

# Physics 115/242

## Monte Carlo simulations in Statistical Physics

Peter Young

(Dated: May 2, 2013)

For additional information on the statistical Physics part of this handout, the first two sections, I strongly recommend *Thermal Physics* by Kittel and Kroemer.

### I. INTRODUCTION TO STATISTICAL PHYSICS

In equilibrium statistical physics (often called “statistical mechanics”) we consider the behavior of a system of a large number of particles in equilibrium with its surroundings (a “heatbath”). The energy of the system is not exactly fixed because it can exchange energy with the bath, and so the state of the system is also clearly not fixed. Statistical mechanics tells us that the system has a certain probability that the system can be in any of the states. The probability of being in state  $|n\rangle$  with energy  $E_n$  is given by the “Boltzmann distribution”

$$P(n) = \frac{1}{Z} e^{-E_n/k_B T}, \quad (1)$$

where  $T$  is the absolute temperature and  $k_B$  is called Boltzmann’s constant. The normalizing factor  $Z$ , given by

$$Z = \sum_{\ell} e^{-E_{\ell}/k_B T}, \quad (2)$$

is called the “partition function”. The partition function provides a convenient link between the microscopic picture involving the states and energy levels and the macroscopic (or thermodynamic) picture, since the free energy of the system,  $F$ , is related to it by

$$F = -k_B T \ln Z. \quad (3)$$

Recall that the free energy is related to the energy  $U$  and entropy  $S$  by

$$F = U - TS. \quad (4)$$

One calculates observable quantities as averages over states with the Boltzmann weight. The average of  $A$  say, (which could be the energy or the magnetization for example) is given by

$$\langle A \rangle = \sum_n P(n) A_n = \frac{\sum_n A_n e^{-E_n/k_B T}}{\sum_{\ell} e^{-E_{\ell}/k_B T}}, \quad (5)$$

where  $A_n$  is value of  $A$  in state- $n$  (to be precise it is the quantum mechanical expectation value  $\langle n|A|n\rangle$ ).

For example, the average energy is given by

$$U = \frac{\sum_n E_n e^{-E_n/k_B T}}{\sum_\ell e^{-E_\ell/k_B T}}, \quad (6)$$

which can be written as

$$U = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta} = \frac{\partial}{\partial \beta} (\beta F), \quad (7)$$

where

$$\beta = 1/k_B T. \quad (8)$$

## II. A TOY MODEL

In order to illustrate the Monte Carlo method it is useful to have a simple example where things can be worked out explicitly. A good model to take is the *Ising* model of magnetism. The magnetic moment (spin) of an individual atom can be one of two possibilities:  $\uparrow$  (“up”) or  $\downarrow$  (“down”). It is convenient to assign a numerical value to these two states by a variable  $S_i$  (for site  $i$ ) which takes value  $+1$  in the up state and  $-1$  in the down state, i.e.

$$S_i = \begin{cases} +1 & (\uparrow) \\ -1 & (\downarrow) \end{cases} \quad (9)$$

There are  $N$  spins,  $S_i$ ,  $i = 1, 2, \dots, N$ , and, in statistical mechanics,

we are interested in the situation where  $N$  is large.

The total number of states is clearly  $2^N$ , which is therefore *enormous*. In general, and not just for this model, it is true that the number of states grows exponentially with the number of degrees of freedom in the problem. This will turn out to be important.

For example, in a solid  $N$  might be of order Avogadro’s number,  $6 \times 10^{23}$ . Two to power of this number is so large that it is impossible even to visualize it. Even if for smaller sizes, to enumerate all the states on a PC would take more than the age of the universe (about  $4 \times 10^{17}$  seconds) for  $N$  greater than about 80, not a very large number. Clearly enumerating all the states is not feasible for large  $N$  ( $2^{80} \simeq 10^{24}$ ). If we are going to determine averages in Eq. (5) numerically we need to do a *sampling* of some (small) fraction of the states. How to do this sampling is the topic of this handout.

We shall discuss extensively the **total magnetization of the system  $M$** , which is given by

$$M = \sum_{i=1}^N S_i. \quad (10)$$

Let us **initially** assume all states have the same energy, so we can forget about the Boltzmann distribution, and average by giving *equal weight* to all states. Later on we will give the states different energies and put in the Boltzmann factor.

