An Introduction to Basic Statistical Mechanics for Mathematicians I: Background

Paul Martin

(work in progress!)
DRAFT FOR SMDG ONLY.
NOT FOR DISTRIBUTION!

Contents

0.1 Preface

Part 1 is a version of some talks given to the Leeds Statistical Mechanics Discussion Group. The brief was to introduce basic statistical mechanics, so as to explain the common setting in which the various different interactions with Mathematics sit.

The notes do not require the reader to have read any of [9]. However this is a useful companion work.

0.1.1 What/Why Statistical Mechanics?

We are going to have to assume some basic Physics, and hopefully move on from there. ...So where does Statistical Mechanics fit in to Physics?

> *"The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved."* **Paul Dirac**

In this quote Dirac points out that the problems of Physics do not end, by any means, with the determination of fundamental principles. They include such fundamental problems; and also problems of computation.
(Indeed for the subject we are going to describe here, its original historical development was assumed to be on the fundamental side. Only a better understanding of its setting later showed otherwise.)
An example of the laws that Dirac is refering to would be Newton's laws,

which do a good job of determining the classical dynamics of a single particle moving through a given force-field. Two-body systems are also manageable but after that, even though it may well still be Newtonian (or some other well-understood) laws that apply in principle, exact dynamics will simply not be computationally accessible.

Do we really need to know about many-body dynamics? Yes. At least some understanding of the modelling of many-body systems is needed in order to work with a number of important materials (magnets, magnetic recording materials, LCDs, non-perturbative QFT etc). In each such case, the key dynamical components of the system are numerous, and interact with each other. Thus the force fields affecting the movement of one, are caused by the others; and when it moves, its own field changes, moving the others.

The solution:
The equilibrium Statistical Mechanical approach to such problems is to try to model only certain special types of observation that could be made on the system. One then models these observations by weighted averages over all possible instantaneous states of the system. In other words dynamics is not modelled directly (questions about dynamics are not asked directly). As far as is appropriate, dynamics is encoded in the weightings – the probabilities asigned to states.

The first problem is to describe these states, and determine appropriate probabilities.

It is most convenient to pass to an example. We shall have in mind a bar magnet. [1] We shall assume that the metal crystal lattice is essentially fixed (the formation of the lattice is itself a significant problem, but we will have enough on our plate). The set of states of the system that we shall allow are the possible orientations of the atomic magnetic dipoles (not their positions, which shall be fixed on the lattice sites).

What next?

---

[1] This provides a number of simplifications of the general problem, without trivialising the key features.

Chapter 1

Background

1.1 Towards the partition function

1.1.1 Classical reminders

A good rule of thumb when analysing a physical system is: "follow the energy". (This begs many questions, all of which we ignore.)

The kinetic energy of a system of N point particles with masses m_i and velocities v_i is

$$E_{kin} = \sum_{i=1}^{N} \frac{1}{2} m_i v_i^2$$

What can affect a particle's subsequent velocity, and hence change its kinetic energy? That is, what causes $\frac{dv}{dt}$ to be non-zero? A force can do this:

$$F = m\frac{dv}{dt}$$

Thus we also need to understand the forces acting on the particles.

For example: If they are really pointlike then they interact pairwise via the Coulomb force

$$F_1 = \frac{q_1 q_2}{4\pi\epsilon_0} \frac{\underline{r}_{12}}{r_{12}^3} = -F_2$$

Here q_1, q_2 are the charges (perhaps in coulombs); ϵ_0 is a constant (depending on that unit choice); and $\underline{r}_{12} = \underline{r}_1 - \underline{r}_2$.

For a moment we can think of this as a force field created by the second particle, acting on any charged first particle. This is a conservative force field; meaning that there is a function $\phi(\underline{r})$ such that

$$F = -\nabla\phi$$

The function $\phi(\underline{r})$ is part of the potential energy of the first particle. In other words its 'total energy' is of the form

$$E = \frac{1}{2}mv^2 + \phi$$

In practice, since ϕ is only defined up to an additive constant, E itself is not so significant as *changes* in E.

1.1.2 Stats/Gibbs canonical distribution

Notice that system energy E depends on the velocities and positions of all the atoms in the system. There are 10^{23} or so atoms in a handful of Earthbound matter, so we are not going to be able to keep track of them all (nor do we really want to). We would rather know about the bulk, averaged behaviour of the matter.

Let us call the inaccessible complete microscopic specification of all positions and velocities in the system a 'microstate'. Then for each microstate σ we know, in principle, the total energy $E(\sigma)$. We could ask: What is the probability P of finding the system, at any given instant, in a specific microstate?

Then we could compute an expected value for some bulk observation $\mathcal{O}$ by a weighted average over the microstates:

$$\langle\mathcal{O}\rangle = \sum_\sigma \mathcal{O}(\sigma)P(\sigma) \tag{1.1}$$

In principle the probability P could depend on every aspect of σ. This would make computation very hard. At the other extreme, P could be independent of σ. But this turns out to be a problematic assumption for a number of Mathematical and Physical reasons. Another working assumption would be that two microstates are equally likely if they have the same energy; i.e. that P depends on σ just through E. That is, that P depends only on the total energy of the system. Let us try this.

The next question is: How does P depend on E? What is the function $P(E)$?

If we have a large system, then we could consider describing it in two parts (left and right side, say), separated by some notional boundary, with the total microstate σ being made up of σ_L and σ_R. These halves are in contact, of course, along the boundary. But if the system is also in contact with other systems (so that energy is not required to be locally conserved), then it is plausible to assume that the states of the two halves are independent variables. In this case

$$P(\sigma) = P(\sigma_L)P(\sigma_R)$$

as for such probabilities in general. Similarly, the total energy

$$E = E_L + E_R + E_{int}$$

(where E_{int} is the interaction energy between the halves) is reasonably approximated by

$$E \sim E_L + E_R$$

(Why is this reasonable?!... Clearly the kinetic energy is localised in each of the two halves. The potential energy is made up of contributions from all pairs, including pairs with one in each half. But we assume that the pair potential is greater for pairs that are closer together; and that the boundary is a structure of lower dimension that the system overall. In this sense E_{int} is localised in the boundary (pairs that are close together but in separate halves are necessarily close to the boundary); while being part of the overall potential energy, which is spread with essentially constant density over the whole system. Thus E_{int} is a vanishing proportion of the whole energy for a large system. (We shall return to these core Physical assumptions of Statistical Mechanics later. They imply an intrinsic restriction in Statistical Mechanics to treating interactions that are, in a suitable sense, short-range. Fortunately this seems Physically justifiable.))

The L and R subsystems will each have their own 'energy-only' probability function. Thus we have something like

$$P(E_L + E_R) = P_L(E_L)P_R(E_R) \tag{1.2}$$

In this expression E_L and E_R are independent variables, so

$$\frac{\partial P(E_L + E_R)}{\partial E_L} = \frac{\partial P(E_L + E_R)}{\partial E_R}$$

so $P_L'(E_L)P_R(E_R) = P_L(E_L)P_R'(E_R)$, so

$$\frac{P_L'(E_L)}{P_L(E_L)} = \frac{P_R'(E_R)}{P_R(E_R)}$$

This separates. We write $-\beta$ for the constant of separation. We have $P_L'(E_L) = -\beta P_L(E_L)$ (and similarly for R). This is solved by a function of form

$$P(E) = C\exp(-\beta E) \tag{1.3}$$

where C is any constant. In our case C is determined by

$$\sum_\sigma P(E(\sigma)) = 1$$

The separation constant β is interesting, since it is the only thing (other than the form of the function itself) that connects the subsystems. We will see later that this connection corresponds (inversely) to a notion of *temperature*.

1.1.3 Partition Function

The normalisation function for our system (1.3)

$$Z(\beta) = \sum_\sigma \exp(-\beta E(\sigma))$$

(Z for *zustatensummen*, or some such name due to Boltzmann) is called the partition function. That is, for given β,

$$P(E) = \frac{\exp(-\beta E)}{Z}$$

Recall that, by our derivation, β represents the effect of thermal (energetic) contact with the universe of other systems. Our usual notion of the bulk contribution of neighbouring systems on the energetics of a given system, at least where long-time-stable (equilibrium) properties are concerned, is the notion of temperature. Thus β encodes temperature. How specifically does it do this? See later.

First we want to consider the pay-off for the analysis we have made so far. The idea was that we would be able to compute time-averaged bulk properties of the system.

Z is 'just' a normalising factor. But

$$\frac{d \ln Z}{d\beta} = -\frac{1}{Z} \sum_{\sigma} E(\sigma) \exp(-\beta E(\sigma))$$

the *internal energy* (or compute $\frac{1}{N}$ of this for the energy *density*), so its analysis contains Physics!

Suppose

$$E : S \to \pm\mathbb{N}$$

(i.e. the energy is quantised).[1] Then Z is polynomial in $\exp(\mp\beta)$. Its only analytic structure is *zeros*. However, we will see how these zeros are indeed physically significant.

Figure 1.1 gives a quick glimpse of the distribution of zeros in the complex $x = \exp(\beta)$ plane for a three-dimensional Ising model (what this is will be explained later).

Question: What does this tell us?

The *specific heat* tells us the rate of change of internal energy with temperature (or, in practice, the other way round — we measure the amount of energy we have to put into a fixed mass of material to change its temperature by 1 degree):

$$\mathcal{S} = \frac{\partial^2 \frac{1}{N} \ln Z}{\partial \beta^2}$$

Consider the formulation

$$Z = \prod_j (x - z_j)$$

In terms of the complex zeros $\{z_j = x_j + iy_j\}$ we have:

$$\frac{\partial \frac{1}{N} \ln Z}{\partial \beta} = \frac{x}{N} \sum_j \frac{1}{x - z_j}$$

$$\mathcal{S} = \frac{\partial^2 \frac{1}{N} \ln Z}{\partial \beta^2} = \frac{x}{N} \left(\sum_j \frac{1}{x - z_j} - x \sum_j \frac{1}{(x - z_j)^2} \right)$$

---

[1] This is far from always true, but it is not unrealistic.

