

3 The renormalization group for Ising spins

I couldn't find a funny quote, so instead I offer the following gross understatement.

The actual process of explicitly constructing a useful renormalization group is not trivial.

Michael Fisher

3.1 The one-dimensional Ising model

The one-dimensional Ising model with nearest neighbor interactions is the only Ising type system for which one can do an exact, simple renormalization group calculation. Of course, this model is not of a great deal of interest since it has no phase transition. Nonetheless, it is a good place to start.

For reasons that will be apparent in a moment, we use a chain which runs from $-N$ to N with N equal to a power of 2, namely 2^n . The partition function is

$$Z = \sum_{\sigma} \exp\left(\sum_{i=-N+1}^N \sigma_{i-1}\sigma_i\right) \quad (1)$$

where the sum over σ is shorthand for the sum over all the spin variables $\sigma_{-N}, \sigma_{-N+1}, \dots, \sigma_N$. Rather than sum over all the spins at once, we will implement the renormalization group philosophy of summing over the modes associated with the short length scales first. Several implementations of this philosophy have been developed. We will begin with one known as decimation. (We will see later that decimation is fatally flawed in more than one dimension.) Decimation is simple. We fix the spins σ_i with i even and only sum over the spins σ_i with i odd. The resulting quantity will depend on the spins $\sigma_{-2^n}, \sigma_{-2^n+2}, \dots, \sigma_{-2}, \sigma_0, \sigma_2, \dots, \sigma_{2^n}$ and is obviously positive, so we can write it as

$$\exp(-H_1(\dots, \sigma_{-2}, \sigma_0, \sigma_2, \dots)) = \sum_{\dots, \sigma_{-3}, \sigma_{-1}, \sigma_1, \sigma_3, \dots} \exp\left(\sum_{i=-N+1}^N \sigma_{i-1}\sigma_i\right) \quad (2)$$

The hope is that the sum in (2) will be easier to compute than the sum in (1). (The reason for the subscript 1 on H_1 will become apparent in a moment.)

Given $H_1(\cdots, \sigma_{-2}, \sigma_0, \sigma_2, \cdots)$ we can still recover the full partition function by the equation

$$Z = \sum_{\cdots, \sigma_{-2}, \sigma_0, \sigma_2, \cdots} \exp(-H_1(\cdots, \sigma_{-2}, \sigma_0, \sigma_2, \cdots))$$

However, this sum will be as difficult to compute as the original sum (1). So we proceed instead by iterating the decimation procedure. We sum over half of the remaining spins.

$$\exp(-H_2(\cdots, \sigma_{-8}, \sigma_{-4}, \sigma_0, \sigma_4, \sigma_8, \cdots)) = \sum_{\cdots, \sigma_{-6}, \sigma_{-2}, \sigma_2, \sigma_6, \cdots} \exp(-H_1(\cdots, \sigma_{-2}, \sigma_0, \sigma_2, \cdots)) \quad (3)$$

We then sum over half of the remaining spins to define $H_3(\cdots, \sigma_{-16}, \sigma_{-8}, \sigma_0, \sigma_8, \sigma_{16}, \cdots)$, and so on.

At this point it may seem that we just have a big mess, so some philosophical comments are in order. When one is near a critical point it is very difficult to compute the partition function numerically or by various expansion techniques. This is because of the presence of many length scales in the system. Sums like (2) and (3) should be much easier to compute than the full partition function since they hopefully only involve summing over modes associated with short length scales. Of course in the one-dimensional case everything may be computed. The computation of the “renormalized” Hamiltonian H_1 is especially easy. In principle H_1 could be a complicated function of the spins $\cdots, \sigma_{-2}, \sigma_0, \sigma_2, \cdots$. In our one dimensional model it has a very simple form:

$$-H_1(\cdots, \sigma_{-2}, \sigma_0, \sigma_2, \cdots) = \sum_{i=-2^{n-1}+1}^{2^{n-1}} [c_1 + \beta_1 \sigma_{2(i-1)} \sigma_{2i}] \quad (4)$$

The reason for this simple result is that the sum in (2) factors:

$$\begin{aligned} & \exp(-H_1(\cdots, \sigma_{-2}, \sigma_0, \sigma_2, \cdots)) \\ &= \sum_{\cdots, \sigma_{-3}, \sigma_{-1}, \sigma_1, \sigma_3, \cdots} \dots e^{\beta(\sigma_{-2}\sigma_{-1} + \sigma_{-1}\sigma_0)} e^{\beta(\sigma_0\sigma_1 + \sigma_1\sigma_2)} e^{\beta(\sigma_2\sigma_3 + \sigma_3\sigma_4)} \dots \\ &= \dots \sum_{\sigma_{-1}} e^{\beta(\sigma_{-2}\sigma_{-1} + \sigma_{-1}\sigma_0)} \sum_{\sigma_1} e^{\beta(\sigma_0\sigma_1 + \sigma_1\sigma_2)} \sum_{\sigma_3} e^{\beta(\sigma_2\sigma_3 + \sigma_3\sigma_4)} \dots \end{aligned}$$

It is a simple exercise to show that

$$\sum_{\sigma_1} e^{\beta(\sigma_0\sigma_1 + \sigma_1\sigma_2)} = \exp[h(\beta) + g(\beta)\sigma_0\sigma_2]$$

where

$$h(\beta) = \frac{1}{2} \ln[4 \cosh(2\beta)] \quad (5)$$

$$g(\beta) = \frac{1}{2} \ln[\cosh(2\beta)] \quad (6)$$

Thus H_1 is given by (4) with $c_1 = h(\beta)$ and $\beta_1 = g(\beta)$. Note that the renormalized Hamiltonian H_1 has the same form as the original Hamiltonian, i.e., it only contains nearest neighbor couplings. (This is very special to decimation in one dimension.) Of course the nearest neighbor pair of sites in H_1 are actually two lattice spacing apart. The calculation of H_2 is exactly like that of H_1 . H_2 again contains only a nearest neighbor coupling β_2 and a constant term c_2 given by

$$\begin{aligned} c_2 &= 2c_1 + h(\beta_1) = 2h(\beta) + h(g(\beta)) \\ \beta_2 &= g(\beta_1) = g(g(\beta)) \end{aligned}$$

The factor of 2 in the equation for c_2 comes about as follows. In H_1 there is a c_1 for each nearest neighbor bond. This is a constant term that does not depend on any spins, so it is unchanged by the next decimation transformation. However, in H_2 the number of nearest neighbor bonds is reduced by a factor of 2. So the contribution of c_1 to the constant term per nearest neighbor bond in H_2 is $2c_1$. Another application of the decimation transformation yields yet another nearest neighbor Hamiltonian H_3 with

$$\begin{aligned} c_3 &= 2[2h(\beta) + h(g(\beta))] + h(\beta_2) = 4h(\beta) + 2h(g(\beta)) + h(g(g(\beta))) \\ \beta_3 &= g(\beta_2) = g(g(g(\beta))) \end{aligned}$$

After k transformations we have

$$c_k = \sum_{i=0}^{k-1} 2^{k-1-i} h(g^i(\beta)) \quad (7)$$

$$\beta_k = g^k(\beta) \quad (8)$$

where g^i denotes g composed with itself a bunch of times, not g raised to the power i .

What can we learn from the exact renormalization group transformation that we have derived? Suppose we know that for small β the model is in a

high temperature phase, i.e., boundary conditions don't matter in the infinite volume limit. Then we can use the renormalization group transformation to show that the model is in a high temperature phase for all finite β . A graph of the function $g(\beta)$ is shown in figure 1. The only fixed point of $g(\beta)$ is the trivial fixed point $g(0) = 0$. For all $\beta > 0$, $g(\beta)$ is strictly less than β . For large β , $g(\beta) \approx \beta - \frac{1}{2} \ln(2)$. Suppose that we impose + boundary conditions on the system. We could do this in the previous calculations by simply fixing σ_{-2^n} and σ_{2^n} to be -1 . The calculations then go through as before. The renormalized systems all still have + boundary conditions since σ_{-2^n} and σ_{2^n} are never summed over. It is easy to show that the expected value of the spin at the origin in the original system is the same as the expected value of the spin at the origin in the renormalized system. No matter how large β is, after a sufficient number of applications of the renormalization group transformation the renormalized β , namely $g^k(\beta)$, will be small enough that we are in the regime where we know the model is in the high temperature phase. Thus the expected value of the spin at the origin is zero in the infinite volume limit.

In addition to showing that there is no phase transition, we can also use the renormalization group to study how the correlation length depends on β . We saw in the previous chapter that the correlation length is finite for all β , but diverges to ∞ as $\beta \rightarrow \infty$. We will write the correlation length as $\xi(\beta)$ to emphasize its dependence on β . Recall that this length gives the exponential decay rate of the two point correlation function

$$\langle \sigma_0 \sigma_l \rangle \sim \exp(-l/\xi(\beta))$$

as $l \rightarrow \infty$.

