S_1}T_{S_1S_2}\cdots T_{S_{N-1}S_0}\,,& \\ \text{where}\quad T_{S_iS_j} = \exp\left[\beta\big(JS_iS_j + \frac{1}{2}h(S_i+S_j)\big)\right].& \end{aligned} \tag{2.3}$$

The matrix T is the transfer matrix. This procedure is in spirit very similar to the way in which we derive the path integral from the quantum mechanical transition amplitude: we cut up the amplitude in a number of steps, glue those steps together, and sum over the states at the intermediate points. Here that sum over 'intermediate states' is given by the sum over spins.

In more explicit form, we can write the matrix elements of T as

$$T = \begin{pmatrix} e^{\beta(J-h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J+h)} \end{pmatrix}. \tag{2.4}$$

Since the trace of a matrix is equal to the sum of its eigenvalues, we can express the partition sum as

$$Z = \mathrm{Tr}\left[T^N\right] = \lambda_+^N + \lambda_-^N\,, \tag{2.5}$$

where $\lambda_\pm$ are the eigenvalues of T, with $\lambda_+ > \lambda_-$. It is not hard to compute these eigenvalues explicitly; from (2.4) we get

$$\lambda_\pm = e^{\beta J}\left(\cosh(\beta h) \pm \sqrt{\sinh^2(\beta h) + e^{-4\beta J}}\right)\,. \tag{2.6}$$

When we take the continuum limit $N \to \infty$, the largest of these two eigenvalues will tend to dominate the physics.

We can immediately compute a number of observables, following the general logic outlined in section 1.4. The simplest is the free energy F. In the $N \to \infty$ limit, this one becomes

$$F = -k_B T N \ln \lambda_+ = -JN - k_B T N \ln\left(\cosh(\beta h) + \sqrt{\sinh^2(\beta h) + e^{-4\beta J}}\right)\,. \tag{2.7}$$

The free energy in the continuum limit is determined by the largest eigenvalue of the transfer matrix.

It is an extensive quantity (scaling with the size of the system), but the strict linear scaling with N holds true only in the large-N limit. When $h = 0$ this reduces to

$$F = -k_B T N \ln\left[2\cosh(\beta J)\right]\,. \tag{2.8}$$

This is an analytic function in T except at $T = 0$. Hence there is no phase transition in the one-dimensional Ising model.

The next quantity is the magnetisation. We can compute this in two different ways. Given that we already have the free energy, we can compute it simply by using

$$M = -\left(\frac{\partial F}{\partial h}\right)_T = N\frac{\sinh(\beta h)}{\sqrt{\sinh^2(\beta h) + e^{-4\beta J}}}\,. \tag{2.9}$$

We can alternatively compute it by doing a direct computation of the expectation value of the sum of spins (this method will prove useful later when we compute the correlation length). Since the system is translationally invariant, we can simply compute the expectation value of a single spin and multiply that with N, to get

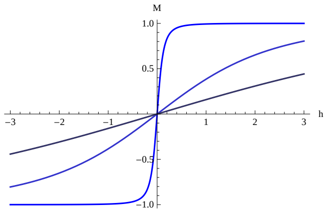
$$\begin{aligned} M = N\langle S_k\rangle &= \frac{N}{Z}\sum_{\{S_i\}} T_{S_0S_1}T_{S_1S_2}\cdots T_{S_{k-1}S_k} \times S_k \times T_{S_kS_{k+1}}\cdots T_{S_{N-1}S_0} \\ &= N\frac{\mathrm{Tr}\left[T^k S T^{N-k}\right]}{\mathrm{Tr}\left[T^N\right]} = N\frac{\mathrm{Tr}\left[S\,T^N\right]}{\mathrm{Tr}\left[T^N\right]}\,, \quad \text{where} \quad S := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\,. \end{aligned} \tag{2.10}$$

Translation invariance is reflected in the fact that this result is independent of k, and related to the cyclic invariance of the trace. In order to compute the trace on the right-hand side of (2.10), we note that T is a symmetric matrix and hence can be diagonalised with an orthogonal matrix P,

$$T = P\Lambda P^{-1}\,, \quad \text{where} \quad \Lambda = \begin{pmatrix} \lambda_- & 0 \\ 0 & \lambda_+ \end{pmatrix}\,. \tag{2.11}$$

If we find P, we can then compute the magnetisation using

$$M = N\frac{\mathrm{Tr}\left[P^{-1}SP\,\Lambda^N\right]}{\mathrm{Tr}\left[\Lambda^N\right]}\,. \tag{2.12}$$



The magnetisation of the one-dimensional Ising model as a function of the external magnetic field. The flatter curves correspond to higher temperatures.

By writing out the right-hand side of (2.11) explicitly and comparing with the components of T on the left-hand side, one finds that

$$P = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}\,, \quad \text{with} \quad \cos(2\phi) = \frac{\sinh(\beta h)}{\sqrt{\sinh^2(\beta h) + e^{-4\beta J}}}\,. \tag{2.13}$$

We then get

$$P^{-1} S P = \begin{pmatrix} -\cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & \cos(2\phi) \end{pmatrix}\,,$$

$$M = N \frac{(\lambda_+^N - \lambda_-^N)\cos(2\phi)}{\lambda_+^N + \lambda_-^N} \overset{N\to\infty}{\longrightarrow} N\cos(2\phi)\,, \tag{2.14}$$

which together with (2.13) reproduces the result obtained in (2.9). Note that for $h = 0$ the magnetisation is always zero, indicating that there is indeed no phase transition in the one-dimensional Ising model. A plot of the magnetisation as a function of the external field, for various values of the temperature, is given in the figure.

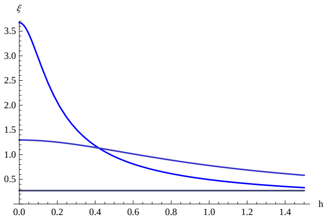
(end of lecture 4)

The correlation length can now be computed using a very similar technique.

$$\begin{aligned} \langle S_0 S_r \rangle &= \frac{1}{Z} \sum_{\{S_i\}} S_0 \times T_{S_0 S_1} T_{S_1 S_2} \cdots T_{S_{r-1} S_r} \times S_r \times T_{S_r S_{r+1}} \cdots T_{S_{N-1} S_0} \\ &= \frac{\text{Tr}\left[S T^r S T^{N-r} \right]}{\text{Tr}\left[T^N \right]} = \frac{\text{Tr}\left[S\,(P\Lambda^r P^{-1}) S (P \Lambda^{N-r} P^{-1}) \right]}{\text{Tr}\left[T^N \right]}\,, \end{aligned} \tag{2.15}$$

Here Λ is again the matrix of eigenvalues as given in (2.11). In the large-N limit the Λ^N matrix can be approximated by keeping only the largest eigenvalue λ_+, but because r is finite, the Λ^r matrix depends on both eigenvalues. A bit of matrix multiplication yields

$$\langle S_0 S_r \rangle = \cos^2(2\phi) + \left(\frac{\lambda_-}{\lambda_+}\right)^r \sin^2(2\phi)\,. \tag{2.16}$$



The correlation length of the one-dimensional Ising model as a function of the external magnetic field. Again, the flatter curves correspond to higher temperatures.

Using our previous result for the magnetisation, the variance becomes

$$\langle S_0 S_r \rangle - \langle S_0 \rangle \langle S_r \rangle = \left(\frac{\lambda_-}{\lambda_+}\right)^r \sin^2(2\phi)\,. \tag{2.17}$$

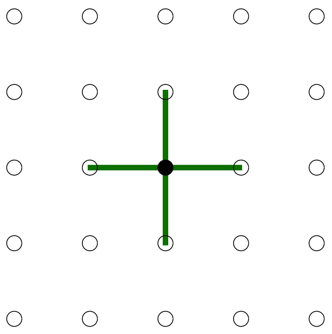
The only r-dependence sits in the ratio of eigenvalues. From the definition of the correlation length (1.19) we then find

$$\xi \propto \frac{1}{\left|\ln\left(\lambda_-/\lambda_+\right)\right|}\,, \qquad \eta = 1\,. \tag{2.18}$$

At zero external magnetic field the correlation length reduces to $\xi = 1/\ln\tanh(\beta J)$. The behaviour as a function of the magnetic field is plotted for various temperatures in the figure.

We state here without proof that for many other models, the correlation length is determined by the same type of expression (2.18), with λ_+ and λ_- the two largest eigenvalues of the transfer matrix. Note that if $\lambda_- = 0$ the correlation length vanishes. This happens when $J = 0$, i.e. when there is no interaction between the spins.

The correlation length is determined in terms of the two largest eigenvalues of the transfer matrix.



The unique configuration which changes 4 bounds (multiplicity N).

A phase transition happens when $\xi \to \infty$, which requires that the two largest eigenvalues become degenerate. For the one-dimensional Ising model this only happens when $h = 0$ and $\beta \to \infty$ ($T \to 0$), but this does not connect two phases.

▸ ***See also:*** The one-dimensional Ising model is discussed in many places, see e.g. [15].

2.1.2 Two-dimensional Ising model

The two-dimensional Ising model can also be solved by the transfer matrix technique, as was shown by Onsager in 1944. However, this requires a much more substantial analysis. Instead, we will in these lectures use a different method to extract properties of the two-dimensional model. This goes under the name of *Kramers-Wannier duality*, and is a relation between the low-temperature perturbative expansion and the high-temperature perturbative expansion. The advantage of doing this is that such expansions and duality relations have analogues in all of the other systems which we will discuss, and hence these techniques are of more general use (in contrast to the transfer matrix method which, although it can sometimes give exact results, is in general too complicated to be of any use). We will in this section only consider the case of a vanishing external magnetic field, i.e. $h = 0$, and for the time being just look at the low- and high-temperature expansions in isolation.