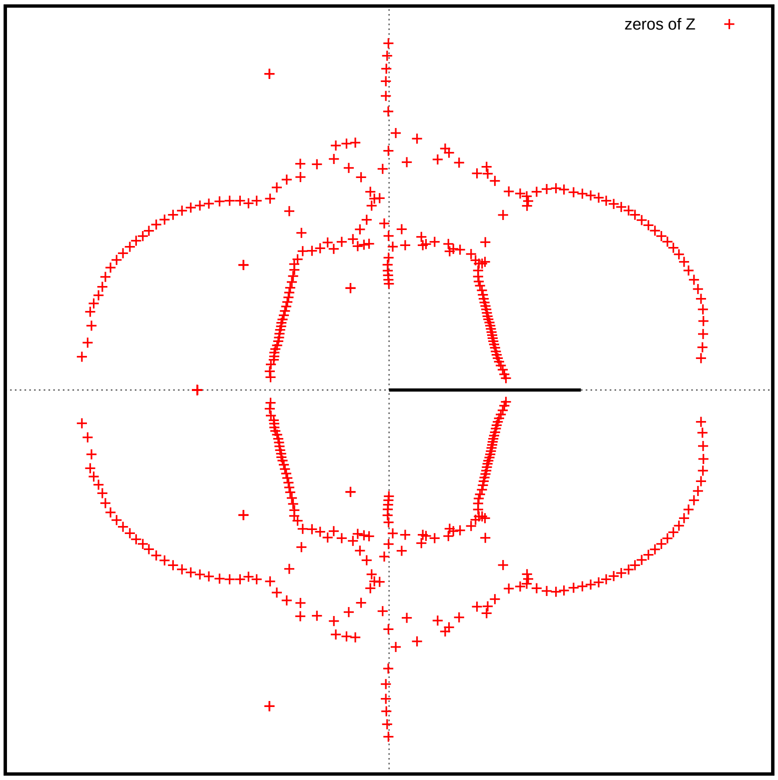


Figure 1.1: Complex zeros of the partition function for a cubical lattice Ising model of about $N = 150$ sites [11, 15].

Note that the complex zeros appear in conjugate pairs. Performing the sum within each conjugate pair this becomes

$$\mathcal{S} = \frac{2x}{N}\left({\sum_j}' \frac{x-x_j}{(x-x_j)^2+y_j^2} - x {\sum_j}' \frac{(x-x_j)^2-y_j^2}{((x-x_j)^2-y_j^2)^2+4(x-x_j)^2y_j^2} \right)$$

Consider the contribution of the j-th term (i.e. from a pair of zeros) to $\mathcal{S}$, at some point on the real x axis. Note that if y_j is large, or if $x-x_j$ is large, then this contribution is small. Meanwhile the contribution is large if the zeros are close to the axis, and x is close to these zeros. In particular the contribution is large if $x = x_j$, whereupon

$$\mathcal{S}_j \sim \frac{2x^2}{Ny_j^2}$$

Simply put, this says that, moving along the real line (real temperature), $\mathcal{S}$ and U go crazy when there are complex zeros close by (as there are at a particular point in Figure 1.1, for example).
Accordingly we shall call a region of the complex plane that is close to the real axis and contains zeros of Z a *critical neighbourhood* of Z.

Let us (very crudely) compare with physical observation.
When we boil a kettle we put roughly equal amount of energy into the water in each unit of time. At first the temperature rises, and the rate of rise does not change very much as the temperature goes up. That is, the specific heat changes slowly and smoothly with temperature. Close to and at the boiling point, however, the temperature rise essentially stops, i.e. the amount of energy required to further change the temperature becomes very large. In the practical experiment there are a number of reasons for this, but one of them is that the specific heat becomes very large. Thus we associate divergent specific heat with a phase transition (in this case the liquid-gas transition at the boiling point).

Another quick picture, of a different model, is shown in Figure 1.2. Once again you are invited to study the distribution of zeros close to the 'physical' region: the real interval $[0, 1]$.

In practice we are never even going to know N accurately for a real physical system. Nor is such an accurate knowledge important to us, since realistic observations do not depend on it. Indeed for intensive quantities we

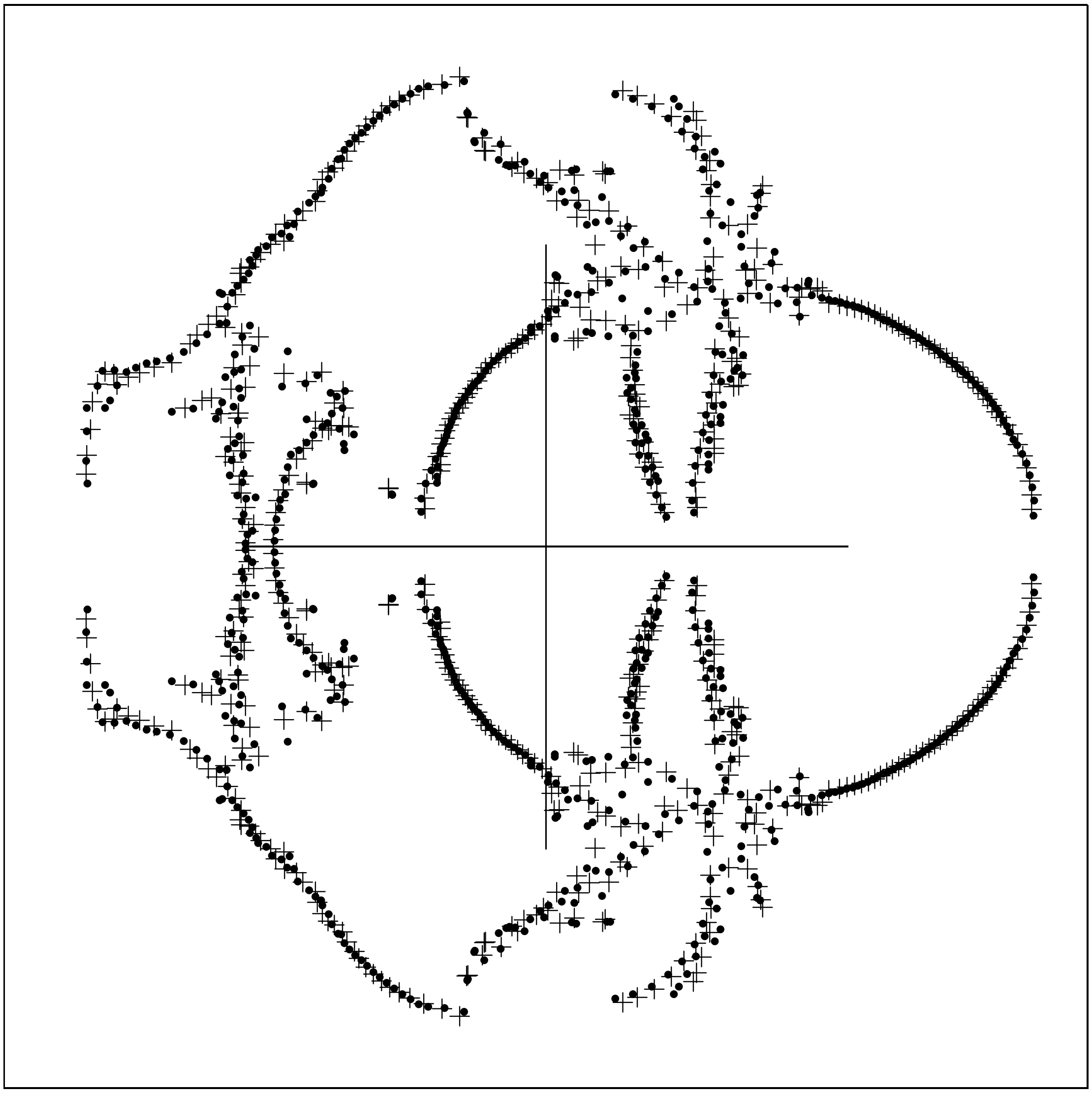


Figure 1.2: Complex zeros of the partition function for a two-dimensional clock model at two different lattice sizes [10].

expect them to be stable under even large changes in N, so long as N is large enough. Thus intensive observations in our model also need to be large N stable. [2]

Our pictures give a clue as to the sense in which this can happen: Different polynomials (different N; different but suitably 'similar' physical systems) could have similar distributions, real accumulation points, etc.; and hence manifest behaviour on the physical line in similar ways.
Can we at least have a model for *this*?

1.2 Models

We now briefly introduce a simple choice for E (the microstate energy function), from a Physics perpective.

As noted in Section 1.1.3, to introduce a choice for E for physical modelling, we must actually introduce one for each of a whole collection of 'similar' systems, and then check the stability of observables across this collection. Mathematically, it is convenient to introduce an E for each of a rather large collection of (nominally but not necessarily adequately similar) systems, then refine this collection by physical considerations post hoc. (We will make all this very precise later.) This is called choosing a 'model'. The choice we shall describe here is called the *Potts model*.

(**1.2.1**) Some statistical mechanics nomenclature: While kinetic-energy-only (non-interacting) models are rather simple, models in which only the potential energy is accounted for in the microstate energy are much richer (partly because aspects of the kinetic energy of the system are encoded in β anyway). Excluding kinetic energy from E means that we are essentially treating our particle positions as fixed (that is, not translating). Instead system dynamics is manifested in other ways. For example we can consider non-pointlike particles, hence with the possibility of magnetic dipole moments. System dynamics in this case can be manifested in variations in magnetic dipole orientation. In such a setting, particles are called *spins*. Also, the microstate energy function is called the *Hamiltonian* (and typically written H not E).

(**1.2.2**) Let Γ be the set of graphs. For $G \in \Gamma$ let V_G be the vertex set of G (sometimes we will simply write G for V_G), and E_G the edge set. We adopt

---

[2] Note that we do not require a large N *limit* per se for Physics. But stability and the existence of such a limit amount to the same thing computationally.

the notation

$$\underline{n} = \{1, 2, ..., n\}$$

For any set S (such as a set of graph vertices) and $Q \in \mathbb{N}$ write

$$Q^S = Hom(S, \underline{Q})$$

for the set of maps from S to $\underline{Q}$ — each map f asigns a state in $q \in \{1, 2, ..., Q\}$ to each vertex s by $f(s) = q$.

1.2.1 Potts and Z_Q-symmetric lattice models

(**1.2.3**) For $Q \in \mathbb{N}$, a *Q-state graph Hamiltonian* is a map asigning to each graph G a map H_G in $Hom(Q^{V_G}, \mathbb{Z})$, or more generally $Hom(Q^{V_G}, \mathbb{R})$. We say that this map H_G gives the 'energy' of state $\sigma \in Q^{V_G}$.