First we need to complete the renormalization group by rescaling. After decimation we are left with spins $\dots, \sigma_{-2}, \sigma_0, \sigma_2, \dots$ i.e., the spins on a lattice whose sites are two units apart. To make the renormalized system look as much like the original system as possible we rescale the lattice to make the spacing between the spins that remain equal to one unit. All this really means is that we relabel the spins $\dots, \sigma_{-2}, \sigma_0, \sigma_2, \dots$ by $\dots, s_{-1}, s_0, s_1, \dots$. So we now have a system on the lattice $-2^{n-1}, -2^{n-1}+1, \dots, 2^{n-1}$ with Hamiltonian $H_1(s_{-2^{n-1}}, \dots, s_{-1}, s_0, s_1, \dots, s_{2^{n-1}})$ such that the partition function of this system is exactly equal to the partition function of our original system. The correlation function of this renormalized system is related to the correlation

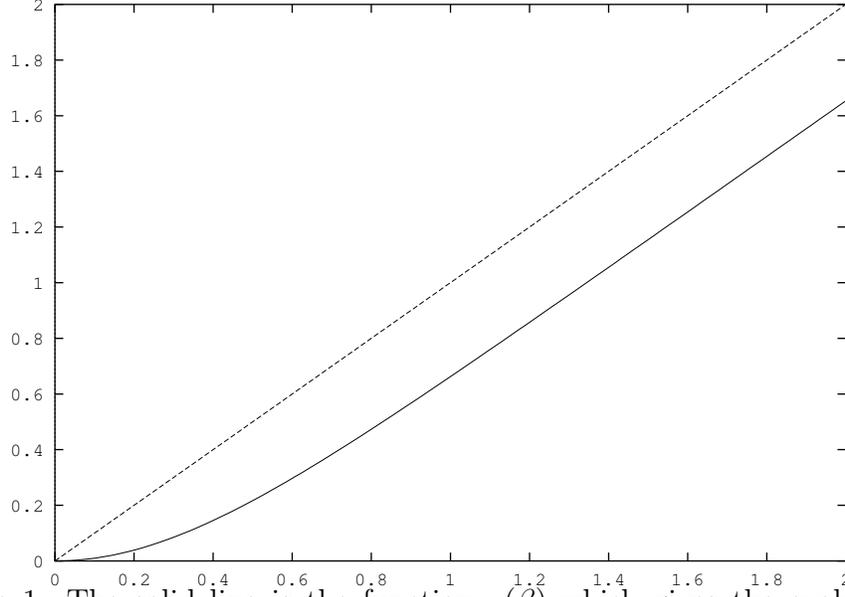


Figure 1: The solid line is the function $g(\beta)$ which gives the evolution of the coupling β under decimation in the one dimensional Ising model. The dashed line is simply the straight line with slope 1. Note that The only fixed point of $g(\beta)$ is the trivial fixed point $\beta = 0$.

function in the original system in a simple way.

$$\begin{aligned}
\langle s_i s_j \rangle_1 &= Z^{-1} \sum_{s_{-2n-1}, \dots, s_{-1}, s_0, s_1, \dots, s_{2n-1}} s_i s_j e^{-H_1(\dots, s_{-1}, s_0, s_1, \dots)} \\
&= Z^{-1} \sum_{\sigma_{-2n}, \dots, \sigma_{-2}, \sigma_0, \sigma_2, \dots, \sigma_{2n}} \sigma_{2i} \sigma_{2j} e^{-H_1(\dots, \sigma_{-2}, \sigma_0, \sigma_2, \dots)} \\
&= Z^{-1} \sum_{\sigma_{-2n}, \dots, \sigma_{-2}, \sigma_0, \sigma_2, \dots, \sigma_{2n}} \sigma_{2i} \sigma_{2j} \sum_{\dots, \sigma_{-3}, \sigma_{-1}, \sigma_1, \sigma_3, \dots} e^{-H(\dots, \sigma_{-1}, \sigma_0, \sigma_1, \dots)} \\
&= \langle \sigma_{2i} \sigma_{2j} \rangle
\end{aligned}$$

If $\langle \sigma_x \sigma_y \rangle \sim e^{-|x-y|/\xi}$, then $\langle s_i s_j \rangle_1 \sim e^{-2|i-j|/\xi}$. So the correlation length in the renormalized system, ξ_1 , is related to the correlation length in the original system, ξ , by $\xi_1 = \xi/2$. The renormalized system is again a nearest neighbor Ising model with coupling $g(\beta)$. Thus $\xi_1 = \xi(g(\beta))$. Hence we have the important result

$$\xi(g(\beta)) = \xi(\beta)/2 \tag{9}$$

Now suppose we start with a very large β . Recall that in this regime each iteration of the renormalization group reduces β by essentially $\frac{1}{2} \ln 2$. Thus the renormalized β will be reduced to something of order 1 after N iterations where N is $\frac{2\beta}{\ln 2}$. By (9) this implies

$$\xi(\beta) = 2^N \xi(g^N(\beta)) = 2^N O(1) = e^{2\beta} O(1)$$

So the renormalization group shows that as $\beta \rightarrow \infty$, the correlation length $\xi(\beta)$ diverges as $e^{2\beta}$.

Exercises:

3.1.1 (this is a fairly straightforward calculation meant to reinforce what we did in this section) Instead of summing over every other spin we could define a modified form of decimation where we sum out two out of three spins. So spins $\sigma_0, \sigma_3, \sigma_6, \dots$ would remain while $\sigma_1, \sigma_2, \sigma_4, \sigma_5, \dots$ are summed out. Show that the renormalization group transformation may still be computed in closed form and find the function $g(\beta)$. Of course the correlation length must still diverge as $e^{2\beta}$ as $\beta \rightarrow \infty$. Show that this new renormalization group transformation also yields this conclusion.

3.1.2 (similar to above) Consider the Ising model on the “ladder”.

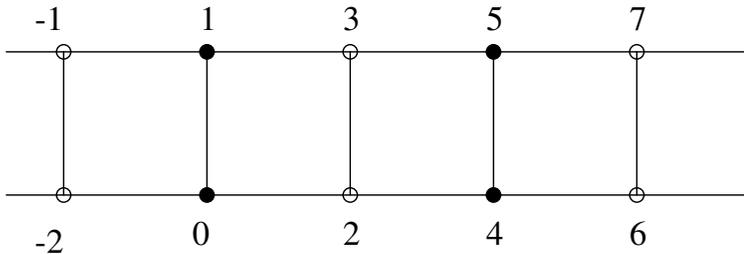


Figure 2:

We can define a decimation transformation for this model as follows. We fix the spins at the sites indicated by a solid circle in the figure and sum over the spins at sites indicated by an open circle. So the new Hamiltonian is a function of spins $\dots, \sigma_0, \sigma_1, \sigma_4, \sigma_5, \dots$. We then rescale the horizontal axis by a factor of 2 to return to the original ladder of spins. This model is essentially a one dimensional model, in particular it has no phase transition. Compute this decimation renormalization group transformation and use it to

show there is no phase transition. The renormalized Hamiltonian will not just contain nearest neighbor interactions. However, the number of different types of terms in the renormalized Hamiltonian remains finite. So the RG map is a map on a finite dimensional space. (Warning: I have not done this computation and it could get complicated.)

3.1.3 (vague and open-ended, but thirty years ago this was publishable research, Fisher and Nelson, the reference should be in the Stellenbosch lectures by Fisher?) The idea is to include a magnetic field in the Hamiltonian and redo everything in this section. The renormalization group transformation may still be explicitly computed. The renormalized Hamiltonian will contain a nearest neighbor interaction and a magnetic field term. So instead of having a map $g(\beta)$ from R to R , there will be a map from R^2 to R^2 which takes the original β, h into the renormalized β, h . Sketch what this map does to the β, h plane.

3.2 The two-dimensional Ising model

We now turn to a more interesting model, the two-dimensional Ising model. As we will see, the decimation transformation generalizes to any number of dimensions. However, there is a simple argument that shows that something must go wrong when one attempts to study the critical point with the decimation transformation. (We will give this argument at the end of this section.) In this section we will look at a rather crude approximation to the decimation transformation which produces qualitatively correct results for the critical phenomena. In the next section we will present some renormalization group transformations which do not suffer from the problem the decimation transformation has and which work quite well numerically. However, from the point of view of a mathematician they are not even known to be well defined.

To define decimation in higher dimensions all we need to do is explain how to interpret even and odd sites in more than one dimension. Let (i_1, i_2, \dots, i_d) be the coordinates of a lattice site in d dimensions. We say it is odd if $i_1 + i_2 + \dots + i_d$ is odd, even if the sum is even. Figure 2 shows the division of the two dimensional lattice into even and odd sites. With this definition the nearest neighbor bonds all connect an even site with an odd site. The decimation transformation is now defined as before

$$\exp[-H_1(\sigma_i : i \text{ even})] = \sum_{\sigma_i : i \text{ odd}} \exp[-H(\sigma)] \quad (10)$$

Of course this equation is meant for a finite volume. For example, in two dimensions we could take a square of sites and make the above definitions with i restricted to the sites in the square.