For the *low-temperature* expansion, we start with the ground state at $T = 0$, which has all spins in one direction (here taken to be 'up'). We have to sum over all configurations in the partition sum. For low temperature, the more spins we flip the more bonds will change their energy, and hence we can look at things order by order in the number of flipped spins. If we flip one spin, 4 bonds are changed. Each such bond yields an energy increase of $2\beta J$. If we flip two spins, we can do it in two different ways, depending on how close the spins are. For the first possibility (top figure) we can place this at any of the N lattice sites, but there are two orientations, so there are $2N$ possibilities, each changing 6 bonds. For the second possibility (bottom figure) we have $N(N-5)/2$ possible ways of placing it, and we change 8 bonds. There are three other configurations which change 8 bonds, obtained by flipping more spins. For the partition sum, we obtain the low-temperature expansion,

$$Z = e^{2N\beta J}\left(1 + Ne^{-8\beta J} + 2Ne^{-12\beta J} + \frac{1}{2}N(N+9)e^{-16\beta J} + \cdots\right). \tag{2.19}$$

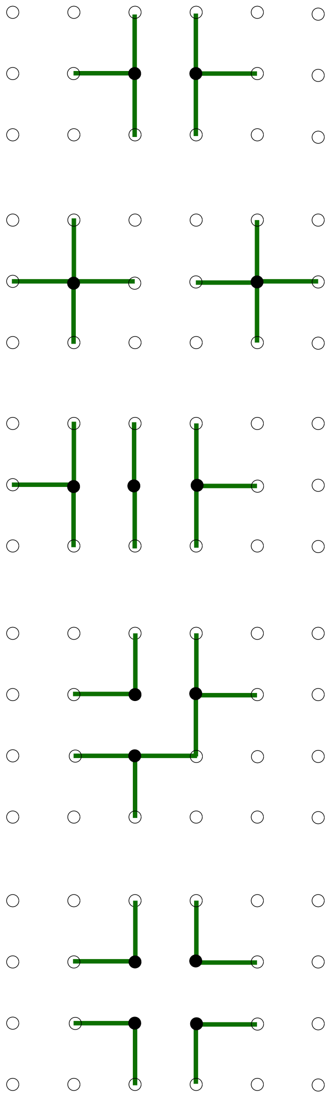
Here $e^{2N\beta J}$ is energy of the all-up state (explained by the fact that there are 2 links for every site, e.g. the one pointing up and to the right). Since T is small, βJ is large and hence $e^{-\beta J}$ is a small number.

Let us now turn to the *high-temperature* expansion. When T is large, βJ is small. A crucial role in finding an expansion in small βJ is played by the following simple identities, valid for a variable S which takes values ± 1 (and easily verified by inserting $S = \pm 1$ explicitly),

$$e^{\beta J S} = \cosh(\beta J) + S\sinh(\beta J)\,, \qquad \sum_{S=\pm 1} S^n = \begin{cases} 0 & \text{if } n \text{ odd} \\ 2 & \text{if } n \text{ even}\,. \end{cases} \tag{2.20}$$

Using these identities we can rewrite the partition sum in the following form,

$$\begin{aligned} Z &= \sum_{\{S_i=\pm 1\}} \prod_{\langle i,j\rangle} \Big(\cosh(\beta J) + S_i S_j \sinh(\beta J)\Big) \\ &= \Big[\cosh(\beta J)\Big]^{2N} \sum_{\{S_i=\pm 1\}} \prod_{\langle i,j\rangle} \Big(1 + S_i S_j \tanh(\beta J)\Big)\,. \end{aligned} \tag{2.21}$$



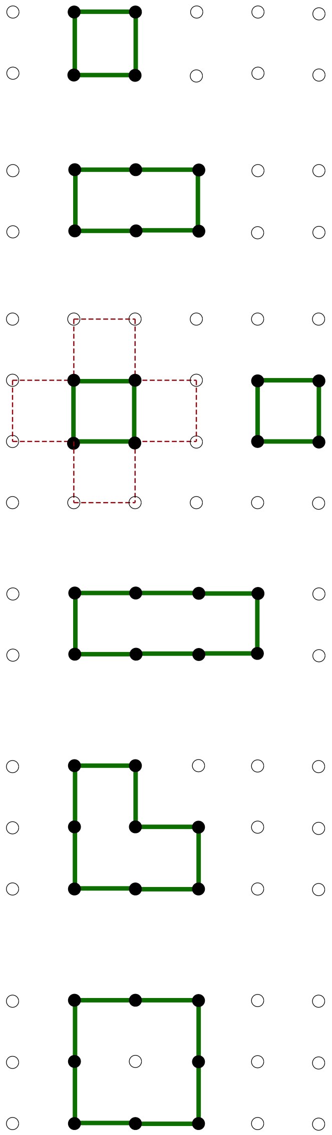
Configurations which change 6 bonds (top figure, multiplicity $2N$) or 8 bonds (other figures, multiplicity $\frac{1}{2}N(N-5)$, $2N$, $4N$ and N respectively).

At this stage we have not done any approximation yet, we have just rewritten the partition sum in a somewhat different form. However, note that $\tanh(\beta J)$ is small at high temperature. We can thus use (2.21) to attempt an expansion in powers of $\tanh(\beta J)$.

(end of lecture 5)

If we expand (2.21) in powers of $\tanh(\beta J)$, the only terms which contribute are those in which all spins which occur do so in even powers (by virtue of the identities in (2.20)). It is convenient to think about such terms in a graphical way, since that will help us to count the number of terms at each order. To do this, first observe that in the expression (2.21), a given product of two spins $S_n S_m$ appears only once. If we denote this product by a line from site n to site m, we need to look for those graphs in which the contours are closed, i.e. do not end in a single spin factor. Multiple closed contours are allowed though. Every line segment contributes a power of $\tanh(\beta J)$, and the total contour length thus sets the order in the perturbative expansion. For the partition sum we find

$$Z = \Big(\cosh\beta J\Big)^{2N} 2^N \Big(1 + N(\tanh\beta J)^4 + 2N(\tanh\beta J)^6 + \frac{1}{2}N(N+9)(\tanh\beta J)^8 + \cdots\Big)\,. \tag{2.22}$$



Various closed contours with non-overlapping edge segments. In the third figure, there are N ways to put the first square and $N-5$ to put the second (since the 5 squares denoted by red dotted lines are disallowed). The multiplicities are N, $2N$, $\frac{1}{2}N(N-5)$, $2N$, $4N$ and N respectively.

The main idea behind the high-temperature expansion we have just seen is that their are only nearest-neighbour couplings in the Hamiltonian. These couplings are the lines in our graphs, and what the high-temperature expansion does is to group lattice variables into linked clusters. For this reason, it is also known as the *cluster expansion* or *hopping expansion*. We will see that it has its use also in much more complicated discrete models, like lattice gauge theory.

2.1.3 Kramers-Wannier duality and dual lattices

Both the low-temperature and high-temperature expansions of the two-dimensional Ising model partition sum are untransparent and complicated. However, we see that these two perturbative series get transformed into each other if we simultaneously replace

$$\begin{aligned} e^{-2\beta J} &\leftrightarrow \tanh(\beta J)\,,\\ \frac{Z_{\text{low}}}{e^{2N\beta J}} &\leftrightarrow \frac{Z_{\text{high}}}{2^N(\cosh\beta J)^{2N}}\,. \end{aligned} \tag{2.23}$$

This is a remarkable property. It says that if we understand the low-temperature, magnetised phase of the model, we will also understand the high-temperature, disordered phase. As we will see shortly, this symmetry is not just a property of the first few terms in the perturbative expansions. In fact, it is possible to start with a partition sum $Z[\beta]$, and rewrite it such that it looks like the partition sum of the Ising model at a different temperature, $Z[-(1/2)\ln\tanh(\beta J)]$. More precisely,

$$Z[\beta] = \Big[\sinh(2\tilde{\beta}J)\Big]^{-N} \times Z[\tilde{\beta}]\,, \tag{2.24}$$

where β and $\tilde{\beta}$ are related by the relation we found from the perturbative analysis,

$$\tilde{\beta}J = -\frac{1}{2}\ln\tanh(\beta J)\,. \tag{2.25}$$

The two-dimensional Ising model is called *self-dual*, in the sense that we can rewrite its partition function as that of another two-dimensional Ising model, but at a different temperature.

The two-dimensional Ising model is self-dual.

(end of lecture 6)

If we assume that the system has only one phase transition, then this had better happen when the inverse temperatures β and $\tilde{\beta}$ are the same. This requires

$$\sinh(2\beta_c J) = 1 \quad \Rightarrow \quad T_c = \frac{2J}{k_B \ln(1+\sqrt{2})}\,. \tag{2.26}$$

Even though we have not yet shown that there is in fact a phase transition, we already know that it has to happen at this particular critical temperature. No exact solution of the model was required.

In this particular case, the result (2.26) for the critical temperature can be verified by making use of Onsager's solution of the two-dimensional Ising model using transfer matrix techniques. Needless to say, it agrees with the explicit computation.