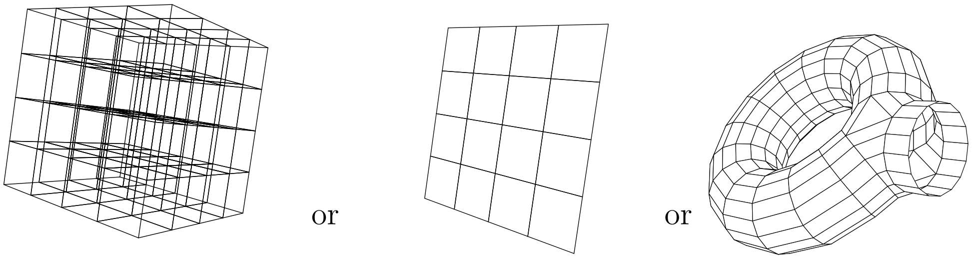
(**1.2.4**) The idea of a model like this (modelling Physics in some given physical space $\mathbb{E}$) is as follows.
(i) one considers a subset of graphs, such that each graph G considered represents (in principle) a given physical system — a collection of degrees of freedom embedded in $\mathbb{E}$.
(ii) in particular the degrees of freedom of the system reside on the vertices of G, and each takes values from a range represented by $\{1, 2, ..., Q\}$; and
(iii) the geometrical relationship of the degrees of freedom is encoded (somehow) in the edges between the vertices.
In other words, the set of microstate variables is V_G (which does not depend on the edges of G), but the interactions H_G between spins will depend on the edges E_G.

(**1.2.5**) Fixing Q (e.g. $Q = 2$) we have, for example,

$$H_G^{Potts}(\sigma) = \sum_{(v,v') \in E_G} \delta_{\sigma(v),\sigma(v')} = \sum_{(v,v') \in E_G} \delta_{\sigma(v)-\sigma(v'),0} \tag{1.4}$$

$$H_G^{Clock}(\sigma) = \sum_{(v,v') \in E_G} \cos(2\pi(\sigma(v) - \sigma(v'))/Q) \tag{1.5}$$

Note that each of these indeed gives a Hamiltonian for each choice of G. However not every G makes sense physically. Here the idea is specifically that V_G represents the set of molecules on some crystal lattice; and E_G determines nearest neighbour molecules on the lattice. For example, see Figure 1.3.



3d metal 2d metal more exotic 2-manifold

Figure 1.3:

Such an embedding implies further useful structural properties, as we shall see shortly. We may write $\overline{G}$ to emphasise the extra structure on G, when it is of such a form.

(**1.2.6**) The 2d square lattice shown, for example, would be part of a sequence of graphs including also larger square lattices such as the one with 56 vertices and 97 edges shown in Figure 1.4(a).

(**1.2.7**) Figure 1.4(b) is an example of a $Q = 2$-state spin configuration on the vertices of the same square lattice. We draw an up-arrow for the spin state 1, and a down-arrow for the spin state 2. In this case there are only two down-arrows out of $N = 56$. Thus $H_G(\sigma) = 97 - 8 = 89$.

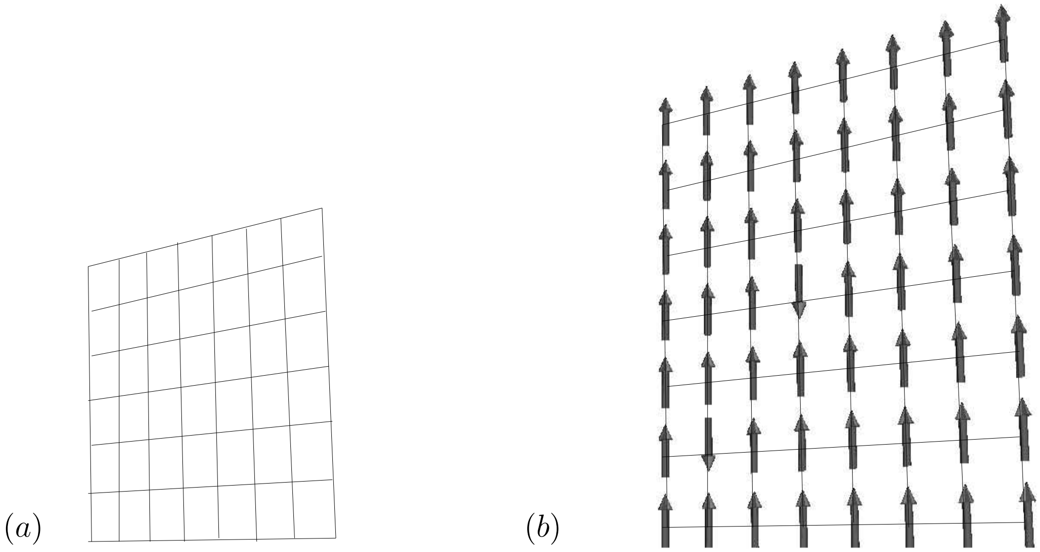


Figure 1.4:

Chapter 2

More on lattice models

Here we look in a little more detail at certain aspects of the 'lattice models' introduced in §1.2.1. This Chapter can be skipped at first reading.

2.1 Geometrical aspects

2.1.1 The dual lattice

Note that an embedded graph of the kind discussed in Section 1.2.1 may be equipped with the structure of a cell complex (see for example Hilton and Wylie [7, §2.12], Spanier [13, Ch.4]). That is, (I) the embedded graph $\overline{G}$ defines sets $s_c(\overline{G})$ of simplices (oriented convex polytopes) of dimensions $c = 0, 1, ..., d$, sometimes called c-cells.

(II) For each dimension c one also considers the free $\mathbb{Z}$-module with basis $s_c(\overline{G})$, the elements of which are called *c-chains*. The *boundary* operator ∂ is a $\mathbb{Z}$-linear map

$$\partial : \mathbb{Z}s_c(\overline{G}) \to \mathbb{Z}s_{c-1}(\overline{G});$$

for each c, such that $\partial^2 = 0$.

For example, if $e_1 = (v_1, v_2)$ is a directed edge of G then $\partial e_1 = v_2 - v_1$ (we also identify $-e_1 = (v_2, v_1)$).

(**2.1.1**) In particular $\overline{G}$ partitions the d-dimensional embedding space $\mathbb{E}$ into points, arcs, plaquettes and so on.

We write $|\overline{G}|$ for the union of all convex cells in the cell complex of $\overline{G}$. Strictly speaking, if it is non-empty we include the complement of $|\overline{G}|$ in $\mathbb{E}$ among the d-cells associated to $\overline{G}$.

(**2.1.2**) From this perspective the Potts spin configurations are $\mathbb{Z}_Q$-valued 0-cochains (regarding $\mathbb{Z}_Q = \mathbb{Z}/Q\mathbb{Z}$ as an additive group). That is, a *c-cochain* is a linear map

$$\sigma : \mathbb{Z}s_c(\overline{G}) \to \mathbb{Z}_Q$$

for any c, determined by the images of the c-cells. (And the images of the vertices are their values in the Potts spin configuration σ.)

The set of c-cochains form an abelian group by pointwise addition (not a very natural operation from the physical perspective, but a useful organisational device).

(**2.1.3**) We define a scalar product on cells by $(S, S') = \delta_{S,S'}$ and extend bi-linearly to chains. This leads to a *coboundary* operator $\partial^* : \mathbb{Z}s_c \to \mathbb{Z}s_{c+1}$ by:

$$(S, \partial^* S') = (\partial S, S')$$

for $S \in s_c$, $S' \in s_{c-1}$. Note $(\partial^*)^2 = 0$.

Consider Figure 1.4(a) for example. Let f_1 be the face in the top left-hand corner, and compute $\partial^* f_1$. We have $(S, \partial^* f_1) = (\partial S, f_1) = 0$ (f_1 is not in the boundary of any cell). Thus $\partial^* f_1 = 0$.
Meanwhule, let e_1 be the edge in the top left-hand corner. We have $(S, \partial^* e_1) = (\partial S, e_1) = \delta_{S,f_1}$ (up to choice of orientation), thus $\partial^* e_1 = f_1$.

We can look for a proper subset of chains that forms a ∂^* subcomplex. Such a subset can be generated by making a choice of 0-cells to include. For example, if we include all the *interior* 0-cells (in the obvious sense) then all the interior 1-cells and all the 2-cells must be included, but we do not need to include any of the *exterior* 0-cells or 1-cells.

(**2.1.4**) From this structure we may define a *Kramers-Wannier dual* graph to $\overline{G}$. This is a graph (with associated cell complex) $D(\overline{G})$ with a vertex v for each d-dimensional component of $\overline{G}$ (i.e. a 0-cell for each d-cell), and an edge between v and v' if their d-dimensional preimages in G share a common $(d-1)$-dimensional component (hence a 1-cell for each $(d-1)$-cell). Indeed

$$D : s_c(\overline{G}) \to s_{d-c}(D(\overline{G}))$$

is a bijection for each c.

For example, the dual of our square lattice graph above is shown in Figure 2.1.

(**2.1.5**) A c-cochain σ defines a dual $(d-c)$-cochain $D(\sigma)$ on $s_{d-c}(D(\overline{G}))$ via

$$D(\sigma)(D(S)) := \sigma(S)$$

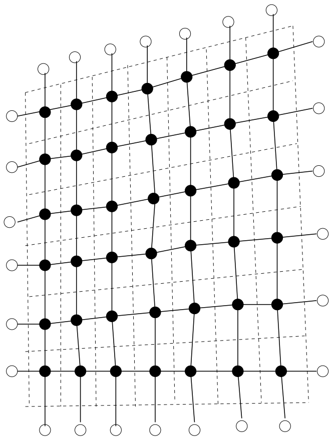


Figure 2.1: Dual lattice of a certain square lattice. All of the exterior vertices are to be identified as the same vertex.

2.1.2 Domain walls

(**2.1.6**) This dual graph $D(G)$ gives us an alternative way to describe spin configurations — in terms of islands of aligned spins, or specifically in terms of the positions of the boundaries of islands of aligned spins. Consider the square lattice case, and note that each (directed) edge of the dual graph is associated to a pair of (ordered) spins, i, j, say, on the original graph. Then for a given configuration σ we can asign a directed weight

$$w_{ij}(\sigma) = \sigma(j) - \sigma(i)$$

(modulo Q) to this dual edge. (We shall orient the dual edges so that, passing from i to j, the positive direction of the dual edge across the edge $i - j$ is from left to right. We asign w as above to this direction.) In particular, if two adjacent spins have the same value in σ then the dual edge weight is zero (and no energy is 'lost' at this edge).

For Q odd, a simple notation is to use the representatives of $\mathbb{Z}/Q\mathbb{Z}$ in the interval $[-Q/2, ..., Q/2]$ for w, and draw w positive-direction pointing arrows on the dual edge if $w \geq 0$, or $-w$ reverse arrows if $w < 0$. The same notation works for Q even, but note that $-Q/2 \equiv Q/2$. In the $Q = 2$ case every dual edge has weight either 1 or 0, for example.

The weight 'variables' are $\mathbb{Z}/Q\mathbb{Z}$-valued, but it is not appropriate to replace the sum over 0-cochains on the original lattice by the sum over arbitrary 1-cochains on the dual lattice, since the map w is not surjective. (It is also not injective, but it is made so if we pick and fix any one spin.) The constraint is that if we traverse a loop around any dual vertex, starting at some spin, then the sum of weights must be congruent to 0 mod. Q (else we do not have a consistent value for the starting spin).