Just as in one dimension, the first application of the decimation transformation may be carried out exactly because the sum in (10) factors into individual sums over the spin at each odd site. In figure 4 we show a site and its four nearest neighbors. Site "0" is the odd site which is being summed out, so sites 1,2,3,4 are even sites at which the spin is fixed. We need to compute

$$\sum_{\sigma_0} \exp[\beta\sigma_0(\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4)]$$

The result will be a symmetric function of $\sigma_1, \sigma_2, \sigma_3, \sigma_4$. It must also be an even function of the σ 's, i.e., if we change the sign of all the σ 's then the

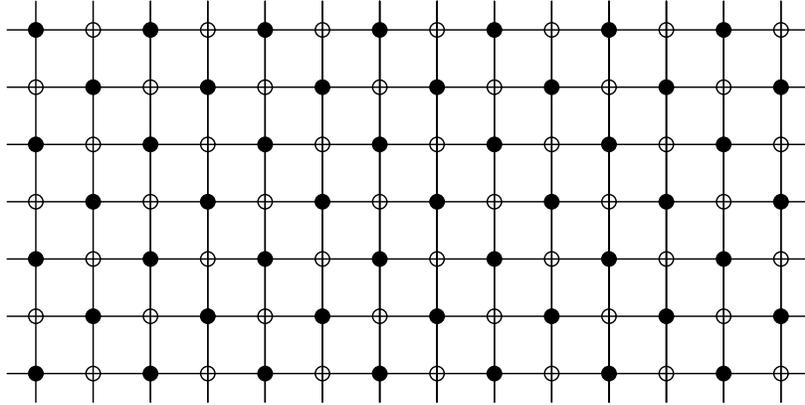


Figure 3: The division of the lattice lattice into even and odd sites in two dimensions.

value of the function is unchanged. Hence it must be of the form

$$\exp[a(\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_4 + \sigma_1\sigma_4) + b(\sigma_1\sigma_3 + \sigma_2\sigma_4) + c\sigma_1\sigma_2\sigma_3\sigma_4 + d]$$

where a, b, c, d are functions of β . It is not hard to see that the coefficient a must equal the coefficient b . We have distinguished between the pairs with separation $\sqrt{2}$ and 2 for latter purposes.

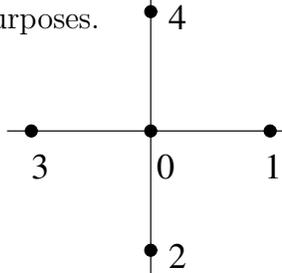


Figure 4: With a nearest neighbor interaction, a spin interacts only with the four spins shown.

In one dimension we were fortunate in that the renormalized Hamiltonian was of the same form as the original Hamiltonian, i.e., it consisted only of nearest neighbor bonds. As we now see this is not the case in two dimensions. The terms corresponding to the a parameter are fine since they are nearest neighbor couplings on the lattice of block spins. The terms corresponding to

the b parameter are a problem. They couple block spins which are not nearest neighbors. The four spin interaction associated with c is also a problem. These non-nearest neighbor terms are problems in the sense that we when try to compute the next iteration of the renormalization group transformation, the sum will not factor into a product of sums over individual spins. In fact the next iteration of the transformation may not be done in closed form. To avoid this problem we will cheat. We drop the four site interaction completely, i.e., we pretend that $c = 0$. Dropping the terms corresponding to b turns out to be too big a cheat. Instead we replace the term $b(\sigma_1\sigma_3 + \sigma_2\sigma_4)$ by $\frac{b}{2}(\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_4 + \sigma_1\sigma_4)$. We use $\frac{b}{2}$ rather than just b to account for the relative number of terms. Thus our cheat may be summarized as follows:

$$\begin{aligned} \exp[a(\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_4 + \sigma_1\sigma_4) + b(\sigma_1\sigma_3 + \sigma_2\sigma_4) + c\sigma_1\sigma_2\sigma_3\sigma_4 + d] \\ \rightarrow \exp[(a + b/2)(\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_4 + \sigma_1\sigma_4) + d] \end{aligned}$$

Of course the reason for this cheating is to end up with a renormalized Hamiltonian which only contains nearest neighbor interactions. For each nearest neighbor bond in the lattice of block spins there are two odd sites in the original lattice for which the above computation will produce the contribution $a + b/2$ for the nearest neighbor bond in the block spin lattice. Thus H_1 is just the nearest neighbor Ising model with

$$\beta_1 = 2a + b$$

We need to find a and b as functions of β . By considering various choices for $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ in ?? we obtain the equations

$$\begin{aligned} 2 &= \exp(-4a + 2b + c + d) \\ 2 &= \exp(-2b + c + d) \\ e^{4\beta} + e^{-4\beta} &= \exp(6a + c + d) \\ e^{2\beta} + e^{-2\beta} &= \exp(-c + d) \end{aligned}$$

Solving for a and b and using $\beta_1 = 2a + b$ yields $\beta_1 = g(\beta)$ with

$$g(\beta) = \frac{3}{8} \ln\left(\frac{e^{4\beta} + e^{-4\beta}}{2}\right)$$

A graph of this function is shown in figure 5

The important feature of this function is that it has a nonzero fixed point $g(\beta_0) = \beta_0$ and this fixed point is unstable. If we start with a β near β_0 but

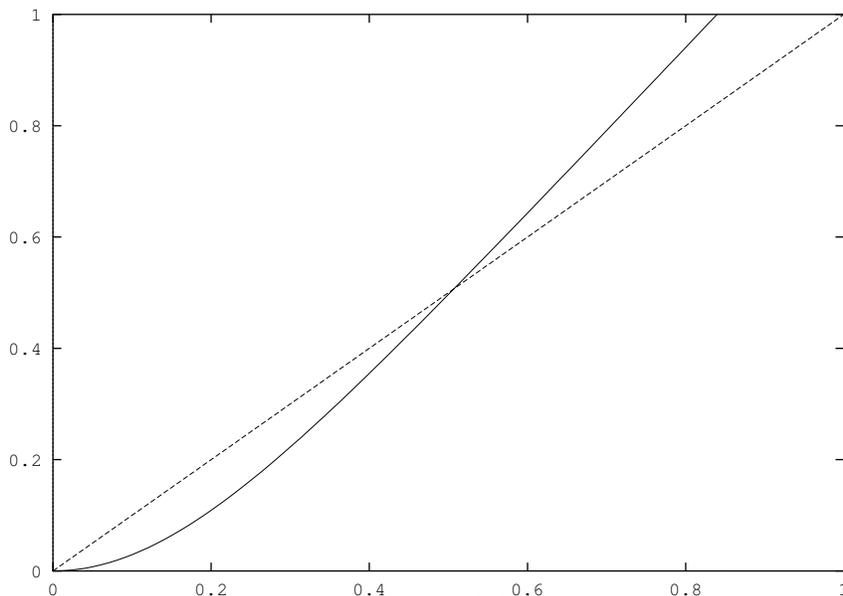


Figure 5: The solid line is the function $g(\beta)$ which gives the evolution of the coupling β under decimation in the approximate transformation for the two dimensional Ising model. The important feature is that there is a nonzero fixed point and it is unstable.

less than β_0 and iterate the map g , then the resulting β 's will move away from β_0 and eventually converge to zero. Likewise, if we start slightly above β_0 the iterations will run off to ∞ . Suppose we know that for very small β the model is in a high temperature phase (boundary conditions do not matter) and for very large β the model is in a low temperature phase (boundary conditions matter). Then the renormalization group transformation shows that any β below β_0 is eventually driven into the high temperature phase and so the original β must be in the high temperature phase. Likewise any $\beta > \beta_0$ must be in the low temperature phase. Thus the fixed point β_0 is the critical point β_c .

The fixed point and the slope at the fixed point are given approximately by

$$\beta_c \approx 0.507, \quad g'(\beta_c) \approx 1.449$$

The value of β_c from the exact solution is $\beta_c = 0.44 \dots$, so our approximation did not do too badly. We are much more interested in the critical exponents rather than the value of β_c . Can we say anything about them? We start

with the correlation length.

After we sum out the odd spins, the remaining even spins form a lattice with spacing $\sqrt{2}$. (It is rotated by 45 degrees with respect to the original lattice.) As in the one dimensional model the correlation functions of the renormalized model are directly related to correlation functions of the original model. The only difference is that the lattice spacing must be rescaled by a factor of $\sqrt{2}$ rather than 2. The argument we will give applies in a general setting, so we will let b denote the factor by which lengths are rescaled by the transformation. For the transformation we are studying, $b = \sqrt{2}$. We have

$$\xi(g(\beta)) = \xi(\beta)/b \tag{11}$$

It is a good idea to write $g'(\beta_c) = b^{y_T}$. (This equation defines y_T . In our example, $y_T = 1.07$.) We assume that for very small β we understand the model, in particular the correlation length, very well. Let β_0 be a fixed small number, and consider the interval $(g(\beta_0), \beta_0]$. A little thought shows that if we take a β in (β_0, β_c) and apply g repeatedly, exactly one of the iterates will lie in the interval $(g(\beta_0), \beta_0]$. So we can define $N(\beta)$ to be the integer such that $g^{N(\beta)}(\beta) \in (g(\beta_0), \beta_0]$.

When β is near β_c , $g(\beta)$ is essentially $\beta_c + g'(\beta_c)(\beta - \beta_c) = \beta_c + b^{y_T}(\beta - \beta_c)$. As long as $g^k(\beta)$ is still close to β_c , $g^k(\beta) \approx \beta_c + b^{ky_T}(\beta - \beta_c)$. If this equation were always true we would have

$$N(\beta) = \frac{-\ln(\beta_c - \beta)}{y_T \ln b}$$

We leave it as an exercise to show that the true value of $N(\beta)$ differs from this result by an additive correction, i.e., there is a constant c such that

$$\left| N(\beta) - \frac{-\ln(\beta_c - \beta)}{y_T \ln b} \right| \leq c$$

for all $\beta \in (\beta_0, \beta_c)$.