Let us now look closer at how the exact duality (2.24) can be proven. What we want to show is how the Ising model partition sum can be rewritten in terms of another partition sum, but at a different temperature. The starting point is the full expression for the partition sum at vanishing magnetic field. Setting $J = 1$ (as only the combination βJ appears), this is

$$Z = \sum_{\{S_i\}} \exp\Big[\beta \sum_{\langle i,j\rangle} S_i S_j\Big]\,. \tag{2.27}$$

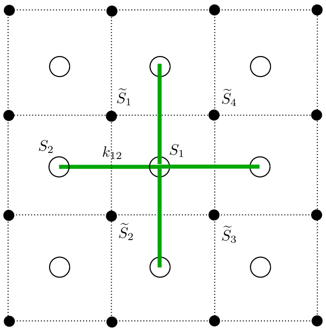
The first step is to rewrite this expression in a way similar to what we did in the first step of the high-temperature treatment, now written in a somewhat more compact form,

$$Z = \sum_{\{S_i\}} \prod_{\langle i,j\rangle} \sum_{k_{ij}=0,1} C_{k_{ij}}(\beta)(S_i S_j)^{k_{ij}}\,,$$
$$\text{where} \quad C_0(\beta) = \cosh\beta\,, \quad C_1(\beta) = \sinh\beta\,. \tag{2.28}$$

The new variables k_{ij} are associated to every nearest-neighbour pair, i.e. to every link on the lattice (they are *link variables*). As we already noted earlier, the only contributions to this sum are those which have an even power of each spin variable S_i. This means that the only non-zero contributions are those in which the distribution of values of the link variables is such that the sum of all links connected to each and every spin S_i is an even number. If it is, the sum over S_i will yield twice the same contribution 'one', so that we can simplify Z as

$$Z = \sum_{\{k_l\}} \prod_l C_{k_l}(\beta) \prod_i 2\,\delta(\text{sum of values of links incident at } i \text{ modulo } 2)\,. \tag{2.29}$$

In other words, we have rewritten the partition sum as a sum over the link variables, but these link variables have to satisfy a particular constraint.



The original Ising lattice (open dots) together with the dual lattice (filled dots). Also indicated are the four links incident on S_1, one of which is k_{12}. The dual variables $\tilde{S}_1$ and $\tilde{S}_2$ are used to define k_{12}.

The trick to solve this constraint is to realise that there is yet another set of variables which we can introduce, such that the link constraint is always satisfied. These are the *dual variables* $\tilde{S}_i$, which sit on the nodes of the dual lattice, as indicated in the figure. The link variables are related according to

$$k_{12} = \frac{1}{2}(1 - \tilde{S}_1 \tilde{S}_2)\,, \tag{2.30}$$

and so on. For the sum of the four link variables incident at S_1, this gives

$$2 - \frac{1}{2}(\tilde{S}_1\tilde{S}_2 + \tilde{S}_2\tilde{S}_3 + \tilde{S}_3\tilde{S}_4 + \tilde{S}_4\tilde{S}_1) = 2 - \frac{1}{2}(\tilde{S}_1 + \tilde{S}_3)(\tilde{S}_2 + \tilde{S}_4)\,. \tag{2.31}$$

This expression is manifestly even for any choice of dual variables.

So we can now write the partition function as

$$Z = \frac{1}{2} 2^N \sum_{\{\widetilde{S}_i\}} \prod_{\langle i,j \rangle} C_{(1-\widetilde{S}_i\widetilde{S}_j)/2}(\beta)\,. \tag{2.32}$$

The extra factor of 1/2 arises because there are two configurations of dual variables which give rise to the same value of the link variable, and thus all configurations of link variables are counted twice. The only thing that is left is to express the $C_k(\beta)$ in terms of the dual variables. This gives

$$\begin{aligned} C_k(\beta) &= \cosh\beta\,\big(1 + k(\tanh\beta - 1)\big) \\ &= \cosh\beta\,\exp\big[k \ln\tanh\beta\big] \\ &= (\cosh\beta\sinh\beta)^{1/2}\,\exp\big[-\tfrac{1}{2}\widetilde{S}_i\widetilde{S}_j \ln\tanh\beta\big]\,. \end{aligned} \tag{2.33}$$

If we insert this into the partition sum we arrive at

$$\begin{aligned} Z = \frac{1}{2}(2\cosh\beta\sinh\beta)^N \sum_{\{\widetilde{S}_i\}} \exp\big[-\tfrac{1}{2}\ln\tanh\beta \sum_{\langle i,j\rangle} \widetilde{S}_i\widetilde{S}_j\big] \\ = \frac{1}{2}(\sinh 2\tilde{\beta})^{-N} \sum_{\{\widetilde{S}_i\}} \exp\big[\tilde{\beta} \sum_{\langle i,j\rangle} \widetilde{S}_i\widetilde{S}_j\big]\,. \end{aligned} \tag{2.34}$$

This is again an Ising model, now in terms of the dual variables, and with a coupling $\tilde{\beta}$ which is related to the original one by the relation we found earlier in the perturbative analysis, namely $\tilde{\beta} = -(1/2)\ln\tanh\beta$.

▸ ***See also:*** A good text on duality in statistical systems and field theory is Savit's report [16]. The Ising model is discussed in e.g. [3].

2.2. Kosterlitz-Thouless model

The next model we will discuss is the Kosterlitz-Thouless model. It is again a lattice model, but instead of the spin variables S_i of the Ising model, which can only take on the values ±1, it has a two-dimensional unit-vector $\vec{S}$ at each site. That is to say, there is one continuous angle variable at each site, with

$$\vec{S}_i = \big(\cos\theta_i,\, \sin\theta_i\big)\,. \tag{2.35}$$

It is thus the same as the classical Heisenberg model in two dimensions. The Hamiltonian is again given by a nearest-neighbour one,

$$H_{\text{KT}} = -J \sum_{\langle i,j\rangle} \vec{S}_i \cdot \vec{S}_j = -J \sum_{\langle i,j\rangle} \cos\big(\theta_i - \theta_j\big) \tag{2.36}$$

The crucial ingredients which set this model apart from the Ising model is that we are dealing with a continuous variable, and that this continous variable is periodic, i.e. takes values in a compact space. These properties will come back again when we discuss lattice gauge theories for QCD, so the Kosterlitz-Thouless model is a nice warm-up example.

One might perhaps have expected that, just like for the Ising model, there should be a high-temperature phase in which the spins are disordered and there is no magnetisation, and a low-temperature phase in which the spins align and there is a net magnetisation (the expectation value of the sum of spins does not vanish). However, that intuition turns out to be wrong. What we will see is that the Kosterlitz-Thouless model never has a net magnetisation, but it does exhibit two different phases, characterised by two different behaviours of the correlation function $\langle \vec{S}_i \vec{S}_j \rangle$.

(end of lecture 7)

At high temperature, we can make an expansion which is very similar to the one in the Ising model. Write the correlation function as

$$\langle e^{i\theta_0} e^{-i\theta_n} \rangle = \frac{1}{Z} \int \prod_m \mathrm{d}\theta_m \, e^{i\theta_0 - i\theta_n} \times \exp\left[\frac{1}{2} \beta J \sum_{\langle i,j \rangle} \left(e^{i\theta_i - i\theta_j} + e^{i\theta_j - i\theta_i} \right) \right] . \tag{2.37}$$

Here we have expanded the cosine in terms of exponentials. For large temperatures, we can expand in powers of β and obtain integrals over exponentials. Using identities which take the role of those in (2.20),

$$\int_0^{2\pi} \mathrm{d}\theta_m = 2\pi \,, \quad \int_0^{2\pi} \mathrm{d}\theta_m e^{i\theta_m} = 0 \,, \tag{2.38}$$

we then see (again by similar logic as in the Ising model at high temperature) that the only terms which will contribute to the correlator are those in which we have exponentials corresponding to a contour which starts at site 0 and ends at site n. Any other contour will have a left-over $e^{i\theta_k}$ for some site, and integrate to zero. Assuming that the sites 0 and n are on an axis, the lowest order contribution comes from a contour that goes straight between these sites, and is of order $(\beta J)^n$. The correlator thus behaves as

$$\langle e^{i\theta_0} e^{-i\theta_n} \rangle \sim \exp\Big[- |n| \ln\big(1/(\beta J)\big) \Big] . \tag{2.39}$$

The Kosterlitz-Thouless model at high temperature has a finite correlation length, which goes to zero as $T \to \infty$.

The correlation length is thus $\xi \sim 1/\ln\big(1/(\beta J)\big)$ at high temperature.

In the low-temperature phase we expect the phases θ_n to vary slowly as a function of the position on the lattice. We can then expand the cosine in the Hamiltonian, and obtain for the correlator

$$\langle e^{i\theta_0} e^{-i\theta_n} \rangle = \frac{1}{Z} \int \prod_m \mathrm{d}\theta_m \, e^{i\theta_0 - i\theta_n} \exp\left[\beta J - \frac{1}{2} \beta J \sum_{\langle i,j \rangle} (\theta_i - \theta_j)^2 \right] . \tag{2.40}$$

In preparation for the jump to continuum field theories, the argument of the Boltzmann weight is often interpreted as a discretised derivative, i.e.

$$\frac{1}{2} \beta J \sum_{\langle i,j \rangle} (\theta_i - \theta_j)^2 = \frac{1}{2} \beta J \sum_{i,\mu} (\nabla_\mu \theta_i)^2 \,, \tag{2.41}$$

where μ labels the direction away from site i, and (in a somewhat sloppy notation),

$$\nabla_\mu \theta_i := \theta_{i+\mu} - \theta_i \,. \tag{2.42}$$

The integral in (2.40) is purely Gaussian, so can be performed exactly. Using the continuum language (2.42), we find[1]

$$\langle e^{i\theta_0} e^{-i\theta_n} \rangle \sim \left| \frac{1}{n} \right|^{\frac{1}{2\pi\beta J}} . \tag{2.44}$$

The Kosterlitz-Thouless model at low temperature has an infinite correlation length, but no spontaneous symmetry breaking.