(**2.1.7**) Note that for $Q = 2$ the sum of dual edge weights for edges incident at a give dual vertex, must be even. In general, the sum (taking care of signs, note) must be congruent to zero modulo Q. For $Q = 2$ that is to say, we can form sequences of dual edges with weight 1 into closed loops. This cannot necessarily be done uniquely (for example if four weight-1 edges are incident at a dual vertex), but all equivalent loop asignments correspond to the same spin configuration (up to a global up-down choice).

This means that we can replace the sum over spin configurations by a sum over 'edge coverings'. Coverings are weight asignments to the edges of the dual lattice satisfying the *covering rule*: only asignments with the signed sum of weights incident at each dual vertex congruent to 0 modulo Q are allowed. Every such weight asignment determines a spin configuration, given only the state of a single spin. (Note that this ambiguity can be resolved freely, since the Hamiltonian (1.4) is invariant under changes in this choice. That is

$$\sum_{spin\ configs} \exp(\beta H) \;=\; Q \sum_{\substack{allowed \\ coverings\ \gamma}} \exp(\beta H)$$

Note also that the Hamiltonian can be expressed simply in terms of the total length of weight-1 boundary in the covering:

$$H(\gamma) = |E_V| - l(\gamma) \tag{2.1}$$

where $l(\gamma)$ is the length of boundary in covering γ.)

For example, the picture of our spin configuration (b) above is the left-hand picture in Figure 2.2. And here $H = |E_V| - 8$. The right-hand picture in Figure 2.2 shows a different spin configuration and its boundary representation.

This representation of spin configurations is sometimes called the *domain wall* representation (for reasons that the figure makes clear).

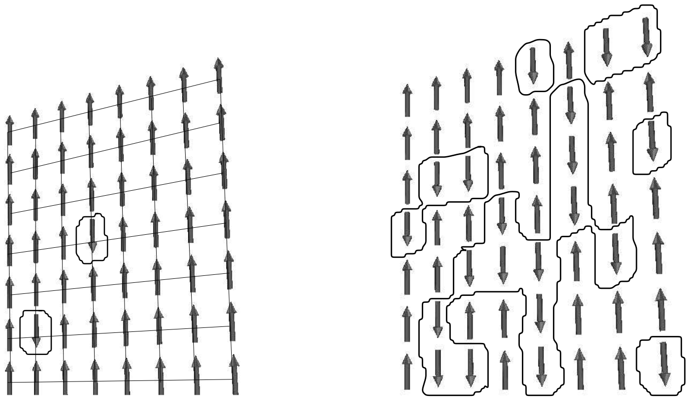


Figure 2.2: Two spin configurations with corresponding coverings.

(**2.1.8**) Note that the covering rule may appear to be modified if one considers only part of a system, or a system with boundary. Then the details of the connection with the exterior may be encoded in sinks and sources — dual sites where the incident weight sum is not congruent to 0. See later.

2.1.3 Trivial examples

Consider the lattice consisting of a single square. The dual lattice consists of two vertices, connected by four edges.

2.1.4 High and low temperature

2.1.5 Kramers–Wannier duality

2.1.6 Graphs with boundary

blah blah!!!

2.A Appendix: Combinatorics of the Potts groundstate

I have removed the appendix from this point, because it does not belong in an introductory text, and it is not finished. See SM-CombinPotts-1.tex.

Chapter 3

Computation

3.1 Transfer matrix formalism

3.1.1 Partition vectors

Suppose that we have fixed a graph Hamiltonian as in (1.2.3). Then for each graph G we have a partition function Z^G, associated to the Hamiltonian H_G.

(**3.1.1**) Let $Z^G_{V|_x}$ be Z^G but with vertex subset V *fixed* to x:

$$Z^G_{V|_x} = \sum_{s \text{ s.t. state } s|V=x} \exp(-\beta H)$$

Then the 'Partition vector' Z^G_V is a vector indexed by configurations of V, whose x-th entry, $(Z^G_V)_x$, is $Z^G_{V|_x}$.

(**3.1.2**) If $G = G' \cup G''$ where $V_{G'} \cap V_{G''} = V$, $E_{G'} \cap E_{G''} = \emptyset$, and H_G is 'local' in the sense that interactions are associated to pairs of vertices defined by edges, then the subgraph partition vector $Z^{G'}_V$ makes sense, and we have

$$Z^G = \sum_x (Z^{G'}_V)_x (Z^{G''}_V)_x \tag{3.1}$$

Typically G has topological properties (perhaps embedded in and representing some manifold), with respect to which V is a boundary, and the situation of equation(3.1) may be illustrated as in Figure 3.1 or 3.2.

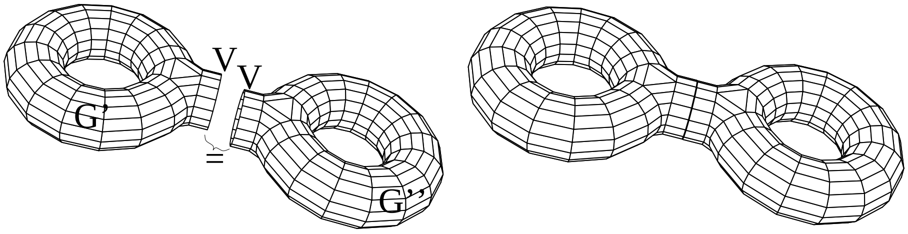
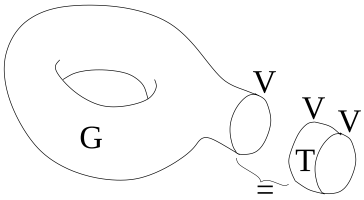


Figure 3.1: $\sum_x Z^{G'}_{V|_x} Z^{G''}_{V|_x} \ = \ Z^{G'\cup G''}$

□



$$\sum_x Z^G_{V|_x} Z^T_{V|_x\ V'|_y} = Z^{G\cup T}_{V'|_y}$$

Figure 3.2: Transfer Matrix $\mathcal{T}_{xy} = Z^T_{V|_x\ V'|_y}$.

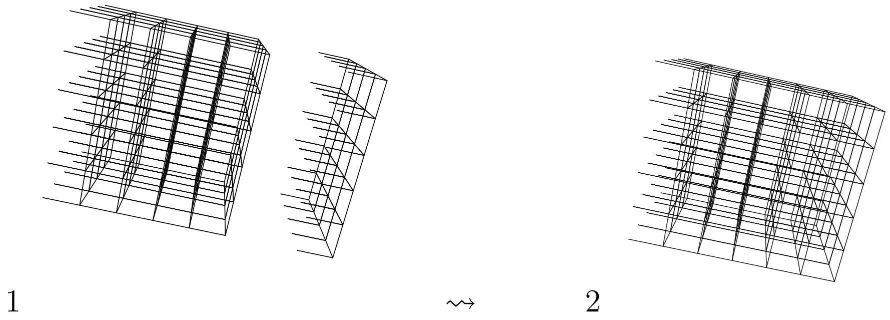


Figure 3.3: 1. Adding a lattice layer; 2. New larger lattice.

3.1.2 Transfer matrices

Figure 3.2 also server, formally, to define the *transfer matrix* $\mathcal{T} = Z^T_{V,V'}$, as a partition vector with two parts to the 'boundary' (note that this is simply an organisational arrangement). Suppose we iterate composition of a suitable $\mathcal{T}$, as illustrated in Figure 3.3. Then we get

$$Z^{big} = \langle \mathcal{T}^n \rangle$$

(for suitable initial and final boundary conditions $\langle - \rangle$). If $\{\lambda_i\}_i$ are the eigenvalues of $\mathcal{T}$ we have

$$Z^{big} = \langle \mathcal{T}^n \rangle = \sum_i k_i \lambda_i^n$$

where the k_is depend on the boundary conditions, but not n. For example with simple periodic b.c.s we have

$$\langle \mathcal{T}^n \rangle = \mathrm{Tr}(\mathcal{T}^n) = \sum_i \lambda_i^n$$

Note that here $\mathcal{T}$ is +ve symmetric, so the Perron–Frobenius Theorem implies

$$Z^{big} \sim k_0 \lambda_0^n \left(1 \; + \; \frac{k_1}{k_0} (\frac{\lambda_1}{\lambda_0})^n + \sum_{i>1} \frac{k_i}{k_0} (\frac{\lambda_i}{\lambda_0})^n \right) \sim k_0 \lambda_0^n$$

where λ_0 is the largest eigenvalue, unless $\lambda_1 \to \lambda_0$ as size$\to \infty$. So the Helmholtz free energy $\frac{1}{N} \ln(Z) \sim \ln(\lambda_0)$.

What about the physical role of other eigenvalues?

3.1.3 Correlation functions

Cold systems tend to be ordered, and hot systems disordered. Neither of these states exhibits long range *correlation* between local states. Thus only in the order/disorder transition region may there be such correlations. Experimentally, correlation of spins over long distance is indeed a signal of phase transition.

- Experimentally, at a fixed T away from T_c, an observation of the correlation of the state of two spins (say) as a function of their separation r, behaves like:

  $$\langle \sigma_i \sigma_{i+r} \rangle \sim e^{-r/\rho}$$

  (length scale $\rho(T)$ measured in terms of lattice spacing).

  As $T \to T_c$, $\rho \to \infty$ (crucial in lattice Field Theory).

- In Stat Mech

  $$\langle \sigma_i \sigma_{i+r} \rangle \sim \frac{(\mathcal{T}^{N_1} \hat{\sigma} \mathcal{T}^r \hat{\sigma} \mathcal{T}^{N_2})_.}{(\mathcal{T}^{N_1+r+N_2})_.} \sim \left(\frac{\lambda_\sigma}{\lambda_0}\right)^r$$

  $$= \exp(-r \underbrace{(\ln(\lambda_0) - \ln(\lambda_\sigma))}_{\frac{1}{\rho}})$$

- So other eigenvalues besides λ_0 have physical significance. (NB labelled by operator content, not N, should not depend on N in limit.)