Iterating eq. (11) N times we have

$$b^N \xi(g^N(\beta)) = \xi(\beta)$$

We assume that we can show that the correlation length is bounded from above and from below on the interval $(g(\beta_0), \beta_0]$, i.e., there are positive constants c_1 and c_2 such that

$$c_1 \leq \xi(\beta) \leq c_2, \quad \beta \in (g(\beta_0), \beta_0]$$

Then with $N = N(\beta)$, the iterate $g^N(\beta)$ is in this interval and so

$$c_1 b^N \leq \xi(\beta) \leq c_2 b^N$$

Up to a factor of e^c , $b^N = (\beta_c - \beta)^{-1/y_T}$, so we have

$$c_1 e^{-c} (\beta_c - \beta)^{-1/y_T} \leq \xi(\beta) \leq c_2 e^c (\beta_c - \beta)^{-1/y_T}$$

and so we have calculated ν in terms of y_T .

$$\nu = 1/y_T \tag{12}$$

For our example this yields $\nu = 0.934$. The exact value of ν is 1, so we didn't do too badly.

In the above argument we didn't assume anything about the behavior of the correlation length near the critical point. We used the RG transformation to get into the very high temperature region and then assumed that we could get upper and lower bounds on the correlation length in that regime. If we are willing to assume that the correlation length has a power law divergence at the critical point, then we can give a very simple derivation of eq. (12). We assume that

$$\lim_{\beta \rightarrow \beta_c} \ln \xi(\beta) - \nu \ln |\beta - \beta_c| = 0$$

Taking the logarithm of eq. (11) we have

$$\ln \xi(g(\beta)) = \ln \xi(\beta) - \ln b$$

and so

$$\lim_{\beta \rightarrow \beta_c} [\nu \ln |g(\beta) - \beta_c| - \nu \ln |\beta - \beta_c| + \ln b] = 0$$

But of course

$$\lim_{\beta \rightarrow \beta_c} [\ln |g(\beta) - \beta_c| - \ln |\beta - \beta_c|] = \lim_{\beta \rightarrow \beta_c} \ln \frac{|g(\beta) - \beta_c|}{|\beta - \beta_c|} = \ln |g'(\beta_c)| = y_T \ln b$$

So we obtain $\nu y_T \ln b = \ln b$, which yields (12).

Next we show how the exponent α can be derived from the renormalization group. The derivation we will give is along the lines of the second derivation of the exponent ν that we gave above. Let $f(\beta)$ be the free energy

per site as a function of β . Recall that α is the power in the divergence of the second derivative of this free energy.

$$f''(\beta) \sim |\beta - \beta_c|^{-\alpha}$$

Letting L^d be the number of sites in our lattice, the free energy is defined by

$$\exp(-\beta f(\beta)L^d) = Z$$

After the renormalization group transformation we obtain a model with nearest neighbor coupling $g(\beta)$. The renormalized Hamiltonian also contains a constant term d . It is some function of β that we did not bother to compute. Let $h(\beta)$ be the constant term per site. The renormalized system and the original system have the same partition function. Keeping in mind that the number of sites in the renormalized system is $L^d/2$, we have

$$Z = \exp(-g(\beta)f(g(\beta))L^d/2 + h(\beta)L^d/2)$$

and so

$$-\beta f(\beta) = -g(\beta)f(g(\beta))/2 + h(\beta)/2$$

We now want to differentiate this equation twice with respect to β . This is a bit of a mess, but we are only interested in the terms which will be singular at β_c . g and h are analytic at β_c while f and f' are continuous at β_c . Hence

$$-\beta f''(\beta) = -\frac{1}{2}g(\beta)f''(g(\beta))[g'(\beta)]^2 + r(\beta)$$

where $r(\beta)$ is continuous. As $\beta \rightarrow \beta_c$, $g(\beta) \rightarrow \beta_c$, so

$$f''(\beta) = \frac{1}{2}f''(g(\beta))[g'(\beta)]^2 + s(\beta)$$

where $s(\beta)$ is another continuous function of β . Now we assume that

$$f''(\beta) \sim c|\beta - \beta_c|^{-\alpha}$$

Then

$$\frac{1}{2}f''(g(\beta))[g'(\beta)]^2 \sim \frac{1}{2}c|g(\beta) - \beta_c|^{-\alpha} b^{2y_T}$$

where we have used the continuity of $g'(\beta)$ to replace $g'(\beta)$ by $g'(\beta_c)$ which is equal to b^{y_T} . Now $|g(\beta) - \beta_c|^{-\alpha} \sim [g'(\beta_c)|\beta - \beta_c|]^{-\alpha} = b^{-\alpha y_T} |\beta - \beta_c|^{-\alpha}$ and thus we must have

$$1 = \frac{1}{2}b^{y_T(2-\alpha)}$$

The factor of $\frac{1}{2}$ was the ratio of the number of sites in the new lattice to the number of sites in the old lattice. In general this is given by b^{-d} where b is the factor by which lengths are rescaled by the transformation. (In our example $b = \sqrt{2}$, so $b^{-d} = \frac{1}{2}$.) Thus the general result is

$$1 = b^{-d} b^{y_T(2-\alpha)}$$

which says

$$\alpha = 2 - \frac{d}{y_T}$$

What about the other critical exponents? To compute γ , β and δ we would need to include a magnetic field in our Hamiltonian. This can be done but we will leave all the fun to the reader. (Exercise 2.2.2) We will return to this problem in a more general setting later. The final exponent η describes the decay of the correlation function at the critical point. As we will see at the end of this section, when we try to compute η we find that the correlation function at the critical point does not decay at all! Something is wrong, which shouldn't be a great surprise given the big cheat.

We now consider the decimation transformation without this big cheat we introduced to force the renormalized Hamiltonian to remain nearest neighbor. The first iteration of the transformation can be computed in closed form. The resulting renormalized Hamiltonian H_1 contains interactions which are not nearest neighbor. This prevents us from computing H_2 in the same way that we computed H_1 . (The sums over the odd spins do not factor.) Nonetheless we can still try to define H_2 by the equation

$$e^{-H_2} = \sum e^{-H_1}$$

where the sum is over half of the spins that appear in H_1 . For a finite volume this equation defines a function H_2 . The highly nontrivial question is what happens in the infinite volume limit.

Even in a finite volume H_2 will contain all sorts of terms besides the nearest neighbor interaction. Let Λ_2 denote the sites that are left after the second transformation, so H_2 is a function of the spins at these sites. The most general such function can be written in the form

$$H_2 = \sum_{X \subset \Lambda_2} c(X) \sigma(X)$$

where X is summed over all subsets of Λ_2 and $\sigma(X)$ is shorthand for $\prod_{i \in X} \sigma_i$. (See exercise 2.2.1) The coefficients $c(X)$ will depend on the finite volume Λ_2 . We might hope that they have an infinite volume limit. However, we need much more than that. In the infinite volume limit our Hamiltonian will have an infinite number of different types of terms. Such systems make sense only if the interactions have some decay properties. For example consider the terms $c(X)$ where X consists of two sites. At the very least one needs that $c(\{i, j\})$ decays fast enough that for each i , $\sum_j |c(\{i, j\})| < \infty$. Assuming that one has shown that the coefficients have infinite volume limits and the resulting Hamiltonian is sufficiently nice, one can then proceed to study the resulting map in the infinite dimensional space of Hamiltonians. In particular one would like to show it has a fixed point. We will return to all these questions later, but for now we will give a simple argument that this program must fail for the decimation transformation.

Suppose that there is a H such that the renormalized Hamiltonian H_1 (after rescaling the lattice) equals H . (So H is a fixed point of the renormalization group.) We showed before that for any H the correlation functions of the original system and the renormalized system are related by

$$\langle s_i s_j \rangle_1 = \langle \sigma_{i'} \sigma_{j'} \rangle_0$$

where $|i - j| = |i' - j'|/b$. But since H is a fixed point and two systems with the same Hamiltonian have the same correlation functions, we have

$$\langle s_i s_j \rangle_1 = \langle \sigma_i \sigma_j \rangle_0$$

and so $\langle \sigma_{i'} \sigma_{j'} \rangle_0 = \langle \sigma_i \sigma_j \rangle_0$. This means that the correlation functions at the critical point don't decay (or are identically zero) and no one expects this to be true.

Luckily, decimation is not the only renormalization group transformation. In the next section we will define several others which do not suffer from the above problem.

Exercises:

3.2.1 (this has nothing to do with the RG, it is here to fill in a gap) Let Λ be a finite set of sites. Let $H(\sigma)$ be any function defined on the spins configurations on Λ . Note that the space of such H is a vector space with dimension $2^{|\Lambda|}$. Define the "Fourier coefficients" of $H(\sigma)$ by

$$c(X) = 2^{-|\Lambda|} \sum_{\sigma} \sigma(X) H(\sigma)$$

where X is any subset of Λ , $\sigma(X)$ is a shorthand for $\prod_{i \in X} \sigma_i$ and the sum is over all spin configurations on Λ . Show that we then have

$$H(\sigma) = \sum_X c(X) \sigma(X)$$

where the sum is over all subsets of Λ . Hint : $\sum_{\sigma} \sigma(X) \sigma(Y)$ equals 0 unless $X = Y$. We say that $H(\sigma)$ is even if $H(-\sigma) = H(\sigma)$ where $-\sigma$ is the spin configuration obtained from σ by changing the sign of the spin at every site. Show that if $H(s)$ is even then the only nonzero $c(X)$ are those for which X contains an even number of sites.