Since there is no exponential suppression, the correlation length is infinite. How-

---

[1] This is similar to the manipulation with path integrals, which states that

$$\int \mathcal{D}\phi \, \exp\left[\int \mathrm{d}x \left(\frac{1}{2} (\partial_\mu \phi(x))^2 + J(x)\phi(x) \right) \right] \propto \exp\left[\frac{1}{2} \int \mathrm{d}x \int \mathrm{d}y \, J(x) \Delta(x-y) J(y) \right] , \tag{2.43}$$

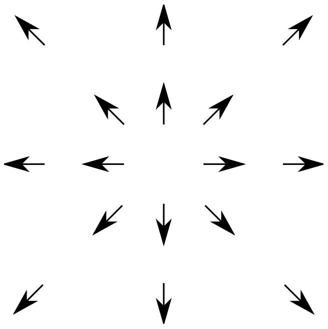
with $\Delta(x-y)$ the inverse of the kinetic operator. In the case here, $J(x) = \delta(x) - \delta(n-x)$, and the propagator in two dimensions behaves as $\Delta(n) \sim -\frac{1}{2\pi} \ln|n|$.

ever, the correlation function does fall off to zero for large distance. Because we have $\langle e^{i\theta_0} e^{-i\theta_n}\rangle \sim \langle e^{i\theta_0}\rangle\langle e^{-i\theta_n}\rangle$ for large distances, this immediately also implies that there is no magnetisation, i.e. no spontaneous symmetry breaking. This is in agreement with the *Coleman-Mermin-Wagner theorem* which says that a continuous global symmetry in two dimensions cannot be spontaneously broken in a system with local interactions.

The question now remains: what is the crucial qualitative difference between the low-temperature and the high-temperature regime? The key ingredient is the periodicity of the θ_i variables. If we consider the continuum formulation of the model (using (2.41)), the solutions to the equations of motion are solutions to the two-dimensional Laplacian, with an additional periodicity condition,

$$\left(\partial_x^2 + \partial_y^2\right)\theta = 0\,, \qquad 0 < \theta < 2\pi\,. \tag{2.45}$$

The solutions to this equation can be labelled by the number of times the phase θ goes around when we circle the origin. These vortices are configurations which have a lower free energy for sufficiently high temperatures. They 'condense' above this T_c, and lead to a finite correlation length.



A vortex of the Kosterlitz-Thouless model, arrows denoting the unit-vectors $\vec{S}_i$.

It is not so easy to show that these vortices indeed dominate the partition sum above T_c. However, it can be made plausible by estimating the free energy of a vortex, and showing that condensation of a vortex lowers the free energy above a certain temperature. The free energy is

$$F = E - TS\,, \tag{2.46}$$

where S is the entropy. The energy can be determined from

$$E = \frac{1}{2}\int \mathrm{d}^2 x\,(\partial_\mu\theta)^2 = \int \mathrm{d}\phi \mathrm{d}r\, r g^{\phi\phi} = \pi J \ln(L/a)\,, \tag{2.47}$$

where we approximated $\theta = \phi$, and have written the answer in terms of the total size of the lattice L and the lattice spacing a. Now the entropy is simply given by k_B times the logarithm of the number of ways in which we can put down the vortex, which is the number of lattice sites $(L/a)^2$. So we find

$$F = \big(\pi J - 2k_B T\big)\ln(L/a)\,. \tag{2.48}$$

Above $T_c = \pi J/(2k_B)$ the system prefers to condense a vortex. This is one of the classic examples of a system in which "entropy wins over energy". Working out the precise details of the condensation mechanism goes beyond the scope of these lectures.

(end of lecture 8)

▸ ***See also:*** Section 2.5 of [4] for a deeper analysis of vortex condensation in the Kosterlitz-Thouless model.

3

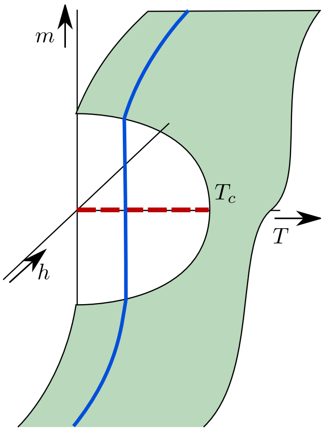
Effective descriptions

In the previous chapter we have seen a number of explicit examples of systems with a phase transition. We have, however, not discussed much of the details of phase transitions themselves, or the techniques to compute properties of these systems close to the phase transition. Exact computations are often hard (as witnessed by e.g. the two-dimensional Onsager solution of the Ising model). In the present chapter we will introduce two ideas for approximate methods to capture the essentials of systems with phase transitions. Before we do that, we will however need a little bit more terminology.

3.1. Characterising phase transitions

A phase transition occurs when the external parameters of a system are such that one or more obervables are singular (i.e. discontinuous or divergent). These points in the phase diagram are called *critical points*. The *critical exponents* show how various quantities vanish or diverge near the critical points. An important property of phase transitions is that various systems which naively look very different may still, when they approach the critical point, have identical behavior. More precisely, many different systems turn out to have the same critical exponents. Examples of such systems are the Ising model on a square and triangular lattice. While the value of the critical temperature depends on the details of the interaction, these two models do turn out to have the same critical exponents.

Phase transitions are commonly classified into roughly three different groups. For *first order* transitions,



The phase diagram of the two-dimensional Ising model. For $T < T_c$ the transition between $h < 0$ and $h > 0$, along the blue curve, is first-order (the magnetisation jumps discontinuously). The line of first-order transitions in the $h - T$ plane (red dashed line) ends in the critical point $T = T_c$.

1. The singular behaviour is manifest in the form of discontinuities.
2. The correlation length ξ stays finite.
3. The latent heat can be non-zero, i.e. going from one phase to the other can release or absorb energy.
4. At the transition two or more phases may coexist in equilibrium (for instance, liquid-ice in water, or domains of magnetisation in ferromagnets).
5. The symmetries of the system in the phases on either side of the transition are often unrelated.

A first order transition occurs for example in the Ising model below the critical temperature (see the figure), when we change the external magnetic field from negative to positive. Typically, the line of first-order transitions ends in a critical point. If we follow a path at zero magnetic field, passing through this point leads to a *second-order* transition.

Second-order transitions are transitions in which

1. The singular behaviour is visible in the form of divergences.
2. The correlation length ξ diverges.
3. The latent heat is zero.
4. There are no mixed phases at the transition point.
5. The symmetries of the system in the phases on either side of the transition are related (usually the symmetries for $T < T_c$ form a subgroup of those at $T > T_c$).

Using again the Ising model as an example, and focussing on $h = 0$, we see that the spontaneous magnetisation for $T < T_c$ breaks the symmetry of the phase at $T > T_c$.

Finally, there exist a class of *infinite order* transitions, in which no symmetries get spontaneously broken. We have actually seen such a transition when we discussed the Kosterlitz-Thouless model.[1]

For the purposes of this course, the second-order transitions are most relevant, as the correlation length diverges and the system loses knowledge about the microscopic details. In the language of discretised quantum field theory, adjusting the parameters such that the correlation length diverges is also known as *taking the continuum limit*. The way in which the correlation length diverges is commonly parameterised by the *critical exponent* ν, defined as

Critical exponents determine the behaviour of observables near a second-order phase transition.

$$\xi \sim \left|T - T_c\right|^{-\nu} . \tag{3.1}$$

In fact, the two-point correlator yields more information than just the correlation length; recalling the definition (1.19),

$$\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle \sim \frac{1}{r^{D+\eta-2}} \exp\left[-\frac{r}{\xi}\right] . \tag{3.2}$$

we also have the exponent η, which sets the power-law behaviour of the two-point correlator. The η exponent is called the *anomalous dimension*. Typically, many other correlation functions will diverge too. The first of those is the magnetisation, which is the *order parameter* of the phase transition (it can be used to determine whether we are in one phase or another). We thus have a number of other standard critical exponents,

Phase transitions are labelled by the value of one or more order parameters.

$$\begin{aligned}
\text{specific heat}: &\quad C \sim \left|T - T_c\right|^{-\alpha} &&(\text{at } h = 0)\,,\\
\text{order parameter}: &\quad M \sim \left|T - T_c\right|^{\beta} &&(\text{at } h = 0)\,,\\
\text{susceptibility}: &\quad \chi \sim \left|T - T_c\right|^{-\gamma} &&(\text{at } h = 0)\,,\\
\text{equation of state}: &\quad M \sim h^{-1/\delta} &&(\text{at } T = T_c)\,.
\end{aligned} \tag{3.3}$$

The specific heat is defined as $C = (\partial^2 F/\partial T^2)_V$ and the susceptibility as $\chi = (\partial M/\partial h)_T$. For systems other than the Ising model, the role of M and h can of course be played by some other parameters. We will see later that not all of these critical exponents are independent; they are related by so-called scaling relations.

The critical exponents $\alpha, \beta, \gamma, \delta, \nu$ and the anomalous dimension η label a universality class.

---

[1] The naming convention of phase transitions is historical and is vaguely connected to the order in derivatives of the thermodynamic potentials at which a discontinuity is present. This historical detail is best forgotten.

Experimentally, it is observed that many quite different systems nevertheless have the same critical exponents. This goes under the name of *universality*. From the point of view of the microscopic theory, the existence of such *universality classes* can be quite suprising. Mean field theory and the Landau-Ginzburg approximation both explain to some extent why universality exists, and how we can make use of it to compute properties of systems near a phase transition.