3.2 Practical calculation

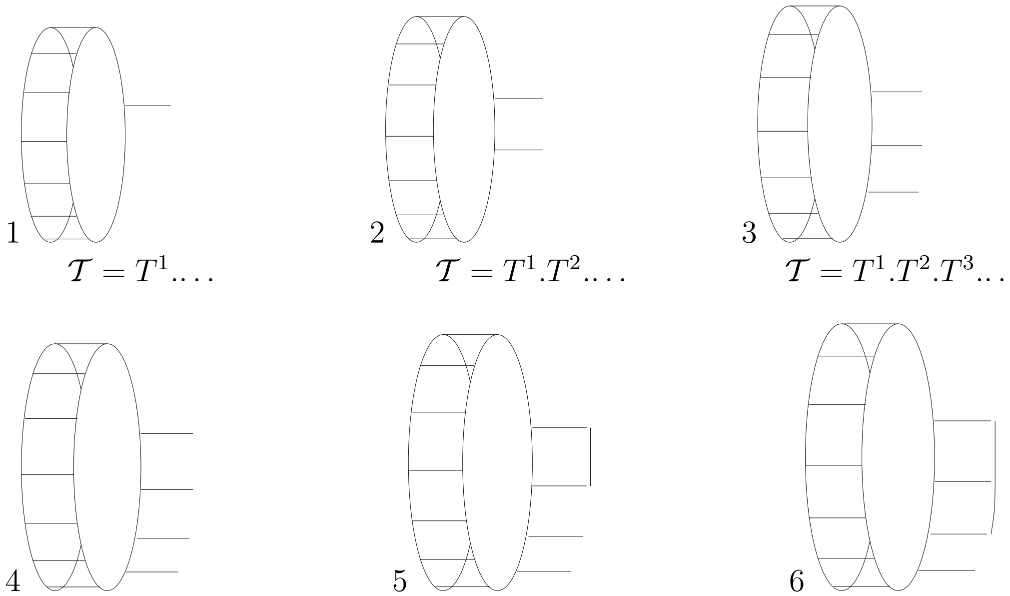
3.2.1 Use the force: transfer matrix algebras

Next idea: We look for an algebra A and a representation R such that we can express

$$\mathcal{T} = R(X)$$

with $X \in A$; then organise the spectrum of $\mathcal{T}$ by simple components of R.

There is no simple recipe for finding A, R, X to make this work. We shall discuss a limited systematisation as we go.



$\mathcal{T} = T^1 \ldots$ $\quad$ $\mathcal{T} = T^1.T^2 \ldots$ $\quad$ $\mathcal{T} = T^1.T^2.T^3 \ldots$

$\mathcal{T} = T^1.T^2.T^3.T^4 \ldots$ $\quad$ $\mathcal{T} = T^1.T^2.T^3.T^4.T^{12} \ldots$ $\quad$ $\mathcal{T} = T^1.T^2.T^3.T^4.T^{12}.T^{23} \ldots$

Figure 3.4: Growing a cylindrical lattice layer one interaction at a time

Local transfer matrix

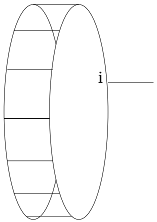
The transfer matrix method (essentially requiring that the lattice can made up of a number of layers) grows the lattice a single layer at a time. Now we go further, and grow the lattice a single interaction at a time.

Let us picture the situation in which we have built some number of complete layers, and now proceed to start building a new layer. We start by adding a single new edge/interaction: see Figure 3.4(1). Proceeding as illustrated, here we get

$$\mathcal{T} = \prod_i T^i \prod_{\langle i,j \rangle} T^{ij} \tag{3.2}$$

What is T^i here?

Consider the following example. Take $H = -\beta \sum_{\langle ij \rangle} \delta_{\sigma_i,\sigma_j}$ (2-state case, say) on a graph made up of closed chain layers (hence a cylindrical lattice, as it were), as in our recent figures. Set $x = e^{\beta}$. Consider the partition vector Z_W^G for some assembly of complete layers of lattice G, relative to some collection of 'boundary' spins W (as in (3.1.1)). One natural arrangement is to take W to be the union of the states in some initial layer (on the left) and the states in the most recent layer grown (on the right) — in which case the partition vector is the transfer matrix $\mathcal{T}^n$ for some n. Alternatively one might consider Z relative only to the states, V say, in the most recently grown layer — i.e. as $\langle \mathcal{T}^n$. But consider (for a moment) the partition vector relative to a *single* spin i in V, preparatory to adding a single new interaction involving that spin, as indicated:



$$Z_{\sigma_i} = \begin{pmatrix} Z_{\sigma_i=1} \\ Z_{\sigma_i=2} \end{pmatrix}$$

Now consider $|V| = m$ so Z_V is a Q^m-component vector

$$Z_V = \begin{pmatrix} Z_{V|\sigma_i=1} \\ Z_{V|\sigma_i=2} \end{pmatrix}$$

where each entry is a Q^{m-1}-component vector. The partition vector for the new system, over the new spin, after the new edge is added, is:

$$Z^+_{\sigma_i} = \begin{pmatrix} xZ_{\sigma_i=1} + Z_{\sigma_i=2} \\ Z_{\sigma_i=1} + xZ_{\sigma_i=2} \end{pmatrix} = \underbrace{\begin{pmatrix} x & 1 \\ 1 & x \end{pmatrix}}_{T^i=(x-1)\mathbb{I}_Q+D_Q} Z_{\sigma_i}$$

Thus the prefactor matrix on the right is the local transfer matrix.

What is T^{ij}?

Similarly

$$Z^{]}_{\sigma_i\sigma_j} = \begin{pmatrix} xZ_{11} & & & \\ & Z_{12} & & \\ & & Z_{21} & \\ & & & xZ_{22} \end{pmatrix} = \underbrace{\begin{pmatrix} x & & & \\ & 1 & & \\ & & 1 & \\ & & & x \end{pmatrix}}_{T^{ij}=\mathbb{I}_{Q^2}+(x-1)C_Q} Z_{\sigma_i\sigma_j}$$

Let us define

$$u_i := \frac{1}{\sqrt{Q}} \mathbb{I}_Q \otimes .. \otimes \underbrace{D_Q}_{i-th} \otimes \mathbb{I}_Q \otimes .. \otimes \mathbb{I}_Q$$

$$u_{ij} := \sqrt{Q}\mathbb{I}_Q \otimes .. \otimes \underbrace{C_Q}_{i-th \text{ and } j-th} \otimes \mathbb{I}_Q \otimes .. \otimes \mathbb{I}_Q$$

NB, these obey

$$u_i^2 = \sqrt{Q} u_i \qquad u_{ij}^2 = \sqrt{Q} u_{ij} \qquad u_i u_{ij} u_i = u_i \qquad u_{ij} u_i u_{ij} = u_{ij} \tag{3.3}$$

$$[u_i, u_j] = [u_{ij}, u_{kl}] = [u_i, u_{kl}] = 0 \qquad \{i\} \cap \{k, l\} = \emptyset \tag{3.4}$$

Algebra

- As abstract relations (3.3-3.4) define *Graph TL algebra (GTLA) for the complete graph K_m*. A sort of TL version of a Coxeter–Artin group[1].
- GTLA for graph G is subalgebra with generators u_i and u_{ij} if $(i,j) \in G$.

  This is generally not finite rank[2] but $G = A_m$ case is $\cong$ ordinary TL algebra.
- Thus in 2d

  $$\mathcal{T} = R\left(\prod_i (\frac{(x-1)}{\sqrt{Q}} 1 + u_i) \prod_{ij} (1 + \frac{(x-1)}{\sqrt{Q}} u_{ij}) \right)$$

  where R is a representation of OTLA.
- Thus spectrum of $\mathcal{T}$ decomposes by irreducible components of R. Thus correlation functions (particles) at least partially indexed by simples of algebra.
- This is the paradigm.

Global limit

- Even fixing the physical model, there is a $\mathcal{T}$, and hence a TMA, for each N.
- But physical observables, and hence spectrum components, defined essentially independently of N. For given N, spectrum components are (partly) indexed by simple module decomposition of

  $$\mathcal{T} = R(X)$$

  , thus these can be indexed independently of N.
- Thus expect global limit to sequence of algebras, and localisation functors picking out fibres of "physically equivalent" modules.
- How change system size? Example:

  Freeze two spins together in transfer matrix layer.

  What does this look like at the level of algebra?

---

[1] (for Coxeter–Artin groups see (Ram's translation of) Brieskorn-Saito)

[2] quite interesting. See Martin-Saleur 93

The partition algebra and the Brauer algebra

For Q-state models the overarching algebra is the partition algebra. The partition algebra P_n has a basis of partitions of 2 rows of n vertices.

- P_n gives a representation of GTLA.
- consider also the Brauer subalgebra (pair partitions).

3.3 Analysis of results I: generalities and very low rank

We have seen quite generally that in the physical temperature region the limit free energy density is $\ln \lambda_0$ where λ_0 is the largest magnitude eigenvalue of the transfer matrix. What becomes of this when we look in the complex $x = \exp(\beta)$ plane, and in particular in our 'critical' neighbourhood of the physical region? (As defined in Section 1.1.3.)

To get a bit more out of the 1d Ising model here consider other boundary conditions. For example, for the A_N graph but with end spin states fixed

$$Z'(A_N) = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} x & 1 \\ 1 & x \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}(\lambda_1^N + \lambda_2^N)$$

while

$$Z(\hat{A}_N) = \mathrm{Tr} \begin{pmatrix} x & 1 \\ 1 & x \end{pmatrix}^N = \lambda_1^N + \lambda_2^N = (x+1)^N + (x-1)^N$$

There are a number of ways that we can recast this simple expression to help think about what might happen in general for large N.

Firstly, we can rewrite

$$(x+1)^N + (x-1)^N = (x-1)^N((\frac{x+1}{x-1})^N + 1)$$

Ignoring the first factor we have

$$Z \sim Y^N + 1 = Y^{N/2}(Y^{N/2} + Y^{-N/2})$$

(where $Y = \frac{x+1}{x-1}$), that is, the remaining zeros are distributed evenly around a circle. It is the same circle for any N, but the line density increases with N. Setting $Y = \exp(\beta')$

$$f = \frac{1}{N} \ln Z = \frac{1}{N}(\ln(2) + \frac{N}{2} \ln Y + \ln(\cosh(\frac{N}{2}\beta')))$$

so

$$U = -\frac{\partial \frac{1}{N}\ln Z}{\partial \beta'} = -1 - \tanh(\frac{N}{2}\beta')$$

That is, the internal energy changes fast (at $\beta' = 0$) for large N.

As we have already seen, the physics is dominated by the zeros close to the real line, so we can approximate

$$\lim_{N\to\infty} f \sim \beta'/2 + \frac{1}{2\pi}\int_{-\infty}^{\infty} a(y)\ln(\beta' + iy)dy$$

where $a(y) = 1$ (in our case) is the line density of zeros. Thus

$$U \sim \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{a(y)dy}{y - i\beta'}$$

The integrand has a simple pole at $y = i\beta$, so the integral changes by $2\pi a(0)$ as β' changes sign. In other words, if the limit line density $a(0) \neq 0$ the internal energy changes discontinuously at this point — a *first order phase transition.*

In practice, in more complicated systems, we can get

$$a(y) \sim |y|^{1-p} \qquad (0 \leq p \leq 1)$$

but we will return to this shortly.