3.2.2 Consider what happens to the transformation defined in this section when we include a magnetic field term in the Hamiltonian. Include the big cheat as we did before. You will even need to cheat some more to make the renormalized Hamiltonian contain only a nearest neighbor term and a magnetic field term. The fixed point occurs at $h = 0$, so the fixed point is as before. The linearization about the fixed point is different now since the RG map is from R^2 into R^2 . Compute the linearization and its two eigenvalues. Find the critical exponents β, γ, δ . More explanation needed here.

3.3 Other renormalization group transformations

We now consider renormalization group transformations other than decimation. One of the most successful for numerical work is the “majority rule” transformation. We first explain how it works in two dimensions with 3 by 3 blocks. We divide the lattice into disjoint 3 by 3 blocks and for each block we introduce a new “block” spin which also takes on the values of only +1 and -1. Thus the block spins may be pictured as living on a lattice with a spacing of 3 units. We define a map from spin configurations on the original lattice to block spin configurations as follows. In each block we look at the values of the spins at the 9 sites in the original lattice. If a majority are +1, then the block spin is taken to be +1. If a majority are -1, then the block spin is taken to be -1. We denote this map from a spin configuration on the original lattice to a spin configuration on the block spin lattice by $\sigma \rightarrow s$. This map is something called “blocking the spins”. In decimation one computes the renormalized Hamiltonian H_1 by fixing the spins that are not being summed over and computing a “partial partition function.” We do something similar for the majority rule. Given a spin configuration s for the

block spins, we compute $H_1(s)$ by only summing over the spin configurations σ on the original lattice which yield s when blocked, i.e., $\sigma \rightarrow s$.

$$e^{-H_1(s)} = \sum_{\sigma: \sigma \rightarrow s} e^{-H(\sigma)}$$

Every spin configuration σ on the original lattice corresponds to exactly one block spin configuration s . Thus if we sum the above equation over all the block spins configurations s we will recover the full partition function.

$$\sum_s e^{-H_1(s)} = Z$$

We complete the renormalization group transformation by rescaling the block spin lattice by a factor of 3. We can then iterate the renormalization group transformation.

We have used 3 by 3 blocks instead of 2 by 2 blocks to avoid ties. In a 2 by 2 block we could have equal numbers of +1's and -1's. One can still define a majority rule transformation for 2 by 2 blocks. One usually says something like "in the case of a tie one chooses randomly between the two values of the block spin." I find this wording confusing since it suggests that some extra element of randomness is being introduced although none is. What is done is the following. First we need a little notation. We will denote sites for the block spins by α . So s_α is a block spin. We label the sites for the original spins with two labels : α, i where α indicates which block the original spin belongs to and $i = 1, 2, 3, 4$ indicates which site within that block. Consider a single block α . We define a kernel

$$k(s_\alpha, \sigma_{\alpha,1}, \sigma_{\alpha,2}, \sigma_{\alpha,3}, \sigma_{\alpha,4}) = \begin{cases} 1, & \text{if } \sigma_\alpha \sum_i \sigma_{i,\alpha} > 0 \\ \frac{1}{2}, & \text{if } \sum_i \sigma_{i,\alpha} = 0 \\ 0, & \text{if } \sigma_\alpha \sum_i \sigma_{i,\alpha} < 0 \end{cases} \quad (13)$$

In words, the kernel is 1 if there is a clear majority in the block and it agrees with the block spin, 0 if there is a clear majority and it is different from the block spin and $\frac{1}{2}$ if there is a tie in the block. We then define a kernel for the entire lattice by

$$K(s, \sigma) = \prod_{\alpha} k(s_\alpha, \sigma_{\alpha,1}, \sigma_{\alpha,2}, \sigma_{\alpha,3}, \sigma_{\alpha,4})$$

where the product is over all the blocks. The renormalized Hamiltonian is now defined by

$$e^{-H_1(s)} = \sum_{\sigma} K(s, \sigma) e^{-H(\sigma)}$$

A single configuration σ can now contribute to several block spin configurations s . However, if we sum over all the block spin configurations we find

$$\sum_s e^{-H_1(s)} = \sum_s \sum_{\sigma} K(s, \sigma) e^{-H(\sigma)} = \sum_{\sigma} \sum_s K(s, \sigma) e^{-H(\sigma)} = \sum_{\sigma} e^{-H(\sigma)} = Z$$

where we have used the property of the kernel

$$\sum_s K(s, \sigma) = \prod_{\alpha} \left[\sum_{s_{\alpha}} k(s_{\alpha}, \sigma_{\alpha,1}, \sigma_{\alpha,2}, \sigma_{\alpha,3}, \sigma_{\alpha,4}) \right] = 1$$

which follows from the easily checked identity

$$\sum_{s_{\alpha}} k(s_{\alpha}, \sigma_{\alpha,1}, \sigma_{\alpha,2}, \sigma_{\alpha,3}, \sigma_{\alpha,4}) = 1$$

Note that the 3 by 3 majority rule transformation may be restated in terms of a kernel. The kernel k for a single block equals 1 if the majority of the spins in the block equal the block spin and equals 0 if they do not.

For the majority rule transformations we were able to compute the original partition function using the renormalized Hamiltonian H_1 by

$$Z = \sum_s e^{-H_1(s)} \tag{14}$$

because the kernel $K(s, \sigma)$ satisfied the equation

$$\sum_s K(s, \sigma) = 1 \tag{15}$$

Given any kernel satisfying (15) we can try to use (3.3) to define a renormalization group transformation. In addition to (15) we should also require that the kernel satisfy $K(s, \sigma) \geq 0$ so that the quantity we are taking the logarithm of is manifestly positive. Just as in the case of decimation this equation defines the renormalized Hamiltonian for a finite volume, but one must then deal with the infinite volume limit.

Another family of transformations are the “quasi-linear” transformations. For a 2 by 2 block the transformation is given by

$$k(s_\alpha, \sigma_{\alpha,1}, \sigma_{\alpha,2}, \sigma_{\alpha,3}, \sigma_{\alpha,4}) = \frac{1}{2} + a s_\alpha (\sigma_{\alpha,1} + \sigma_{\alpha,2} + \sigma_{\alpha,3} + \sigma_{\alpha,4})$$

where a is a parameter. Obviously, $\sum_{s_\alpha} k(s_\alpha, \sigma_{\alpha,1}, \sigma_{\alpha,2}, \sigma_{\alpha,3}, \sigma_{\alpha,4}) = 1$ for each α , so this transformation satisfies (15).

One of the nice features of the decimation transformation was the simple relation between the correlation functions of the transformed or renormalized system and the original system. We now need to investigate this relation for the other transformations we have defined. If the kernel is given by a product over blocks as in (3.3), then we have

$$\begin{aligned} \langle s_\alpha s_\beta \rangle &= Z^{-1} \sum_s s_\alpha s_\beta e^{-H_1(s)} \\ &= Z^{-1} \sum_s s_\alpha s_\beta \sum_\sigma K(s, \sigma) e^{-H(\sigma)} \\ &= Z^{-1} \sum_\sigma \left[\sum_s s_\alpha s_\beta K(s, \sigma) \right] e^{-H(\sigma)} \end{aligned}$$

Since $\sum_{s_\gamma} k(s_\gamma, \sigma_{\gamma,i}) = 1$, we have

$$\sum_s s_\alpha s_\beta K(s, \sigma) = \left[\sum_{s_\alpha} s_\alpha k(s_\alpha, \sigma_{\alpha,i}) \right] \left[\sum_{s_\beta} s_\beta k(s_\beta, \sigma_{\beta,i}) \right]$$

For the quasilinear transformation,

$$\sum_{s_\alpha} s_\alpha k(s_\alpha, \sigma_{\alpha,i}) = \sum_{s_\alpha} s_\alpha \left[\frac{1}{2} + a s_\alpha (\sigma_{\alpha,1} + \sigma_{\alpha,2} + \sigma_{\alpha,3} + \sigma_{\alpha,4}) \right] = 2a (\sigma_{\alpha,1} + \sigma_{\alpha,2} + \sigma_{\alpha,3} + \sigma_{\alpha,4})$$

So

$$\langle s_\alpha s_\beta \rangle_1 = 4a^2 \langle (\sigma_{\alpha,1} + \sigma_{\alpha,2} + \sigma_{\alpha,3} + \sigma_{\alpha,4}) (\sigma_{\beta,1} + \sigma_{\beta,2} + \sigma_{\beta,3} + \sigma_{\beta,4}) \rangle \quad (16)$$

For blocks α and β we define $|\alpha - \beta|$ to be the distance between the block spins measured in units in which nearest neighbor block spins are one unit apart. (This is a natural definition since we will rescale the block spin lattice so that nearest neighbors are one unit apart.) Then the distance from any one

of $\sigma_{\alpha,1}, \sigma_{\alpha,2}, \sigma_{\alpha,3}, \sigma_{\alpha,4}$ to any one of $\sigma_{\beta,1}, \sigma_{\beta,2}, \sigma_{\beta,3}, \sigma_{\beta,4}$ is essentially $2|\alpha - \beta|$. Thus we obtain the same relation between correlation lengths, $\xi_1 = \xi/b$, that we had in the case of decimation. The length rescaling factor b is equal to 2 for the 2 by 2 quasilinear transformation.

growing prefactor ??