3.2. Mean field theory

Mean field theory and Landau-Ginzburg models are phenomenological theories which were developed to describe phase transitions. Both give the same predictions for the critical exponents. Both ignore local fluctuations of spins (or other variables) about their mean value. Hence, neither of them give correct critical exponents in low dimensions, where these fluctuations can be very wild. As the dimension is increased, the predictions of these theories is improved, and they essentially predict the correct critical exponents at an ordinary critical point if $D > 4$.

(end of lecture 9)

In the mean-field approach, we are still dealing with the microscopic degrees of freedom, but instead of coupling them to their neighbours, we couple them to the average 'field' produced by all other degrees of freedom. Let us illustrate this approach at the level of the Ising model. Recall the Hamiltonian (1.14),

$$H_{\text{Ising}} = -J \sum_{\langle i,j \rangle} S_i S_j - h \sum_i S_i \,. \tag{3.4}$$

The idea is now to write $S_i = m + \delta S_i$, i.e. write the spin as a difference with respect to the average spin. We use this to write

$$S_i S_j = -m^2 + m(S_i + S_j) + \delta S_i \delta S_j \,. \tag{3.5}$$

The idea will now be to neglect $\delta S_i \delta S_j$ in all computations (as well as higher order terms in the variation).[2] With this assumption, the Hamiltonian becomes

$$\begin{aligned} H_{\text{MFT}} &= -J \sum_{\langle i,j \rangle} \Big(-m^2 + m(S_i + S_j) \Big) - h \sum_i S_i \\ &= JNDm^2 - (h + 2JDm) \sum_i S_i \,, \end{aligned} \tag{3.6}$$

where we have made use of the fact that in D dimensions, there are D links for each site. The partition sum now becomes

$$\begin{aligned} Z_{\text{MFT}} &= \exp\left[-\beta JNDm^2\right] \sum_{\{S_i\}} \exp\left[\beta(h + 2JDm) \sum_i S_i\right] \\ &= \exp\left[-\beta JNDm^2\right] \Big(2\cosh\big(\beta(h + 2JDm)\big)\Big)^N \,, \end{aligned} \tag{3.7}$$

where we have used once more the identities (2.20). The free energy becomes

$$F_{\text{MFT}} = N\Big[JDm^2 - k_B T \ln\Big(2\cosh\big(\beta(h + 2JDm)\big)\Big)\Big] \,. \tag{3.8}$$

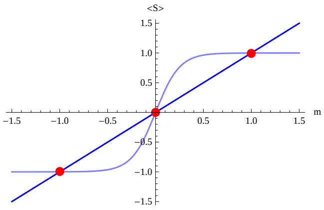
---

[2] As the average spin m can be between plus and minus one, the difference δS_i can be of order one depending on the site. However, the idea is that deviations for which this value is large are rare, and for most spins δS_i will in fact be small.

The magnetisation follows, as in section (2.1.1), either by computing the expectation value of the spin directly,

$$m = \langle S_k \rangle = \frac{\sum\limits_{\{S_i\}} S_k \exp\left[\beta(h+2JDm)\sum\limits_i S_i\right]}{\sum\limits_{\{S_i\}} \exp\left[\beta(h+2JDm)\sum\limits_i S_i\right]} = \tanh\Big(\beta(h+2JDm)\Big)\,, \tag{3.9}$$

or by taking the h-derivative of the partition sum. The value for the magnetisation can be obtained graphically from this equation, see the figure. Plotting the magnetisation as a function of T then gives us the mean-field approximation to the phase diagram of the Ising model.

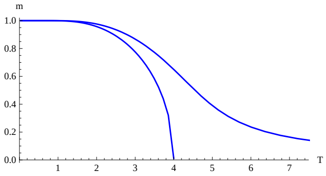


Solving for the magnetisation of the Ising model in the mean-field approximation. At $h = 0$, there are three solutions, $m = 0$ and $m = \pm m_0$.

The critical exponents can be obtained directly. Firstly, the critical point (that temperature above which (3.9) only has the trivial solution $m = 0$) is at $T = 2DJ/k_B$. When $h = 0$ we can expand m near the critical point, to obtain

$$m \approx m(1-t) - \frac{1}{3}m^3(1-t)^3\,, \tag{3.10}$$

where $t = (T - T_c)/T_c$ (so that $\beta \approx \frac{1}{2DJ}(1-t)$. This leads to $m \sim (-t)^{1/2}$ (of course valid only for $t < 0$), so that the critical exponent $\beta = 1/2$. Similar logic leads to the other critical exponents, $\gamma = 1$ and $\delta = -3$.



Magnetisation of the two-dimensional Ising model in the mean-field approximation, for $h = 0$ (positive branch) and $h > 0$, as a function of temperature.

The mean-field approximation gets several things wrong, in particular the fact that it predicts a second-order phase transition for any dimension (whereas we have seen that no such transition exists for the one-dimensional Ising model).

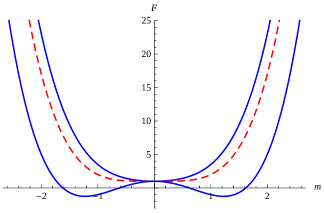
3.3. Landau-Ginzburg

The Landau-Ginzburg approximation which assumes that the only relevant degrees of freedom near a phase transition are those which are associated to an order parameter. Concretely, for the Ising model, this means that it attempts to describe the model purely in terms of the average magnetisation $m(x)$ of the spins. In its crudest form, it assumes that the average magnetisation is the same for all spins, in which case it obtains results which are like those of the mean-field approximation. In contrast to that approximation, however, the Landau-Ginzburg approach does not explicitly deal with the microscopic degrees of freedom. Rather, the goal is to write down the most general expression for the free energy as a function of the order parameter, such that it is compatible with the symmetries of the underlying discrete model. The idea is that the order parameter is in some sense the 'slowest' degree of freedom of the system.

(end of lecture 10)

For the Ising model, the underlying system at $h = 0$ has a symmetry $S_i \leftrightarrow -S_i$, so the Landau-Ginzburg energy has to be symmetric under a sign flip of $m(x)$. We write

$$F = F_0 + F_2 m^2 + F_4 m^4 + \ldots\,. \tag{3.11}$$



Behaviour of the Landau-Ginzburg free energy for positive, zero and negative value of F_2. For negative values, two minima appear, which break the symmetry.

The value $F_2 = 0$ corresponds to the critical temperature. For $F_2 > 0$ only the $m = 0$ minimum exists, but for $F_2 < 0$ we get spontaneous symmetry breaking and two equivalent minima at $m \neq 0$ develop. Note that this happens despite the fact that the free energy is completely smooth. In this broken phase, we can write $F_2 = t\tilde{F}_2$ with t the reduced temperature $t = (T - T_c)/T_c$. The other coefficients stay finite when $t \to 0$. For $t < 0$, the minimum of the free energy is given by

$$\frac{\partial F}{\partial m} = 0 = 2\tilde{F}_2 t m + 4F_4 m^3\,, \tag{3.12}$$

from which we have $m \sim (-t)^{1/2}$. This recovers the mean-field exponent. Of course, this had to happen, because the mean-field expression for the free energy, expanded for small magnetisation, reads at $h = 0$

$$F_{\text{MFT}} = F_0 + DJN(1 - 2DJ\beta)m^2 + \mathcal{O}(m^4)\,. \tag{3.13}$$

Here we also see again that the critical temperature corresponds to the zero of F_2.

We can easily extend this model to a non-vanishing external magnetic field. In this case, the symmetry $S_i \leftrightarrow -S_i$ of the underlying microscopic model is broken, and so we should allow for terms in the free energy which are odd powers of m. The simplest one is of course the hm interaction, so that

$$F = F_0 + hm + F_2 m^2 + F_4 m^4 + \ldots \tag{3.14}$$

There are good reasons to exclude e.g. the m^3 term. Assuming again $F_4 > 0$, we find that for any $h \neq 0$ there is only one global minimum. By changing slowly and continously from $h > 0$ to $h < 0$, the system undergoes a process by which it is temporarily overheated, i.e. sitting in a local minimum with a higher energy than the one in the global minimum. Near the critical point (where $F_2 = 0$) the equation which determines the minimum of the free energy yields $m \sim h^{1/3}$. So we see that the critical exponent $\delta = -3$. This is the same as in mean-field theory.

If $F_4 < 0$, we need to include more terms in the free energy to prevent it from being unbounded from below. Let us briefly discuss the system determined by the Landau-Ginzburg free energy

$$F = F_0 + F_2 m^2 - |F_4| m^4 + F_6 m^6\,, \tag{3.15}$$

where $F_6 > 0$ in order to keep the free energy from running off to minus infinity. The minima are determined by

$$\frac{\partial F}{\partial m} = 0 = m(2F_2 - 4|F_4|m^2 + 6F_6 m^4)\,, \tag{3.16}$$

which yields

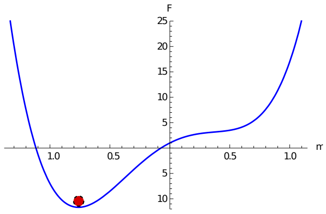
$$m = 0 \qquad \text{or} \qquad m^2 = \frac{|F_4| \pm \sqrt{|F_4|^2 - 3F_6F_2}}{3F_6}\,. \tag{3.17}$$

When $F_2 > |F_4|^2/(3F_6)$ only the $m = 0$ minimum exists. However, as long as $|F_4|$ is not zero, we can always make two minima appear by moving the temperature sufficiently close to T_c (i.e. making F_2 positive but sufficiently small). By tuning F_2, we can make these two extra minima appear at $F = F_0$, hence creating three degenerate ground states.