Notice in our 2×2 transfer matrix example that the distribution of zeros corresponds to the locus of points where the largest eigenvalue is actually degenerate (with the other eigenvalue, regarded as an analytic function of β). In fact a large class of models have a transfer matrix reducible to a 2×2 polynomial matrix T'. As before

$$\lim_{N\to\infty} \frac{\ln Z}{N} = \lim_{N\to\infty} \ln(\lambda_+^N + \lambda_-^N) \overset{*}{=} \ln \lambda_+$$

where * means on the real axis. What happens to the zeros this time?

Consider the general identity

$$C^N + D^N = \prod_{n=-\frac{N-1}{2}}^{\frac{N-1}{2}} (C + \exp(\frac{2\pi i n}{N})D) = \prod_{n=1/2}^{\frac{N-1}{2}} (C^2 + D^2 + 2\cos(2\pi n/N)CD)$$

(the explicit limits are for the case N even — the reader will easily compute the odd case). Using this we can rewrite

$$\lim_{N\to\infty} \frac{\ln Z}{N} = \lim_{N\to\infty} \ln(\lambda_+^N + \lambda_-^N) = \frac{1}{2\pi}\int_0^{\pi} \ln(2(A^2 + B) + 2\cos y(A^2 - B))dy$$

where $A^2 - B = CD = \lambda_+\lambda_-$ and $2(A^2+B) = C^2 + D^2 = \lambda_+^2 + \lambda_-^2$, that is

$$\lambda_\pm = A \pm \sqrt{B}$$

Since T' is polynomial, so are A and B, and hence the limit is (the log of) an infinite product of polynomials. One readily confirms that the zeros of this infinite product are the loci

$$|\lambda_+| = |\lambda_-|$$

and the endpoints of these loci (if any) are the points where $\lambda_+ = \lambda_-$, i.e. at roots of the polynomial B (if B nonvanishing).

(**3.3.1**) REMARKS: This analysis is essentially taken from [9, Ch.11].

3.4 The 2D Ising model: exact solution

Recall from §1.2.1 that a lattice is an embedding (i.e. a positional but not orientational fixing) of a set of spins in an underlying physical space (usually in a regular array). Then a lattice model is a model of the bulk behaviour of such a system of many interacting lattice spins, determined by a spin interaction Hamiltonian.

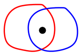
Recall the Potts model Hamiltonian (1.4). The Ising model is the two-state Potts model (up to some trivial Hamiltonian rescalings). It will be convenient to use the equivalent 'Ising form' of the Potts Hamiltonian here. Thus we have the collection of partition functions of form

$$Z = \sum_\sigma \exp(\beta \sum_{ij} (2\delta_{\sigma_i,\sigma_j} - 1))$$

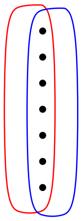
where $\sum_{ij}$ is the sum over pairs of nearest neighbour sites in the lattice. In practice one focusses on a lattice or collection of lattices determined by the embedding space. In 2D this collection of lattices is (at least locally) the $n \times m$ square grids, with n, m large.

Our strategy in computing Z is to determine a transfer matrix $\mathcal{T}$ (acting on the space of states of an n-site layer of the lattice), such that $Z = \langle \mathcal{T}^m \rangle$, and then to compute by finding a basis for the state space in which $\mathcal{T}$ is diagonal.

Let us briefly recall the partition vector/transfer matrix formalism. Fixing the local Hamiltonian one has a tensor (vector or matrix) for each graph G and collection $\{V_1, V_2, ..., V_r\}$ of (possibly intersecting) subsets of the set of vertices. The $i_1, i_2, ..., i_r$ entry of the tensor is the partition function for G with the spins in V_i fixed in state i. Our pictorial notation for this is to draw the graph, together with a loop around each vertex set V_i. For example, in Q-state models one has

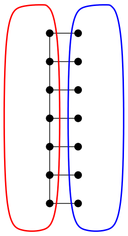


$$= 1_Q$$



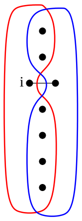
$$= 1_Q^{\otimes n}$$

$$\mathcal{T} =$$



where the last picture is the $\mathcal{T}$ appropriate for a layer in a square lattice.

In the Ising form the local transfer matrix T_i (from (3.2)) for an n-site wide lattice is



$$= T_i = \begin{pmatrix} x & x^{-1} \\ x^{-1} & x \end{pmatrix}_i = x 1_{2^n} + x^{-1}\sigma_i^x = x(1_{2^n} + x^{-2}\sigma_i^x)$$

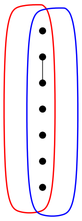
(this means T_i acts non-trivially on the ith factor in the layer configuration space, and acts trivially on all the other $n-1$ factors). Note that for any scalar θ

$$e^{\theta\sigma^x} = \cosh\theta 1 + \sinh\theta\sigma^x = \cosh\theta(1 + \tanh\theta\sigma^x)$$

so if we choose θ so that $\tanh\theta = x^{-2}$ we get

$$T_i = \frac{x}{\cosh(\theta)}\, e^{\theta\sigma_i^x} \;=\; (\cosh(\theta)\sinh(\theta))^{-1/2} e^{\theta\sigma_i^x} \;=\; \sqrt{2\sinh(2\beta)}\, e^{\theta\sigma_i^x} \qquad (3.5)$$

Meanwhile the local transfer matrix T_{ij} is



$$= T_{ij} = \begin{pmatrix} x &&& \\ & x^{-1} && \\ && x^{-1} & \\ &&& x \end{pmatrix}_{ij}$$

(acting on the adjacent factors i, j), which can be written

$$T_{ij} = \exp\left(\beta\begin{pmatrix} 1 &&& \\ & -1 && \\ && -1 & \\ &&& 1 \end{pmatrix}\right) = \exp\left(\beta\begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}\right)$$

Bruria Kaufman's (1949) idea is as follows. One notes that $V_1 = \prod T_i$ and $V_2 = \prod T_{ij}$ can both be equated to certain *spin representations* of rotations. (These are representations on tensor space of dimension 2^n.) We can then use an abstract relation to the eigenvalues of a smaller more manageable representation of the same rotations — the ordinary rotation matrices of dimension $2n$. Recall that these are generated by the simple plane rotations

$$w_{i\,i+1}(\theta) = 1_{i-1} \oplus \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \oplus 1_{2d-i-1} \tag{3.6}$$

where the mixing occurs in the $i, i+1$-positions.

To understand the spin representations it is convenient to introduce *Clifford algebras.*

(**3.4.1**) A set $\{\Gamma_a\}_{a=1,\dots,2n}$ of $2^n \times 2^n$ matrices such that $\Gamma_a^2 = 1$ and

$$\Gamma_a\Gamma_b + \Gamma_b\Gamma_a = 0 \qquad (a \neq b)$$

are said to form a *Clifford algebra.*

(**3.4.2**) For example, with σ_i^x the usual Pauli matrix action on tensor space:

$$\Gamma^\bullet_{2i-1} := \left(\prod_{j=1}^{i-1} \sigma_j^x\right) \sigma_i^z \qquad\qquad \Gamma^\bullet_{2i} := \left(\prod_{j=1}^{i-1} \sigma_j^x\right) \sigma_i^y$$

obey these relations.

In this case note that

$$\Gamma^\bullet_{2i}\Gamma^\bullet_{2i-1} = \sigma_i^y\sigma_i^z = i\sigma_i^x \qquad\qquad \Gamma^\bullet_{2i+1}\Gamma^\bullet_{2i} = \sigma_i^x\sigma_{i+1}^z\sigma_i^y = i\sigma_i^z\sigma_{i+1}^z$$

Thus from (3.5) *et seq*

$$V_1 = \prod_i T_i \;=\; \kappa^n \prod_{i=1}^{n} e^{-i\theta\Gamma^\bullet_{2i}\Gamma^\bullet_{2i-1}} \qquad\qquad V_2 = \prod_i T_{i\,i+1} \;=\; e^{\beta\sigma_n^z\sigma_1^z} \prod_{i=1}^{n-1} e^{-i\beta\Gamma^\bullet_{2i+1}\Gamma^\bullet_{2i}}$$

where $\kappa = \sqrt{2\sinh(2\beta)}$, and at the last we have applied periodic boundary conditions.

[3]

---

[3] (**3.4.3**) Note: (i) If $\{\Gamma_a\}_a$ is a Clifford algebra, then so is $\{S\Gamma_aS^{-1}\}_a$ for any invertible matrix $S \in \mathrm{End}(\mathbb{C}^{2^n})$;
(ii) the matrices Γ'_a obtained from the $\Gamma^\bullet_a$s by swapping the roles of σ^x and σ^z are a Clifford

(**3.4.4**) Fixing any Clifford algebra $\{\Gamma_a\}_a$ we define

$$S(w_{ab}(\theta)) = \cos\frac{\theta}{2}1 - \sin\frac{\theta}{2}\Gamma_a\Gamma_b = \exp(\frac{-1}{2}\theta\Gamma_a\Gamma_b)$$

Example: One can easily find a Clifford algebra $\{\Gamma^\circ_a\}_a$ such that

$$\exp(\frac{-1}{2}\theta\Gamma^\circ_1\Gamma^\circ_2) = \exp(\frac{-1}{2}\theta i\sigma^z_1) = \begin{pmatrix} e^{-i\theta/2} & \\ & e^{i\theta/2} \end{pmatrix} \otimes 1_2 \otimes 1_2 ... \tag{3.7}$$

One can check that these matrices obey

$$S(w_{ab}(\theta))\Gamma_a S^{-1}(w_{ab}(\theta)) = \cos\theta\Gamma_a + \sin\theta\Gamma_b$$

and so on. In other words conjugation by $S(w_{ab}(\theta))$ enacts the rotation $w_{ab}(\theta)$ on the space of Γ-matrices (not to be confused with the space on which the Γ-matrices act).

It follows that conjugation by $S(w_{ab}(\theta))S(w_{cd}(\theta'))$ realises the rotation $w_{ab}(\theta)w_{cd}(\theta')$. From this we have a kind of realisation of the group of rotations in $2n$-dimensions. (This is not quite a representation, since it is a double-cover, but this need not concern us.)

(**3.4.5**) Note from (3.6) that the spectrum of $w_{ab}(\theta)$ is $e^{i\theta}, e^{-i\theta}$ (and possibly some 1s). (Indeed any element in $SO(3)$ can be expressed as a simple rotation about some, not in general coordinate, axis; and hence has eigenvalues of the same form.) Meanwhile, noting (3.7), the spectrum of $S(w_{ab}(\theta))$ is $e^{\frac{i}{2}(\pm\theta)}$ (2^{n-1} copies of each).