While eq. (16) yields the desired result for the correlation length, it was precisely this sort of equation that showed that the decimation transformation could not have a fixed point. Suppose that the quasilinear transformation does have a fixed point. We also assume that at this fixed point the correlation function exhibits power law decay.

$$\langle \sigma_x \sigma_y \rangle \sim \frac{c}{|x - y|^{d-2+\eta}}$$

Since the transformed system has exactly the same Hamiltonian as the original system, $\langle s_\alpha s_\beta \rangle_1$ decays in exactly the same way

$$\langle s_\alpha s_\beta \rangle_1 \sim \frac{c}{|\alpha - \beta|^{d-2+\eta}}$$

The distances from the four sites in block α to the four sites in block β are all essentially $2|\alpha - \beta|$. So

$$\langle (\sigma_{\alpha,1} + \sigma_{\alpha,2} + \sigma_{\alpha,3} + \sigma_{\alpha,4})(\sigma_{\beta,1} + \sigma_{\beta,2} + \sigma_{\beta,3} + \sigma_{\beta,4}) \rangle \sim \frac{16c}{[2|\alpha - \beta|]^{d-2+\eta}}$$

Thus eq. (16) can hold only if

$$1 = 4a^2 \frac{16}{2^{d-2+\eta}} = a^2 \frac{64}{2^\eta}$$

so a must be $2^{\eta/2}/8$. Hence the quasilinear transformation can have a fixed point only for one precise value of the parameter a . While this complicates the search for a fixed point, it also give us a method of computing η .

Finally we consider the 2 by 2 majority rule transformation. To compute $\sum_{s_\alpha} s_\alpha k(s_\alpha, \sigma_{\alpha,i})$ it is useful to first note that the kernel can be rewritten as

$$\begin{aligned} k(s_\alpha, \sigma_{\alpha,1}, \sigma_{\alpha,2}, \sigma_{\alpha,3}, \sigma_{\alpha,4}) &= \frac{1}{2} + \frac{3}{16} s_\alpha (\sigma_{\alpha,1} + \sigma_{\alpha,2} + \sigma_{\alpha,3} + \sigma_{\alpha,4}) \\ &\quad - \frac{1}{16} s_\alpha (\sigma_{\alpha,2} \sigma_{\alpha,3} \sigma_{\alpha,4} + \sigma_{\alpha,1} \sigma_{\alpha,3} \sigma_{\alpha,4} + \sigma_{\alpha,1} \sigma_{\alpha,2} \sigma_{\alpha,4} + \sigma_{\alpha,1} \sigma_{\alpha,2} \sigma_{\alpha,3}) \end{aligned}$$

So we have

$$\sum_{s_\alpha} s_\alpha k(s_\alpha, \sigma_{\alpha,1}, \sigma_{\alpha,2}, \sigma_{\alpha,3}, \sigma_{\alpha,4}) = \frac{3}{16}(\sigma_{\alpha,1} + \sigma_{\alpha,2} + \sigma_{\alpha,3} + \sigma_{\alpha,4}) - \frac{1}{16}(\sigma_{\alpha,2}\sigma_{\alpha,3}\sigma_{\alpha,4} + \sigma_{\alpha,1}\sigma_{\alpha,3}\sigma_{\alpha,4} + \sigma_{\alpha,1}\sigma_{\alpha,2}\sigma_{\alpha,4} + \sigma_{\alpha,1}\sigma_{\alpha,2}\sigma_{\alpha,3})$$

Now we are in trouble. The two spin correlation function $\langle s_\alpha s_\beta \rangle_1$ may be expressed in terms of correlation functions in the original system, but in addition to two spin correlation functions we also find four and six spin correlation functions. To obtain the relation $\xi_1 = \xi/b$ we must now assume that all correlation functions decay with the same length.

MORE ???

For all of the transformations we have just introduced the renormalization group map may not be computed in closed form even in one dimension. We must now make some sort of approximation. For example, rather than working in the infinite dimensional space of Hamiltonians we could choose some finite dimensional space of Hamiltonians and attempt to compute the RG map projected onto this subspace in some way. The infinite volume limit is still demanding to be taken. Typically what is done is to compute the map in several finite volumes and then extrapolate to the infinite volume. The majority rule transformation has probably been the most studied numerically. Based on these calculations and ??? the following is believed to happen in two and three dimensions.

There is a nontrivial fixed point. The coefficients of the terms in this fixed point appear to have good decay properties. In two dimensions the two largest interactions are the nearest neighbor interaction (coefficient is about 0.352) and the diagonal interaction between two sites (coefficient is about 0.097). In three dimensions the coefficients of these terms are ??? This fixed point belongs to a stable manifold which is of codimension two in the space of Hamiltonians. (The stable manifold is the set of Hamiltonians for which the iterations of the renormalization group converge to the fixed point.) All of the Hamiltonians on this stable manifold describe systems at a critical point. The usual nearest neighbor Ising model Hamiltonian belongs to this manifold when $\beta = \beta_c$. If we start with a Hamiltonian in this manifold, then iterations of the renormalization group will drive it to the fixed point. This is the explanation of universality. While there are infinitely many critical Hamiltonians, the long distance properties of them are all described by the same fixed point Hamiltonian. (Of course one can ask where the global

structure of this map comes from. Why is there a fixed point with a large stable manifold, etc. ? It is sometimes said that the renormalization group does not explain universality, but rather assumes it.)

The fact that the manifold is of codimension two reflects the need to “tune” two parameters to put a system at its critical point. (In the usual nearest neighbor model in addition to $\beta = \beta_c$ one must also take $h = 0$ to obtain criticality.) Suppose we added another term to the Ising Hamiltonian, for example consider the Hamiltonian

$$H = K_1 \sum_{i,j:|i-j|=1} \sigma_i \sigma_j + K_2 \sum_{i,j:|i-j|=\sqrt{2}} \sigma_i \sigma_j + h \sum_i \sigma_i$$

If we fix K_2 and vary the nearest neighbor coupling K_1 and the magnetic field h then we obtain a two dimensional manifold of Hamiltonians which will intersect the stable manifold in one point which corresponds to the critical values of K_1 and h for this modified model.

The linearization of the transformation about the fixed point has spectrum strictly contained in the unit disc except for two real eigenvalues which are strictly greater than 1. These eigenvalues are usually written in the form b^{y_T} and b^{y_H} where b is the length rescaling factor, e. g., $b = 2$ for majority rule with 2 by 2 blocks in two dimensions or 2 by 2 by 2 blocks in three dimensions.

If we start with a Hamiltonian that is very close to critical, then under the renormalization group transformation it will flow towards the fixed point for a large number of iterations. The initially small distance to the stable manifold will grow and eventually the unstable directions will dominate and drive the Hamiltonian away from the fixed point. The closer we start to the criticality, the more iterations we will spend near the fixed point. A key tenet of the renormalization group is that the renormalization group map is smooth (at the very least twice differentiable). The singular behavior of the model at the critical point arises from iterating the map infinitely many times. (We already saw this in our simple RG map for the two dimensional Ising model.) The critical properties are determined by how much time the Hamiltonian spends near the fixed point. This is determined by the linearization about the fixed point, more precisely on y_T and y_H . All six of the critical exponents are determined by these two “eigenvalues.” Accordingly, numerical studies often concentrate on calculating these two numbers to as many decimal places as possible and do not bother to systematically investigate

the global nature of the RG map, etc. (That is both a shot at the physicists and a suggestion for a thesis problem.)

In the following we let R denote the renormalization group transformation. So R is a map on the space of Hamiltonians. If we start with the usual nearest neighbor Ising Hamiltonian with β close to β_c , then after a few iterations it will be close to the fixed point and we can then use the linearization of the renormalization group transformation. (Numerical people find that about five or six iterations get you close to the fixed point.) In the subspace of even interactions there is only one unstable direction. So when the iterations eventually move away from the fixed point, they do so in this unstable direction. Let H_0 denote the fixed point of the transformation and H_T the eigenvector of the linearization associated with the unstable direction. Let $H(\beta)$ denote $-\beta \sum_{\langle i,j \rangle} \sigma_i \sigma_j$, the usual nearest neighbor Ising Hamiltonian. Then we expect that for a moderately large but fixed integer n we have

$$R^n(H(\beta)) = H_0 + c(\beta - \beta_c)H_T \quad (17)$$

where c is a constant that depends on n . (Note: n does not depend on β . more ??) We are assuming that the iterate $R^n(H(\beta))$ is still close to the fixed point. So if we apply the transformation one more time we can use the linearization of the transformation and so

$$R^{n+1}(H(\beta)) = H_0 + c(\beta - \beta_c)b^{y_T} H_T \quad (18)$$

These two equations imply that

$$R^n(H(\beta_c + b^{y_T}(\beta - \beta_c))) = R^{n+1}(H(\beta)) \quad (19)$$

Note that we are *not* asserting that

$$H(\beta_c + b^{y_T}(\beta - \beta_c)) = R(H(\beta)) \quad (20)$$

This equation is grossly untrue. The intuition behind (19) is as follows. The renormalization group trajectories that start at $H(\beta)$ and $H(\beta_c + b^{y_T}(\beta - \beta_c))$ are not the same, but at each iteration the distances to the stable manifold of the two trajectories are related by a factor of b^{y_T} . Thus if we apply the transformation one extra time to the closer trajectory in the region where the linearization applies, we end up at the same Hamiltonian that the farther trajectory ends at.