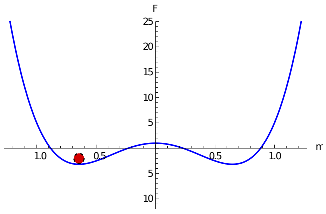
There are thus two temperature scales that play a role. For $T > T_0$ (where T_0 is the temperature at which the three minima become degenerate), there is only one global minimum at $m = 0$. For $T_c < T < T_0$ there are three minima, one of which sits at $m = 0$, but the other two have lower free energy. Finally, for $T < T_c$ the minimum at $m = 0$ disappears (turns into a maximum).

In the $F_4 \to 0$ limit, the two local minima that are typically present just below the critical temperature disappear. The special point at $F_2 = F_4 = 0$ is the *tricritical* point. At this point, the equation of motion yields $m \sim h^{1/5}$ so that $\delta = -5$. This is quite different from the mean-field behaviour at the ordinary critical point. This behaviour is found in e.g. a mixture of ^{3}He and ^{4}He.

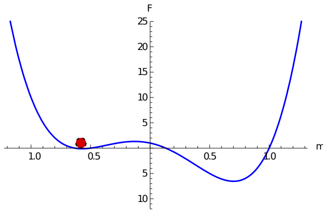
Finally, the assumption that m is a constant is only the first step as far as the Landau-Ginzburg approach is concerned. It makes perfect sense to assume that in fact $m = m(x)$ and include terms in the free energy which contain gradients of the average magnetisation.



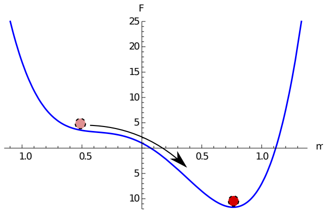
Free energy in the $F_4 > 0$ case, for large positive h



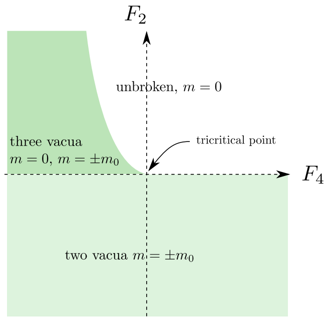
Free energy in the $F_4 > 0$ case for $h = 0$



Free energy for small negative h; the system is 'overheated' and still sits in a local rather than global minimum.



Free energy for large negative h; the system has now rolled to the global minimum.



Phase diagram of the model with $F_6 = 1$, showing the tricritical point. At $F_4 = 0$, the system sits on the vertical axis, with the unbroken phase at $T < T_c$ and the broken phase at $T > T_c$, as before.

(end of lecture 11)

4

Universality and renormalisation

We have seen that if we restrict ourselves to large distance physics, the idea of Landau-Ginzburg is that we can restrict ourselves to the dynamics of the slow, long-wavelength modes, and this then essentially forces us to conclude that many systems which are quite different at microscopic scale will behave in the same way when we get close to a phase transition. This phenomenon is called *universality*. The only things which really seem to matter are the symmetries of the underlying model, and possibly the dimension of the system.

While the Landau-Ginzburg approach captures this idea nicely in the form of a mathematical model, it would be much nicer if we could *derive* this universality property from the microscopic models. The techniques used to do this go under the name of the *renormalisation group*. They allow us to start from a microscopic model and, in successive steps, integrate out microscopic variables to end up with a model which describes the physics at a somewhat larger distance scale. Following this many times, we can then see which microscopic interactions determine the behaviour at large distances, and which ones are irrelevant for the purpose of describing a phase transition.

4.1. Kadanoff block scaling

Before we discuss more general aspects, let us first illustrate the renormalisation group idea at the level of the one-dimensional Ising model. We start with a lattice with spacing a. The idea is now that we are only interested in what happens at larger distances, e.g. $\geq 2a$. We can then try to construct a new *effective* model with lattice spacing $2a$, in such a way that we reproduce all physics at these larger distances from the simpler model with half the number of degrees of freedom.

The Ising partition sum is

$$Z = \sum_{\{S_i\}} \prod_{i=1}^{N} \exp\Big[\beta J S_i S_{i+1} + \beta h S_i + \beta F_0\Big] = \sum_{\{S_i\}} T_{S_1 S_2} T_{S_2 S_3} \cdots T_{S_N S_1} = \text{Tr}\, T^N\,, \quad (4.1)$$

with the transfer matrix

$$T = C \begin{pmatrix} AB & A^{-1} \\ A^{-1} & AB^{-1} \end{pmatrix} \quad A = e^{\beta J}\,,\ B = e^{\beta h}\,,\ C = e^{\beta F_0}\,. \quad (4.2)$$

We have added a constant F_0 to the energy, as we will shortly see that a renormalisation group analysis requires us to keep track of it. The idea is now to do a so-called block-spin transformation, in which we group the dynamics of two spins together, i.e. integrate over every second spin. We can do that by grouping the transfer matrices in groups of two, and define

$$\tilde{T}(\tilde{A},\tilde{B},\tilde{C}) = T^2(A,B,C)\,. \tag{4.3}$$

This gives rise to

$$\begin{aligned}\tilde{C}\tilde{B}\tilde{A} &= C^2(A^2B^2 + A^{-2})\,,\\ \tilde{C}\tilde{A}^{-1} &= C^2(B + B^{-1})\,,\\ \tilde{C}\tilde{A}\tilde{B}^{-1} &= C^2(A^2B^{-2} + A^{-2})\,.\end{aligned} \tag{4.4}$$

These are *discrete renormalisation group equations*: they relate the parameters (coupling constants) of the original model to those of the new model. The number of degrees of freedom has been reduced by a factor of two, but the long-range behaviour of these two models is the same.

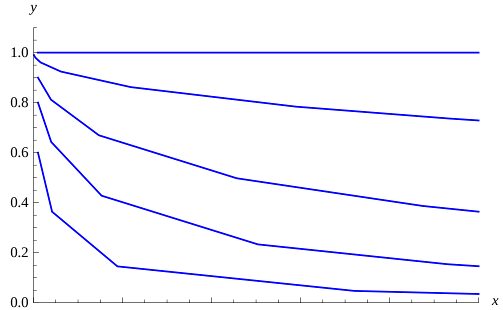
Isolating the parameters of the coarse grained model, and introducing the following shorthands,

$$x = e^{-4\beta J}\,,\quad y = e^{-2\beta h}\,,\quad z = C^{-4}\,, \tag{4.5}$$

the system of equations becomes

$$\begin{aligned}\tilde{x} &= \frac{x(1+y)^2}{(x+y)(1+xy)}\,,\\ \tilde{y} &= \frac{y(x+y)}{1+xy}\,,\\ \tilde{z} &= \frac{z^2xy^2(1+y)^2}{(x+y)(1+xy)}\,.\end{aligned} \tag{4.6}$$

Starting from some initial point (x,y,z), we can now follow the images of this point under successive renormalisation group transformations. Restricting ourselves to the $x-y$ plane, we get the following plot,



Under n blocking transformations, the correlation length (expressed in terms of the number of lattice sites) should scale as

$$\tilde{\xi} = 2^{-n}\xi\,. \tag{4.7}$$

However, at a fixed point of the renormalisation group flow, the parameters do not change, and hence the system will have the same correlation length before and after the transformations,

$$\text{fixed point}:\quad \tilde{\xi} = \xi\,. \tag{4.8}$$

This can only be true if $\xi = 0$ (a trivial fixed point) or $\xi = \infty$ (a non-trivial fixed point).

(end of lecture 12)

We have here seen the simplest way in which we can implement the renormalisation group. The process of summing over every other spin is called *decimation*. It has the property that the remaining variables, i.e. the variables in which the new system is expressed, are again Ising spins. The one-dimensional case is even more special in the sense that the new system is not just expressed in terms of Ising spins, but does not have anything more than the nearest-neighbour interactions that we started from.

In general, things are not so simple. In two dimensions, for example, you might be tempted to sum over all spins on odd lattice sites. While this leaves us with a system of Ising spins (the even lattice spins), it turns out to introduce new couplings, e.g. next-to-nearest neighbour ones.

▸ ***See also:*** This standard example appears e.g. in [3].

4.2. Renormalisation group flow

4.2.1 Scaling near the critical point

Let us now look at the renormalisation group in a somewhat more formal setting. The renormalisation group equations in general map a set of parameters (couplings) to another one; we will denote the set of parameters by K and its individual members by K^i,

$$K^i = \{J, h, \ldots\}^i \quad \overset{\text{ren.group}}{\longrightarrow} \quad \tilde{K}^i = \{\tilde{J}, \tilde{h}, \ldots\}^i = R^i(K)\,. \tag{4.9}$$

A fixed point K^i_* of the equations is one where $\tilde{K}^i = K^i = K^i_*$. In general, the renormalisation group transformation is highly non-linear. Close to a fixed point, however, we can expand the transformation in powers of the small deviation V^i from the fixed point, or more formally,

$$\tilde{K}^i = \tilde{K}^i_* + \tilde{V}^i = K^i_* + \sum_j \frac{\partial R^i(K_*)}{\partial K^j} V^j\,, \quad \text{or} \quad \tilde{V}^i = \sum_j R^i{}_j(K_*)\, V^j\,, \tag{4.10}$$

which implicitly defines the matrix components $R^i{}_j$. The problem becomes linear (because R is evaluated at the fixed point).