Further, if $w = \prod_i w_{a_i b_i}(\theta_i)$ with all the $\{a_i, b_i\}$ distinct; and $S(w) = \prod_i S(w_{a_i b_i}(\theta_i))$, then the $2n$ eigenvalues of w are $\{e^{\pm i\theta_j}\}_j$; and, since the factors in $S(w)$ commute, the 2^n eigenvalue of $S(w)$ are

$$\text{Spectrum}(S(w)) \;=\; \{e^{\frac{i}{2}\sum_{j=1}^{n}\pm\theta_j}\}. \tag{3.8}$$

---

algebra.
(iii) arbitrarily permute the numbering of the Γ_as in any Clifford algebra, and the Clifford relations will still be obeyed.
(iv) Also if $\{\Gamma_a\}$ is any Clifford algebra then $\Gamma' = r\Gamma_1 + s\Gamma_2$ obeys

$$\Gamma_a\Gamma' + \Gamma'\Gamma_a = \Gamma_a(r\Gamma_1 + s\Gamma_2) + (r\Gamma_1 + s\Gamma_2)\Gamma_a = r(\Gamma_a\Gamma_1 + \Gamma_1\Gamma_a) + s(\Gamma_a\Gamma_2 + \Gamma_2\Gamma_a) = 0$$

for all $a > 2$, for any r, s; and

$$(c\Gamma_1 + s\Gamma_2)(s\Gamma_1 - c\Gamma_2) + (s\Gamma_1 - c\Gamma_2)(c\Gamma_1 + s\Gamma_2) = 2(cs - sc) = 0$$

and so on. Following these calculations one eventually checks that Γ_1, Γ_2 can be replaced by the indicated linear combinations, so long as $c = \cos\theta$ and $s = \sin\theta$ for some θ. Evidently one can compose such transformations, so we have an action of the $2n$ dimensional rotation group transforming between realisations of the Clifford relations.

(Later we shall generalise this correspondence to elements of $SO(2n)$ that are products of commuting rotations about arbitrary sets of orthogonal axes, not just the nominal coordinate axes.)

(**3.4.6**) In these terms, writing $S^\bullet(w)$ for $S(w)$ in the $\Gamma^\bullet$ case, we have

$$V_1 = \kappa^n \prod_{i=1}^{n} S^\bullet(w_{2i\;2i-1}(2i\theta))$$

$$V_2 = \chi \prod_{i=1}^{n-1} S^\bullet(w_{2i+1\;2i}(2i\beta))$$

where χ is the periodic boundary operator. For a suitable treatment of the boundary (not quite simple periodic, but close enough) we have $\chi = S^\bullet(w_{1\;2n}(2i\beta))$.

(**3.4.7**) The idea now is to replace $\mathcal{T} = V_1V_2$ with the corresponding product of rotation matrices $W = W_1W_2$. We then find the eigenvalues of this product W. Since every rotation group element can be expressed as a product of commuting rotations (not necessarily respecting the initial axes) these eigenvalues will define a set of rotation angles $\{\theta_i\}_i$ as above. We claim that this set then give the spectrum of $\mathcal{T}$, as in (3.8).

We have (in a representative small example, $n = 4$), with $c = \cos(2i\theta)$, $s = \sin(2i\theta)$, and so on,

$$W_1 = w_{12}(2i\theta)w_{34}(2i\theta)w_{56}(2i\theta)w_{78}(2i\theta) = \begin{pmatrix} c & s & & & & & & \\ -s & c & & & & & & \\ & & c & s & & & & \\ & & -s & c & & & & \\ & & & & c & s & & \\ & & & & -s & c & & \\ & & & & & & c & s \\ & & & & & & -s & c \end{pmatrix}$$

$$W_2 = \begin{pmatrix} c' & & & & & & & -s' \\ & c' & s' & & & & & \\ & -s' & c' & & & & & \\ & & & c' & s' & & & \\ & & & -s' & c' & & & \\ & & & & & c' & s' & \\ & & & & & -s' & c' & \\ s' & & & & & & & c' \end{pmatrix}$$

Note that W_1, W_2 are both fixed by the square of the matrix position shift operator (and that this is true for general n). Accordingly we can fourier transform to block diagonalise W. That is, we make a change of basis as follows. Define $f_z = (1, z, z^2, z^3, ..., z^{n-1})^t$, $v_z = (1,0)^t \otimes f_z$ and $v'_z = (0,1)^t \otimes f_z$. With $z^n = 1$ we have

$$W_1 \begin{pmatrix} v_z \\ v'_z \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} v_z \\ v'_z \end{pmatrix}$$

$$W_2 \begin{pmatrix} v_z \\ v'_z \end{pmatrix} = \begin{pmatrix} c' & zs' \\ -z^{-1}s' & c' \end{pmatrix} \begin{pmatrix} v_z \\ v'_z \end{pmatrix}$$

$$W \begin{pmatrix} v_z \\ v'_z \end{pmatrix} = \begin{pmatrix} cc' + z^{-1}ss' & zcs' - sc' \\ sc' - z^{-1}cs' & cc' + zss' \end{pmatrix} \begin{pmatrix} v_z \\ v'_z \end{pmatrix}$$

Noting that the determinant of the image matrix $W^{(z)}$ on the right is 1 we write the eigenvalues as $\lambda_\pm = e^{\pm l_z}$. Thus

$$e^{l_z} + e^{-l_z} = \text{Trace}(W^{(z)}) = 2\left(cc' + \frac{z+z^{-1}}{2} ss'\right)$$

That is,

$$\cosh(l_z) = \left(\cosh(2\theta)\cosh(2\beta) - \frac{z+z^{-1}}{2}\sinh(2\theta)\sinh(2\beta)\right)$$

Since z is any solution to $z^n = 1$, the complete set of l_zs is obtained from $z = e^{2\pi ik/n}$ with $k = 1, .., n$. One then finds that each l_z is positive for physical parameters. It follows that the largest among the eigenvalues for $\mathcal{T}$ that this gives:

$$\lambda = \exp\left(\frac{1}{2}\sum_{k=1}^{n} \pm l_{e^{2\pi ik/n}}\right)$$

is the case

$$\lambda_0 = \exp\left(\frac{1}{2}\sum_{k=1}^{n} l_{e^{2\pi ik/n}}\right)$$

Recall that this is for an n-site wide lattice. If the lattice is m sites long then

$$Z \sim \lambda_0^m \left(1 + \sum_{i\neq 0}(\frac{\lambda_i}{\lambda_0})^m\right) \sim \lambda_0^m,$$

or more usefully, the free energy is $f = \frac{1}{nm}\ln Z \sim (1/n)\ln\lambda_0$, with this approximation getting better as m gets bigger and becoming an equality in the large m limit.

Now we follow [9, §4.1] for the rest of the analysis. Using that

$$\sinh(2\beta) = \sinh(2\theta)^{-1} \qquad \text{and} \qquad \coth(2\beta) = \cosh(2\theta)$$

we have

$$\cosh(l_z) = \left(\coth(2\beta)\cosh(2\beta) - \frac{z+z^{-1}}{2} \right)$$

which it is convenient to express via the integral representation[4]

$$l_z = \frac{1}{\pi} \int_0^\pi dy \ln(2(\coth(2\beta)\cosh(2\beta) - \frac{z+z^{-1}}{2}) - 2\cos(y))$$

giving

$$\ln \lambda_0 = \frac{1}{2} \sum_{k=1}^{n} \frac{1}{\pi} \int_0^\pi dy \ln(2\coth(2\beta)\cosh(2\beta) - 2\cos(2\pi k/n) - 2\cos(y))$$

Note how the difference in the way we have treated m and n is manifested here. We are looking at the free energy in the infinite-m limit (hence the integral), but with n still finite. We can rigorously bring the treatment of the two directions onto the same footing by taking the large n limit (which will covert the sum to a matching integral); or we can roughly discretise the integral to a sum over m terms, using the sum over n terms as a guide:

$$\int_0^\pi f(y)dy \sim \frac{\pi}{M} \sum_{r=1}^{M} f(\pi r/M)$$

(any $M >> 1$).

In the latter case we get (see e.g. [**?**])

$$Z_{mn} = \prod_{r=1}^{m} \prod_{s=1}^{n} \left\{ 1 - \frac{K}{2}(\cos(2\pi r/m) + \cos(2\pi s/n)) \right\}$$

where

$$K = 4\frac{\exp(-2\beta)(1-\exp(-4\beta)))}{(1+\exp(-4\beta))^2}$$

---

[4] recall $\cosh^{-1} x = \ln(x + \sqrt{x^2-1}) = \ln x + \ln(1+\sqrt{1+x^{-2}})$ and

$$\pi \ln(1+\sqrt{1-t^2}) = \int_0^\pi \ln(1+t\cos w)dw$$

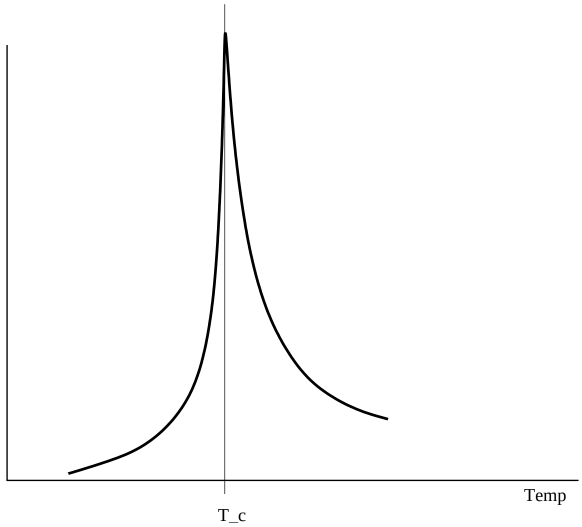


Figure 3.5: Sketch plot of specific heat versus temperature for the 2D Ising model.

Note that each factor in Z_{nm} is invariant under $x \to -x$, and

$$\frac{1}{x^2} \to \tanh\beta = \frac{e^\beta - e^{-\beta}}{e^\beta + e^{-\beta}} = (x^2-1)/(x^2+1)$$

The fixed point of this transformation is given by

$$-(x^2+1) + (x^2-1)x^2 = x^4 - 2x^2 - 1 = (x^2 - (-\sqrt{2}+1))(x^2 - (\sqrt{2}+1)) = 0$$

$$x^2 = \exp(2\beta) = 1 + \sqrt{2}$$

Note that this also determines the ferromagnetic critical point.

A very rough sketch of the second log-derivative (the specific heat) versus $T = 1/\beta$ is given in Figure 3.5. The divergence is logarithmic: close to T_c we have

$$S(T) \sim -\ln|T_c - T|$$

Chapter 4

Spin chains

This Chapter is largely unfinished, but for now contains some useful notation.