Now we consider the exponent for the correlation length. We assume that the renormalization group transformation reduces the correlation length by a factor of b . If we let $\xi(H)$ denote the correlation length as a function of the Hamiltonian, then we are assuming that

$$\xi(R(H)) = \xi(H)/b \tag{21}$$

Using (18) this implies

$$b \xi(\beta_c + b^{y_T}(\beta - \beta_c)) = \xi(\beta)$$

where we have denoted $\xi(H(\beta))$ by simply $\xi(\beta)$. If we assume that

$$\xi(\beta) \sim c|\beta - \beta_c|^{-\nu}$$

then we must have

$$bb^{-y_T\nu} = 1$$

and so

$$\nu = \frac{1}{y_T}$$

Next we rederive the exponent α . Recall that the free energy per site is given by

$$\lim_{\Lambda \rightarrow \infty} -\frac{1}{\beta} \frac{1}{|\Lambda|} \ln Z_\Lambda$$

It is convenient to define

$$f(H) = \lim_{\Lambda \rightarrow \infty} -\frac{1}{|\Lambda|} \ln Z_\Lambda$$

so $f(H)$ differs from the free energy by a factor of β . The exponent α describes how the specific heat diverges as $\beta \rightarrow \beta_c$, and the specific heat is given by the second derivative of the free energy. This second derivative and the second derivative of f differ by a function that does not diverge, so it is enough to determine how the second derivative of $f(H(\beta))$ with respect to β diverges. The renormalization group transformation preserves the partition function, but reduces the number number of sites by a factor of b^d . Hence we have

$$b^d f(H) = f(R(H)) + h(H) \tag{22}$$

where $h(H)$ is the spin independent term that comes out of the renormalization group transformation. Eq. (19) implies

$$f(R^n(H(\beta_c + b^{y_T}(\beta - \beta_c)))) = f(R^{n+1}(H(\beta)))$$

A key tenet of the renormalization group philosophy is that the renormalization group map, including $h(H)$, is smooth. Thus if we apply (22) we have

$$f(H(\beta_c + b^{y_T}(\beta - \beta_c))) = b^d f(H(\beta)) + \text{smooth}$$

where *smooth* is some smooth function of β . We denote $f(H(\beta))$ by simply $f(\beta)$, so the above becomes

$$f(\beta_c + b^{y_T}(\beta - \beta_c)) = b^d f(\beta) + \text{smooth}$$

Now we differentiate twice to obtain

$$b^{2y_T} f''(\beta_c + b^{y_T}(\beta - \beta_c)) = b^d f''(\beta) + \text{smooth}$$

Since $f''(\beta) \sim c|\beta - \beta_c|^{-\alpha}$, we must have

$$b^{2y_T - y_T \alpha} = b^d$$

which yields

$$\alpha = 2 - \frac{d}{y_T}$$

To derive the exponents β, δ and γ we need to include a magnetic field. Let H_h denote the eigenvector corresponding to the eigenvalue b^{y_h} . So H_h is an odd spin Hamiltonian. Let

$$H(\beta, h) = -\beta \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i$$

As before if we start with β close to β_c and h close to 0, then the iterates of the transformation will flow close to the fixed point H_0 , but then eventually move off in the unstable directions. The displacement in the two unstable directions should be a linear function of $\beta - \beta_c$ and h . The most general linear function would yield

$$R^n(H(\beta, h)) = H_0 + c_1(\beta - \beta_c)H_T + d_1 h H_T + c_2 h H_h + d_2(\beta - \beta_c)H_h \quad (23)$$

where c_1, c_2, d_1, d_2 are constants which depend on n . (As before n is a moderately large, but fixed integer.) If $h = 0$ then the starting Hamiltonian is even, and so all its iterates remain even. Hence d_2 must be zero. It is possible to show that d_1 must also be zero, but this takes more work and so will be left to a homework problem. So now we have

$$R^n(H(\beta, h)) = H_0 + c_1(\beta - \beta_c)H_T + c_2hH_h$$

Assuming that $R^n(H(\beta, h))$ is still close to the fixed point, we have

$$R^{n+1}(H(\beta, h)) = H_0 + c_1(\beta - \beta_c)b^{y_T}H_T + c_2hb^{y_h}H_h$$

Hence we have

$$R^n(H(\beta_c + b^{y_T}(\beta - \beta_c), b^{y_h}h)) = R^{n+1}(H(\beta, h))$$

Let $f(\beta, h)$ be minus the logarithm of partition function per site as above. Then as before we have

$$f(\beta_c + b^{y_T}(\beta - \beta_c), b^{y_h}h) = b^d f(\beta, h) + \text{smooth} \quad (24)$$

Now we can find formulae for the exponents β, γ and δ . With our definitions of $f(H)$ and $H(\beta, h)$, the magnetization per site is equal to $\frac{\partial f}{\partial h}$. Letting $m(\beta, h)$ denote the magnetization per site, eq. (24) yields

$$b^{y_h}m(\beta_c + b^{y_T}(\beta - \beta_c), b^{y_h}h) = b^d m(\beta, h) + \text{smooth} \quad (25)$$

The exponent β describes how this magnetization converges to zero as $\beta \rightarrow \beta_c$ with $h = 0$,

$$m(\beta, 0) \sim c|\beta - \beta_c|^\beta$$

Hence we must have

$$b^{y_h}b^{y_T\beta} = b^d$$

and so $y_h + y_T\beta = d$ which yields

$$\beta = \frac{d - y_h}{y_T}$$

The exponent δ describes how the magnetization at the critical temperature goes to zero as the magnetic field is turned off.

$$m(\beta_c, h) \sim h^{\frac{1}{\delta}}$$

Thus eq. (25) implies

$$b^{y_h} b^{y_h/\delta} = b^d$$

and so we obtain

$$\delta = \frac{y_h}{d - y_h}$$

Finally, the exponent γ describes how the magnetic susceptibility diverges as $\beta \rightarrow \beta_c$. To obtain the susceptibility χ we need to take the derivative of the magnetization with respect to the magnetic field. Differentiating eq. (25) produces

$$b^{2y_h} \chi(\beta_c + b^{y_T}(\beta - \beta_c), b^{y_h} h) = b^d \chi(\beta, h) + \text{smooth} \quad (25)$$

Combining this with

$$\chi(\beta, 0) \sim |\beta - \beta_c|^{-\gamma}$$

we conclude

$$b^{2y_h} b^{-y_T \gamma} = b^d$$

and so

$$\gamma = \frac{2y_h - d}{y_T}$$

UNIVERSALITY HYPERSCALING

HOW does RG determine η

power law interactions have different exponents. But a finite dimensional approximation cannot distinguish power law vs exponential decay. So you may get spurious eigenvalues ???

Exercises:

3.3.1 (Research problem) If we do decimation in one dimension but start with a Hamiltonian that contains more than nearest neighbor couplings, the renormalized Hamiltonian H_1 will typically contain infinitely many types of couplings, i.e., we are forced into an infinite dimensional space of Hamiltonians. (With majority rule in one dimension we are forced into an infinite dimensional space of Hamiltonians even if we start with a nearest neighbor Hamiltonian.) Show that it is possible to rigorously define majority rule or decimation in this more general setting where the RG map is not finite dimensional. In particular this will defining an appropriate Banach space of Hamiltonians, showing that the transformation may be iterated infinitely many times and hopefully using all this to prove that the correlation length diverges with β in this more general setting

3.3.2 (reasonable) Show that $d_2 = 0$ in (23).

3.4 Numerical computation of Ising RG maps

We begin this section by explaining why the numerical computation of the renormalized Hamiltonian H_1 should be a better behaved computation than numerically computing the free energy of the original Ising model at a critical point.

Fix a block spin configuration s and look at the probability measure associated with this fixed choice of block spins:

$$\langle F(\sigma) \rangle_s = \frac{\sum_{\sigma} F(\sigma) K(s, \sigma) e^{-H(\sigma)}}{\sum_{\sigma} K(s, \sigma) e^{-H(\sigma)}} \quad (26)$$

Take $H(\sigma)$ to be at criticality. We conjecture that this constrained system is in the high temperature phase for all choices of the block sping configuration s . It fact, we expect it to have a correlation length of order 1 for all choices

of s . So computing H_1 does not require simulating or numerically studying a critical system.

In the simplest Monte Carlo simulation of the original probability measure $\langle \cdot \rangle$, there is a rule for flipping individual spins which is designed so that the time average over this flipping process of an observable or random variable converges to its expected value in $\langle \cdot \rangle$. One can ask how long it takes to generate an effectively new (or independent) configuration? If we are working in a box of dimension L it will take at least $O(L^d)$ such flips. Well away from the critical point it should in fact take no more than this. However, as $\beta \rightarrow \beta_c$, the time will grow and at the critical point it is of order L to a bigger power. This is known as “critical slowing down.”