The eigenvalues λ_j and eigenvectors $\vec{e}_j$ of R play an important role.[1] We will write the eigenvalues as

$$\lambda_j \equiv b^{y_j}\,, \tag{4.11}$$

where b is the scale factor associated to the renormalisation group transformation (in the Ising case we discussed, $b = 2$). We can now expand a point K^i close to the fixed point in terms of the eigenvectors,

$$K^i = K^i_* + \sum_j g_j\, e_j{}^i\,. \tag{4.12}$$

The expansion coefficients g_j are called the *scaling fields*. Applying one step of the renormalisation group transformation gives

$$\tilde{g}_j = b^{y_j} g_j\,. \tag{4.13}$$

---

[1] In general there is of course no guarantee that R is symmetric, so that the eigenvalues do not have to be real and the eigenvectors do not have to span a complete basis. We will assume, however, that they do.

This shows that there are particular combinations of the parameters of the theory which scale with a simple factor under the renormalisation group. The exponents y_j are directly related to the critical exponents we have seen in (3.3); we will see this illustrated explicitly on the Ising model shortly.[2]

It is important to emphasise once more that the partition function does not change under the renormalisation group flow. We simply change from a partition function expressed in terms of N spins to one of $\tilde{N}$ spins, with modified couplings,

$$\mathrm{Tr}_{\{S_i\}}\, e^{-\beta N f(\{K\})} = \mathrm{Tr}_{\{\tilde{S}_i\}}\, e^{-\beta \tilde{N} f(\{\tilde{K}\})}\,. \tag{4.15}$$

Because $\tilde{N} = b^{-D} N$, we have that the free energy *per spin* has to satisfy

$$f(\{K\}) = b^{-D} f(\tilde{K})\,. \tag{4.16}$$

The free energy near a critical point is a generalised homogeneous function, with exponents related to the eigenvectors and -values of the renormalisation group matrix at that point.

If we express the couplings using the scaling fields g_j instead of the components K^i, and then use (4.13), we get

$$f(g_1, g_2, \ldots) = b^{-D} f(b^{y_1} g_1, b^{y_2} g_2, \ldots)\,. \tag{4.17}$$

Functions satisfying this property are called *generalised homogeneous*. We see that the free energy near a critical point is constrained by the eigenvectors and -values of the renormalisation group transformation matrix.

4.2.2 Critical surface and universality

We have seen that a good way to parameterise the system near a critical point is in terms of the expansion coefficients g_j, defined in (4.12). These linear combinations of the parameters in K scale in a nice way under scale transformations, as we have seen in (4.13). We can distinguish three different cases:

$y_i > 0$

The operators multiplying these coefficients are called *relevant*. As we do more and transformations, the coefficients grow larger and larger, driving the system away from the critical point.

$y_i < 0$

These operators are called *irrelevant*. If you give them a small non-zero coefficient, a renormalisation group transformation will drive these coefficients back to zero, i.e. back to the critical point.

$y = 0$

These are *marginal*. In order to understand what happens, you have to go beyond the linear approximation.

In the vicinity of a critical point, the irrelevant directions span a subspace of points which are attracted to the critical point. By continuity, this subspace will exist also in some larger neighbourhood of K_*. We call this subspace the *critical surface*. On the critical surface, the correlation length is infinite. This follows because the

---

[2] For the one-dimensional Ising model, we find from (4.6) that the matrix R at the critical point $(x, y) = (0, 1)$ is

$$R_*(0,1) = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}\,. \tag{4.14}$$

The eigenvectors are simply $\vec{e}_1 = (1, 0)$ and $\vec{e}_2 = (0, 1)$, with eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 2$ respectively, so that $y_1 = 2$ and $y_2 = 1$. The scaling fields are x and y. In terms of the nomenclature to be introduced below, these couplings (or equivalently, βJ and βh) are both *relevant*.

correlation length can only decrease under renormalisation group transformations, but it has to be infinite at the critical point.

The directions orthogonal to the critical surface are spanned by the relevant parameters. If you want to achieve infinite correlation length, these parameters have to be tuned perfectly, because any small deviation from the critical surface will get amplified by successive renormalisation group transformations. (end of lecture 13)

▸ ***See also:*** [2] chapter 3, [1].

4.3. Momentum space flow and quantum field theory

We have so far discussed the renormalisation group transformation in terms of 'coordinate space', explicitly constructing blocks of microscopic degrees of freedom and clumping them together, integrating out the fluctuations within the block. In quantum field theory, we are more accustomed to a discussion in momentum space, as it is there that an interpretation of asymptotic states (free particles with some given momentum) becomes most transparent. So let us transcribe the renormalisation group idea to that setting.

In coordinate space, the renormalisation group has been implemented by summing over fluctuations with wavelength between a and ba, so that we obtain a new lattice system with lattice spacing ba. All distances (measured in units of the lattice spacing) get scaled down, i.e. $\tilde{\xi} = b^{-1}\xi$. If we do more complicated transformations than the simple decimation we have seen in one dimension, it may also be necessary to rescale the remaining variables by a factor, in order to get back to the original form of the Hamiltonian (plus possible new interaction terms). In momentum space the story will be very similar. Summing over the small wavelength fluctuations is now replaced by summing or integrating over high-momentum modes. After that, all distances are scaled with b^{-1}, and thus all momenta are scaled with b. Sometimes a rescaling of the variables is required.

Let us make this more explicit, by considering a scalar field on a lattice. The Hamiltonian is given by (we will start straight in the continuum)

$$H = \int \mathrm{d}^D x\, \left[\frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m_0^2\phi^2 + \frac{1}{4!}\lambda_0\phi^4 + \dots\right] , \tag{4.18}$$

and of course it is possible to write down an arbitrary number of higher-order terms. Here we immediately see that we will have to rescale the field ϕ after the blocking transformation: if $\tilde{x} = b^{-1}x$ then $\tilde{\nabla} = b\nabla$ and in order to keep the same normalisation of the gradient term, we need to scale

$$\tilde{\phi}(\tilde{x}) = b^{(D-2)/2}\phi(x) =: b^{d_\phi}\phi(x) . \tag{4.19}$$

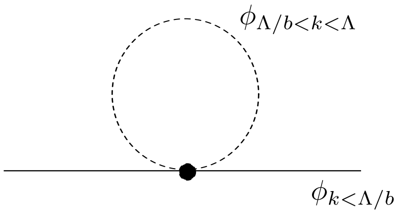
The exponent $d_\phi = (D-2)/2$ is called the *engineering dimension* of the field, or perhaps more useful, the *mass dimension* (if the mass dimension is one, the field scales in a way opposite to length). In a similar way we can work out the dimensions of the coupling constants that appear in the Hamiltonian. If we now go to momentum space using a Fourier transformation, we get

$$\phi(\vec{k}) := \int \frac{\mathrm{d}^D x}{L^{D/2}} e^{i\vec{k}\cdot\vec{x}}\phi(\vec{x}) . \tag{4.20}$$

The normalisation $L^{-D/2}$ is there also in the inverse transform. From this we find that the field in momentum space transforms as

$$\tilde{\phi}(\tilde{k}) = b^{d_\phi - D/2}\phi(k) . \tag{4.21}$$

The idea of the momentum space renormalisation group transformation is now to split the Hamiltonian in modes with momentum below Λ/b and those between Λ/b and Λ. We will then integrate over the latter. This integration is trivial when $\lambda_0 = 0$, because we then just have a Gaussian integral. If the coupling is non-zero, we generate additional terms. It is not easy to do the integral in that case, but one way to compute them is to resort to perturbation theory in λ_0. At the very lowest order, there is only one diagram that contributes, namely



(4.22)

This generates an effective interaction term for the long wavelength modes,

$$\tilde{H} = \int \mathrm{d}^D\tilde{x}\, b^{D-2d_\phi-2}\left[\frac{1}{2}(\tilde{\nabla}\tilde{\phi})^2 + \frac{1}{2}b^2\Big(m_0^2 + \frac{\lambda_0}{2}G\Big)\tilde{\phi}^2 + b^{2-2d_\phi}\frac{\lambda_0}{4!}\tilde{\phi}^4\right]. \qquad (4.23)$$

This gives us for the renormalisation group transformation

$$\begin{aligned}\tilde{m}_0^2 &= b^2\Big(m_0^2 + \frac{\lambda_0}{2}G\Big)\,,\\ \tilde{\lambda}_0 &= b^{4-D}\lambda_0\,.\end{aligned} \qquad (4.24)$$

There is a simple fixed point of these equations, at $m_0^2 = \lambda_0 = 0$. To make things more explicit and prepare for linearisation, we at this stage in perturbation theory only need the integral

$$\int_{\Lambda/b}^{\Lambda} \frac{k^{D-1}}{k^2}\mathrm{d}k = \frac{\Lambda^{D-2}}{D-2}(1 - b^{2-D})\,. \qquad (4.25)$$

Linearising around the fixed point gives us a renormalisation group matrix

$$R = \begin{pmatrix} b^2 & B(b^2 - b^{4-D}) \\ 0 & b^{4-D}\end{pmatrix}\,,\quad B = \frac{K_D\Lambda^{D-2}}{2(D-2)}\,. \qquad (4.26)$$

The eigenvectors and -values are

$$\vec{e}_1 = (1,0)\,,\ \lambda_1 = b^2\,,\ y_1 = 2\,,\qquad \vec{e}_2 = (-B,1)\,,\ \lambda_2 = b^{4-D}\,,\ y_2 = 4-D\,. \qquad (4.27)$$

For $D > 4$ this has one critical exponent $y_1 = 2$ and one exponent $y_2 = 4 - D < 0$. That is, there is one relevant direction, corresponding to the m_0 parameter, and one irrelevant direction, corresponding to λ_0.