4.1 Anisotropic limit

Here we consider a ‘continuous time’ approximation to the Ising model. (Here ‘time’ is a misnomer for the layering direction in the transfer matrix formalism.) In this case one takes an infinite lattice limit in the layering direction first, but adjusts the couplings anisotropically to compensate. There is then a non-trivial simplification of the general problem arising. This has the flavour of a quantum Hamiltonian limit of the transfer matrix (which will lead conveniently to the later consideration of such Hamiltonians *per se*).

— TO DO! —

4.2 XXZ spin chain

In some formalisms, such as quantum theory, the energy function E (the ‘classical Hamiltonian’) is replaced by an operator on the space of states. In such cases, the first mathematical challenge is to determine the eigenvalues and eigenvectors of this Hamiltonian operator.

We describe the Heisenberg chain Hamiltonian and a number of variations thereof. The eigenvalue problem is hard for these systems, for anything other than very small chain length. In the absence of a complete solution, Physical interest may justify trying to solve just for the lowest (respectively highest) eigenvalue, its degeneracy, and the gap to the next lowest.

4.2.1 Notations and conventions

Define

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad i\sigma^y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

so that

$$\sigma^+ = \frac{1}{2}(\sigma^x + i\sigma^y) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Our convention in the Kronecker product $A \otimes B$ is to order so that

$$\sigma^+ \otimes \sigma^- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \sigma^z \otimes 1_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Now fix L and define operators σ^x_i (and so on) on $(\mathbb{C}^2)^{\otimes L}$ so that:

$$\sigma^x_i \sigma^x_{i+1} = 1_2 \otimes \cdots \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \cdots = 1_2 \otimes \cdots \otimes \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \otimes \cdots$$

$$\sigma^z_i \sigma^z_{i+1} = 1_2 \otimes \cdots \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdots = 1_2 \otimes \cdots \otimes \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdots$$

4.2.2 Hamiltonia

$$u^{XYZ}_{ij}(k_x, k_y, k_z) := \frac{1}{2}(k_x \sigma^x_i \sigma^x_j + k_y \sigma^y_i \sigma^y_j + k_z \sigma^z_i \sigma^z_j)$$

$$u^{XYZ}_i(k_x, k_y, k_z) := u^{XYZ}_{i\,i+1}(k_x, k_y, k_z)$$

(by u_L we shall intend the periodic closure $u_{L,1}$)

$$u^{XXZ}_i = u^{XYZ}_i(1, 1, -\frac{q + q^{-1}}{2}) = \begin{pmatrix} -\frac{\delta}{4} & 0 & 0 & 0 \\ 0 & \frac{\delta}{4} & 1 & 0 \\ 0 & 1 & \frac{\delta}{4} & 0 \\ 0 & 0 & 0 & -\frac{\delta}{4} \end{pmatrix}_i$$

where $\delta = q + q^{-1} \in \mathbb{C}$; $u_i^{XXX} = u_i^{XYZ}(1,1,1)$; $u_i^{XY} = u_i^{XYZ}(1+a, 1-a, 0)$ ($a \in \mathbb{R}$);

$$u_i = u_i^{XXZ} + \frac{1}{2}\frac{q+q^{-1}}{2} - \frac{1}{2}\frac{q-q^{-1}}{2}(\sigma_i^z - \sigma_{i+1}^z) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & q & 1 & 0 \\ 0 & 1 & q^{-1} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The simple periodic Heisenberg chain (without external field) is the Hamiltonian

$$H^{XXX} = u_L^{XXX} + \sum_{i=1}^{L-1} u_i^{XXX}$$

Generalisations include

$$H_{pp'}^{XXZ} := \sum_{i=1}^{L-1} u_i^{XXZ} + p\sigma_1^z + p'\sigma_L^z \qquad\qquad H^{XXZ} := \sum_i u_i$$

(see e.g. [**?**]). We shall be interested in the spectrum of these matrices. The matrix H^{XXZ} differs from a sum of the u_i^{XXZ} by an additive scalar and boundary terms, so both are forms of XXZ spin chain Hamiltonian.

The anisotropic limit Potts model Hamiltonian is (roughly) in the same universality class as the Potts model. It also takes the same form as H^{XXZ}, except for being in a different representation of the TL algebra. For this reason (among others) one is interested in the spectrum of H^{XXZ}.

(Potts model unsolved.)

4.2.3 Normal Bethe ansatz

We are interested in the spectrum of $H = H_{pp'}^{XXZ}$ (or specifically H^{XXZ}, or indeed any similar case).

We take words in $\{1,2\}$ of length L as a basis for $(\mathbb{C}^2)^L$. For any fixed but arbitrary L, let us write $v_x = v(x_1, ..., x_J)$ for the vector with 2s in the J indicated positions. Clearly the subspaces spanned by vectors with J fixed are invariant under H, so for each J we may look for eigenvectors w obeying

$$Hw = E_w w$$

with

$$w = \sum_{1 \leq x_1 < ... < x_i < x_{i+1} < ... < x_J \leq L} c_x v_x$$

(We call the excess of 1s over 2s, given by $L-2J$, the *charge*. Whether Physical interest resides mainly in low or high charge depends on the details. Here we shall concetrate on what can be said about eigenvalues in general.)

It will be evident that there is only one vector (hence eigenvector) with $J=0$. We have

$$H^{XXZ}\ 111...1 \ = \ 0$$

For $J=1$ we have

$$Hv_1 = (u_1+u_2+...+u_{L-1})v_1 = q^{-1}v_1+v_2$$
$$Hv_i = v_{i-1}+(q+q^{-1})v_i+v_{i+1} \qquad (1<i<L)$$
$$Hv_L = qv_L+v_{L-1}$$

so that

$$Hw = \sum_j c_j Hv_j = E_w \sum_j c_j v_j$$

gives the following. Coefficient of v_1 (then v_i, then v_L):

$$Ec_1 = c_1 q^{-1} + c_2$$
$$Ec_i = c_{i-1}+c_i(q+q^{-1})+c_{i+1} \qquad (1<i<L) \tag{4.1}$$
$$Ec_L = c_L q + c_{L-1}$$

All three forms coincide if we use the boundary conditions

$$c_0 = -qc_1 \qquad\qquad c_{L+1} = -q^{-1}c_L$$

Following Alcaraz et al one tries for a solution of form

$$c_j = A(k)e^{ijk} + A(-k)e^{-ijk}$$

where viable choices for k, and for each k the constant A, are to be determined. Substituting into (4.1) we have

$$E_k - (q+q^{-1})(A(k)e^{ijk}+A(-k)e^{-ijk})$$
$$= A(k)e^{ijk}e^{-ik} + A(-k)e^{-ijk}e^{ik} + A(k)e^{ijk}e^{ik} + A(-k)e^{-ijk}e^{-ik}$$

giving

$$E_k = (q+q^{-1}) + 2\cos(k)$$

CHECK THIS!

It remains to impose the boundary conditions, which, if they can be satisfied, will determine the possible values of k. The c_0 condition gives

$$A(0)+A(0) = -q(A(1)e^{ik}+A(-1)e^{-ik})$$

4.2.4 Abstract Bethe ansatz

4.2.5 Gram matrices

Acknowledgements

Thanks to Yogi Valani for many useful discussions.
Thanks to my coorganisers of Leeds SMDG, Jeanne Scott and Sathish Sukumaran for all their considerable efforts to make it work.

.1 Appendix: What is Physics?

Statistical Mechanics is part of Physics. *Physics* might be characterised, in the large, as the scientific exercise (as opposed to involuntary reflex) of modelling of the observable physical world. That is, the representation of part of the physical world by something ‘simpler’, which nonetheless captures some of the physical world’s humanistically essential features.
There are various phases to this exercise, such as:
(i) deciding which toy is the model;
(ii) working out what the model itself does; and
(iii) interpreting this behaviour as a prediction for the physical world.

The simple toys at our disposal include *real* toys (scale models of bridges and so on), and systems of equations in mathematics. [1] In particular Scientists have had notable success summarizing large amounts of observational data from the physical world with certain relatively simple mathematical models. A very successful such model is, reasonably, regarded as close to nature itself; and hence fundamental. Key to this is the expectation that such a model, pushed into an as yet unobserved (but suitably nearby) regime, will correctly predict the result of observations subsequently made there. There has been notable success too in this predictive aspect of Physics, and great technological benefits have accrued.

Bibliography

Some relevant texts are, for Statistical Mechanics: Baxter[2], Domb and Green[3], Levy[8], Martin[9], and Wannier[16];

---

[1] Note that these toys either exist themselves in the physical world, or are abstractions formulated by creatures living in the physical world (and whose thought processes are tied to making sense of this experience). This is of course unavoidable, but singular. Other problems in representation theory are not constrained to have all representations built inside the thing being represented.

for generalities: Abramowitz and Stegun[1], Feynman[6].

Bibliography

[1] Milton Abramowitz and Irene A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Dover, New York, 1964.

[2] R J Baxter, *Exactly solved models in statistical mechanics*, Academic Press, New York, 1982.

[3] C Domb and M S Green (eds.), *Phase transitions and critical phenomena I*, Academic Press, London, New York, 1972.

[4] M B Einhorn, R Savit, and E Rabinovici, *A physical picture for the phase transition in Z_N symmetric models*, Nucl Phys B **170** **[FS1]** (1980), 16.

[5] S Elitzur, R B Pearson, and J Shigemitsu, *The phase structure of discrete abelian spin and gauge systems*, Phys Rev D **19** (1979), 3698.

[6] Richard P Feynman, *The Feynman lectures on physics: Volume 3*, Addison-Wesley, Boston, 1963.

[7] P J Hilton and S Wylie, *Homology theory*, Cambridge, 1965.

[8] R A Levy, *Principles of Solid State Physics*, Academic Press, New York, 1968.

[9] P P Martin, *Potts models and related problems in statistical mechanics*, World Scientific, Singapore, 1991.

[10] P P Martin and A Elgamal, *Ramified partition algebras*, Math. Z. **246** (2004), 473–500, (math.RT/0206148).

[11] P P Martin and Y Valani, *Non-integrable Potts models revisited*, In preparation (2008).

[12] G Pólya and R C Read, *Combinatorial enumeration of groups, graphs, and chemical compounds*, Springer-Verlag, New York, 1987.

[13] E H Spanier, *Algebraic topology*, McGraw–Hill, New York, 1966.

[14] R M Thompson and R T Downs, *Systematic generation of all nonequivalent closest-packed stacking sequences of length n using group theory*, Acta Crystallographica **B57** (2001), 766–771.

[15] Y Valani, *Temperature zeros of the Potts model partition function revisited*, Ph.D. thesis, City University, In preparation.

[16] G H Wannier, *Statistical physics*, Wiley, 1966.