Now consider simulating the probability measure $\langle \cdot \rangle_s$. It is not critical so there is no critical slowing down. So it is relatively easy to compute $H_1(s)$ for a given spin configuration. Of course, we want a lot more. We want some representation of the form

$$H_1(s) = \sum_X c_1(X) s(X) \quad (27)$$

How do you compute the $c_1(X)$? This is not easy. We must introduce some sort of approximation or “projection” onto a finite dimensional space of Hamiltonians. I’ll present an approach that is not in the literature.

Real space renormalization group calculations are usually done using the spin variables $\sigma_i = \pm 1$. Our method is based on what are sometimes called the lattice gas variables $n_i = (1 - \sigma_i)/2$ which take on the values 0, 1. Note that we have made the convention that a spin value of +1 corresponds to a lattice gas value of 0. We will use σ ’s for spin variables taking on the values ± 1 , and n ’s for lattice gas variables taking on the values 0, 1. We indicate renormalized spins or variables with a bar over them, e.g., $\bar{\sigma}_i, \bar{n}_i$. We use σ to denote the entire spin configuration $\{\sigma_i\}$. Likewise, $n, \bar{\sigma}$ and \bar{n} denote the corresponding collections of variables.

Initially we will work entirely in the lattice gas variables, both for the original Hamiltonian and the renormalized Hamiltonian. We write the renormalized Hamiltonian as

$$\bar{H}(\bar{n}) = \sum_Y c(Y) \bar{n}(Y) \quad (28)$$

where the sum is over all finite subsets including the empty set, and

$$\bar{n}(Y) = \prod_{i \in Y} \bar{n}_i$$

Consider the block variable configuration of all 0's. Our method for computing the renormalized Hamiltonian uses only block variable configurations which differ from this configuration at a finite number of sites. For a finite subset X , let \bar{n}^X denote the block variable configuration with all block variables in X equal to 1 and the rest equal to 0. Then

$$\exp(-\bar{H}(\bar{n}^X)) = \sum_n T(\bar{n}^X, n) e^{-H(n)}$$

Note that $\bar{H}(\bar{n}^\emptyset) = c(\emptyset)$. In particular, $c(\emptyset)$ will grow as the size of the finite volume. The other coefficients $c(Y)$ should have finite limits in the infinite volume limit. If we define $f(X)$ by

$$f(X) = \bar{H}(\bar{n}^X) - \bar{H}(\bar{n}^\emptyset)$$

then $f(X)$ should have a finite limit in the infinite volume limit, and it should be given by

$$f(X) = \sum_{Y: \emptyset \neq Y \subset X} c(Y) \tag{29}$$

The system of equations (29) can be explicitly solved for the $c(Y)$. We claim that for $X \neq \emptyset$,

$$c(X) = \sum_{Y: \emptyset \neq Y \subset X} (-1)^{|X|-|Y|} f(Y) \tag{30}$$

This is a standard inversion trick. To verify (30), define $c(X)$ by (30). Then for a given $X \neq \emptyset$,

$$\begin{aligned} \sum_{Y: \emptyset \neq Y \subset X} c(Y) &= \sum_{Y: \emptyset \neq Y \subset X} \sum_{Z: \emptyset \neq Z \subset Y} (-1)^{|Y|-|Z|} f(Z) \\ &= \sum_{Z: \emptyset \neq Z \subset X} f(Z) \sum_{Y: Z \subset Y \subset X} (-1)^{|Y|-|Z|} \end{aligned} \tag{31}$$

The sum over Y is 1 if $X = Z$. If Z is a proper subset of X , we claim this sum is 0. To see this:

$$\sum_{Y:Z \subset Y \subset X} (-1)^{|Y|-|Z|} = \sum_{W:W \subset X \setminus Z} (-1)^{|W|} = \prod_{i \in X \setminus Z} (1 + (-1)) = 0$$

Thus (31) collapses to $f(X)$.

The important feature of eq. (30) is that the coefficient $c(X)$ only depends on a finite number of free energies $f(Y)$, specifically those with $Y \subset X$. As we will see, these free energies can be computed extremely accurately. So individual coefficients $c(X)$ in the lattice gas variables can be computed extremely accurately. Moreover, this computation does not depend on how many terms we decide to keep in the renormalized Hamiltonian. If we increase the number of terms we keep, then the coefficients we have already computed will not change.

We have shown that in the lattice gas variables there is a natural way to compute the coefficients $c(Y)$ in the expansion (28) for \bar{H} . We now consider the renormalized Hamiltonian in the spin variables:

$$\bar{H}(\bar{\sigma}) = \sum_Y d(Y) \bar{\sigma}(Y) \quad (32)$$

with $\bar{\sigma}(Y) = \prod_{i \in Y} \bar{\sigma}_i$. The sum over Y is over all finite subsets.

We can use $\bar{n}_i = (1 - \bar{\sigma}_i)/2$ to express the spin coefficients $d(Y)$ in (32) in terms of the lattice gas coefficients $c(X)$ in (28).

$$\begin{aligned} \bar{H}(\bar{\sigma}) &= \sum_X c(X) \bar{n}(X) = \sum_X c(X) 2^{-|X|} \prod_{i \in X} (1 - \bar{\sigma}_i) \\ &= \sum_X c(X) 2^{-|X|} \sum_{Y:Y \subset X} (-1)^{|Y|} \bar{\sigma}(Y) \\ &= \sum_Y \bar{\sigma}(Y) (-1)^{|Y|} \sum_{X:Y \subset X} c(X) 2^{-|X|} \\ &= \sum_Y d(Y) \bar{\sigma}(Y) \end{aligned} \quad (33)$$

where the spin coefficients $d(Y)$ are given by

$$d(Y) = (-1)^{|Y|} \sum_{X:Y \subset X} c(X) 2^{-|X|} \quad (34)$$

The problem is that to compute the spin coefficient $d(Y)$ we need the lattice gas coefficients $c(X)$ for infinitely many X , and so we need the free energies $f(X)$ for infinitely many X 's. So we must introduce some sort of approximation.

Let \mathbb{Y}_∞ be a collection of finite subsets of the renormalized lattice such that one set from each translation class is contained in \mathbb{Y}_∞ . We can rewrite (32) as

$$\bar{H}(\bar{\sigma}) = \sum_{Y \in \mathbb{Y}_\infty} d(Y) \sum_t \bar{\sigma}(Y + t)$$

where the sum over t is over the translations for the renormalized lattice. Here $Y + t$ denotes $\{i + t : i \in Y\}$.

Now let \mathbb{Y} be a *finite* subcollection of \mathbb{Y}_∞ . We want to compute an approximation to the above of the form

$$\bar{H}(\bar{\sigma}) \approx \sum_{Y \in \mathbb{Y}} d(Y) \sum_t \bar{\sigma}(Y + t) \quad (35)$$

We will consider two methods which we will refer to as the “partially exact” method and the “uniformly close” method.

For the partially exact method, we compute the free energies $f(X)$ for $X \in \mathbb{Y}$, and then use them to compute the coefficients $c(Y)$ in the lattice gas variables for $Y \in \mathbb{Y}$. We then approximate \bar{H} by

$$\bar{H}(\bar{n}) \approx \sum_{Y \in \mathbb{Y}} c(Y) \sum_t \bar{n}(Y + t)$$

We then convert this Hamiltonian to the spin variables with no approximation. The result is of the form (35) with

$$d(Y) = (-1)^{|Y|} \sum_{X: Y \subset X, X+t \in \mathbb{Y}} c(X) 2^{-|X|} \quad (36)$$

where the notation $X + t \in \mathbb{Y}$ means some translation of X (possibly X itself) is in \mathbb{Y} . Thus this method is equivalent to truncating the exact formula (34) by restricting the sum over X to sets in \mathbb{Y} and their translates. In the lattice gas variables our approximation to \bar{H} agrees with the true \bar{H} for all n^X such that $X \in \mathbb{Y}$. The change from lattice gas to spin variables did not involve any approximation, so our approximation to \bar{H} in the spin variables agrees

exactly with the true \bar{H} for all configurations $\bar{\sigma}^Y$ for $Y \in \mathbb{Y}$. This is the reason for calling this method “partially exact.” It is exact for some of the block spin configurations.

For the uniformly close method let \mathbb{X} be another finite collection of finite subsets which contains at most one set from each translation class. We compute the free energies $f(X)$ for $X \in \mathbb{X}$, i.e., we compute $\bar{H}(\bar{\sigma}^X)$. We define the error of a set of coefficients $\{c(Y) : Y \in \mathbb{Y}\}$ to be

$$\max_{X \in \mathbb{X}} |\bar{H}(\bar{\sigma}^X) - \sum_{Y \in \mathbb{Y}} c(Y) \sum_t \bar{\sigma}^X(Y+t)|$$

where $\bar{\sigma}^X$ is the spin configuration which is -1 on X and $+1$ on all other sites. We then choose the coefficients $c(Y)$ to minimize the above error. This is a standard linear programming problem which we solve by the simplex algorithm. We call this the uniformly close approximation since we have a uniform bound on the difference between our approximation and the exact \bar{H} for the block spin configurations $\bar{\sigma}^X$ for $X \in \mathbb{X}$. (For other X we cannot say anything about how well the approximation does.) If $\mathbb{X} = \mathbb{Y}$, then the partially exact approximation makes the above error zero. We only use the uniformly close approximation for \mathbb{X} which are larger than \mathbb{Y} .

3.5 Philosophical remarks

introducing the block spins should move the system off the critical point

there are many RG's. The goal is not to see how many don't work, but rather to find one that does