(end of lecture 14)

5

Lattice gauge theory

To tie the various element of this course together, let us finally have a look at a lattice formulation of the theory of strong interactions, described by a SU(3) gauge theory. We will see various things which we have encountered in earlier chapters some back in a somewhat more complicated setting, in particular the strong coupling expansion and the condensation of vortex-like configurations in certain phases of the model.

5.1. Lattice action

To formulate gauge theory on the lattice, we need to find a discretised version of the action

$$S = \frac{1}{4}\int \mathrm{d}^4x\ \mathrm{Tr}(F_{\mu\nu}F^{\mu\nu})\,, \tag{5.1}$$

where the gauge field strength is given by

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_{\text{YM}} f^{abc} A^b_\mu A^c_\nu\,. \tag{5.2}$$

Here f^{abc} are the structure constants of the group. This action is invariant under the non-abelian gauge transformations

$$\begin{aligned} A_\mu &\to g^{-1}A_\mu g + \frac{1}{g_{\text{YM}}} g^{-1}\partial_\mu g\,,\\ F_{\mu\nu} &\to g^{-1}F_{\mu\nu}g\,. \end{aligned} \tag{5.3}$$

So we better make sure that this invariance exists in the lattice version of the theory as well.

Let us first introduce some lattice language. We will denote a point on the lattice by

$$x^\mu = m^\mu a\,, \quad m^\mu = 0, 1, \ldots, N-1\,. \tag{5.4}$$

For derivatives of a field at a point x, we write

$$\partial_\mu \phi_x = \frac{1}{a}(\phi_{x+a\hat\mu} - \phi_x)\,, \tag{5.5}$$

where $\hat\mu$ is a unit-vector in the μ direction.

The idea will now be to incorporate gauge invariance by associating the vector potential A_μ not to the nodes of the lattice, but to the links connecting them. More precisely, we will attach a *link variable* U to each link, given by

$$U_\mu(x) = e^{i g_{\text{YM}} a A_\mu(x+a\hat\mu/2)}\,. \tag{5.6}$$

Here a is, as always, the lattice spacing, and the relation to the continuum A_μ is made by evaluating the latter at the midpoint between the nodes x and $x + a\hat{\mu}$. If we now perform a gauge rotation $g(x)$ at each site, the link variables transform as

$$U_\mu(x) \rightarrow g(x) U_\mu(x) g^{-1}(x + a\hat{\mu})\,. \tag{5.7}$$

This is precisely what we need to know to construct a gauge invariant action: if we string together a set of link variables such that the contour is closed, we obtain a gauge invariant object.

The simplest such object is the *plaquette action*, which is obtained by going around a small two-dimensional square of the lattice and multiplying all link variables that we encounter,

$$S_{\text{plaq}} = \frac{1}{2g_{\text{YM}}^2} \Re \operatorname{Tr} \Big(U_{ij} U_{jk} U_{kl} U_{li} \Big)\,. \tag{5.8}$$

We can then construct an action by summing over this expression over all possible plaquettes in the lattice. This indeed reproduces, in the small-a limit, the continuum action. To see that, use the definition of the link variable and expand for small a. Using CBH then gives

$$UUUU \approx e^{i a^2 g_{\text{YM}} F_{\mu\nu}}\,. \tag{5.9}$$

Upon inserting this in the trace and expanding to second order in a we get the required continuum result.

5.2. Wilson loops

Having access to a lattice action, how now do we go about and construct quantities which correspond to interesting observables?[1] One thing we would like to compute is the interaction potential between two static quarks. We can compute that by creating a quark-antiquark state, evolving it in time and annihilating it again,

$$\langle q\bar{q} | e^{-\hat{H}T} | q\bar{q} \rangle = e^{-V(R)T}\,, \tag{5.10}$$

where R is the distance between the quarks. In a path integral setting, what we need to compute is the path integral in the presence of a source, which peaks at the two positions of the quarks, i.e.

$$\int \mathcal{D}A_\mu^a \, \exp\Big[- S + i g_{\text{YM}} \int A_\mu^a J_\mu^a \mathrm{d}^4 x \Big]\,, \tag{5.11}$$

with suitably chosen $J_\mu^a(x)$. Incorporating the quark-antiquark creation process, we are computing the expectation value of a link variable evaluated along the closed path,

$$V(R) = -\lim_{T\to\infty} \frac{1}{T} \ln \langle \operatorname{Tr} \mathrm{P} e^{i g_{\text{YM}} \oint_C A_\mu \mathrm{d}x^\mu} \rangle \tag{5.12}$$

The object on the right-hand side is called a *Wilson loop*. The symbol 'P' denotes path ordering, i.e. the matrices $A_\mu \mathrm{d}x^\mu$ are multiplied together in the order in which they appear on the contour C. In the lattice version, this becomes a bit more well-defined: we now have

$$V(R) = -\lim_{T\to\infty} \frac{1}{T} \langle \prod_C U_\mu(n) \rangle\,, \tag{5.13}$$

---

[1] We cannot simply compute the expectation value of a link variable, because it is not gauge invariant, and so it has to vanish. The simplest gauge-invariant object is obtained by stringing together link variables in a closed loop.

that is, the expectation value of the products of link variables, evaluated in the order in which they appear on the contour C.

If this expression scales linearly with distance, we have a confining potential,

$$V(R) = \kappa R + \dots\,. \tag{5.14}$$

We will see that this behaviour is indeed present when we look at the strong-coupling limit of the model, for which we can use techniques seen earlier (the high-temperature expansion).

5.3. Strong coupling expansion and confinement

We would like to see that at strong coupling, the potential is a linear function of R. This can be done using methods very similar to the strong coupling expansion we have seen in the Ising model. What we need to compute is

$$\langle\prod_C U_\mu(n)\rangle = \int \prod_{n,\mu} \mathrm{d}U_\mu(n) \prod_C U_\mu(n) e^{-S} \Big/ \int \prod_{n,\mu} \mathrm{d}U_\mu(n) e^{-S}\,. \tag{5.15}$$

This involves integrating over the group for every link variable. We need some identities for these group integrals,

$$\int \mathrm{d}U = 1\,,\quad \int \mathrm{d}U\, U_{ij} = 0\,,\quad \int \mathrm{d}U\, U_{ij} U^\dagger_{kl} = \frac{1}{2}\delta_{il}\delta_{jk}\,. \tag{5.16}$$

These are similar, in some sense, to the relations we had for the exponential integrals in the Kosterlitz-Thouless model (2.38). Since we have one link variable for every link in the contour C, we will need to bring down, from the action, a set of plaquettes which fills the contour, as in the figure. Since the plaquette action has an overall factor $1/g^2_{\text{YM}}$ (see (5.8)), this gives

$$\langle\prod_C U_\mu(n)\rangle = \left(\frac{1}{g^2_{\text{YM}}}\right)^{N_c} = \exp(-\text{Area}\cdot \ln g^2)\,. \tag{5.17}$$

We thus see that we indeed obtain an area law, and the string tension at strong coupling is given by

$$\kappa = \ln g^2_{\text{YM}} + \dots\,. \tag{5.18}$$

The conclusion is that for large bare coupling g_{YM}, the colour force is confining.

▸ ***See also:*** [4] section 3.2.

5.4. Monopole condensation

Unfortunately, the fixed point in QCD is the one where the bare coupling goes to zero. That is to say, the continuum limit requires us to go away from the strong-coupling regime just analysed. It then becomes less clear what is responsible for the confining force (which, in contrast to the bare coupling, is a physically measurable quantity). However, if we just do perturbation theory at small coupling, no sign of the confining force arises.

We thus need some new effects, very similar to what we have seen in the Kosterlitz-Thouless model, to understand what makes the regime of small bare coupling produce a confining force. Electron condensation leads to magnetic flux tubes connecting monopoles. Monopole condensation leads to electric flux tubes connecting electrons.

(end of lecture 15/16)

▸ ***See also:*** See e.g. [10] section 5.8 for more literature pointers.

Bibliography

[1] M. Le Bellac, “Quantum and statistical field theory”, Oxford University Press, 1991.

[2] J. L. Cardy, “Scaling and renormalization in statistical physics”, Cambridge University Press, 1996.

[3] J. M. Yeomans, “Statistical mechanics of phase transitions”, Oxford University Press, 1992.

[4] J. B. Kogut and M. A. Stephanov, “The phases of quantum chromodynamics: From confinement to extreme environments”, 2004.

[5] A. M. Polyakov, “Gauge fields and strings”, Harwood, 1987. Contemporary concepts in physics, volume 3.

[6] M. H. Thoma, “New developments and applications of thermal field theory”, `hep-ph/0010164`.

[7] G. P. Lepage, “What is Renormalization?”, `hep-ph/0506330`.

[8] M. Creutz, “Quarks, gluons and lattices”, Cambridge University Press, 1983.

[9] I. Montvay and G. Münster, “Quantum fields on a lattice”, Cambridge University Press, 1994.

[10] J. Smit, “Introduction to quantum fields on a lattice: A robust mate”, Cambridge Lect. Notes Phys., 2002.

[11] J. Zinn-Justin, “Quantum field theory and critical phenomena”, *Int. Ser. Monogr. Phys.* **113** (2002) 1–1054.

[12] J. Zinn-Justin, “Quantum field theory at finite temperature: An introduction”, `hep-ph/0005272`.

[13] A. Zee, “Quantum field theory in a nutshell”, 2003.

[14] K. Peeters, “A guide to biophysics and biomathematics”, `http://maths.dur.ac.uk/~dma6kp/biophysics.html`.

[15] K. Huang, “Statistical mechanics”, Wiley, 1987.

[16] R. Savit, “Duality in field theory and statistical systems”, *Rev. Mod. Phys.* **52** (1980) 453.