We would like to know *how many* states have a particular value of the magnetization  $M$ . First of all we note that the *largest* value of  $M$  is  $N$  and, and the least is  $-N$ . We can get an idea of the *typical* value of  $M$  by calculating averages of  $M$  and  $M^2$  (giving equal weight to all states as mentioned above). Now

$$\langle M \rangle = \sum_i \langle S_i \rangle = 0 \quad (11)$$

since  $S_i$  takes values  $\pm 1$  and so, giving equal weight to these two possibilities, we have

$$\langle S_i \rangle = 0 \quad (12)$$

$$\langle S_i^2 \rangle = 1. \quad (13)$$

The average of  $M^2$  is given by

$$\langle M^2 \rangle = \sum_{i,j} \langle S_i S_j \rangle. \quad (14)$$

Now if  $i \neq j$  the average is zero since the four possibilities for  $S_i S_j$

$$S_i = +1, S_j = +1, \quad S_i S_j = +1 \quad \text{probability } 1/4 \quad (15)$$

$$S_i = -1, S_j = +1, \quad S_i S_j = -1 \quad \text{probability } 1/4 \quad (16)$$

$$S_i = +1, S_j = -1, \quad S_i S_j = -1 \quad \text{probability } 1/4 \quad (17)$$

$$S_i = -1, S_j = -1, \quad S_i S_j = +1 \quad \text{probability } 1/4 \quad (18)$$

average to zero. However, if  $i = j$  then  $\langle S_i^2 \rangle = 1$  (see Eq. (13)) and so Eq. (14) gives

$$\langle M^2 \rangle = N, \quad (19)$$

which means that  $\langle M^2 \rangle^{1/2} = N^{1/2}$ . Hence a *typical* value of  $M$  is of order  $N^{1/2}$ , which, for large  $N$ , is very much less than the maximum value, which is  $N$ .

One can easily write down the total number of states with a given  $M$ . If there are  $N/2 + k$  states with spin up<sup>1</sup> and  $N/2 - k$  states with spin down, then

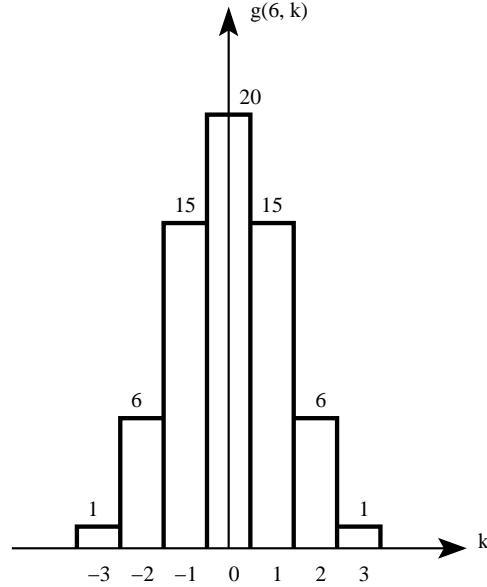
$$M = 2k, \quad (20)$$

where  $k$  runs over *consecutive* integers. Furthermore, the number of states with a given value of the spin imbalance factor  $k$  is equal to the number of ways arranging the up spins (say) among the  $N$  sites. This is given by the binomial coefficient

$$g(N, k) = \frac{N!}{\left(\frac{N}{2} + k\right)! \left(\frac{N}{2} - k\right)!}. \quad (21)$$

For example, for  $N = 6$  we have

$k$	$M$	$g(N, k)$
-3	-6	1
-2	-4	6
-1	-2	15
0	0	20
1	2	15
2	4	6
3	6	1

(22)



---

<sup>1</sup> We assume here that  $N$  is even. As one might expect the difference between even and odd is unimportant for  $N$  large.

The distribution is clearly symmetric, and peaked about zero. From the central limit theorem we expect it to become Gaussian for large  $N$  and, from Eq. (19), have a width of order  $\sqrt{N}$ . This is indeed the case. Taking the log of  $g(N, k)$ , i.e.

$$\ln g(N, k) = \ln(N!) - \ln \left[ \left( \frac{N}{2} + k \right)! \right] - \ln \left[ \left( \frac{N}{2} - k \right)! \right], \quad (23)$$

and using **Stirling's approximation**

$$\ln N! \simeq N \ln N - N + \frac{1}{2} \ln(2\pi N), \quad (24)$$

which is **valid for large  $N$** , one finds, after a bit of algebra and also assuming  $|k| \ll N$ , that

$$g(N, k) = 2^N \sqrt{\frac{2}{\pi N}} \exp \left( -\frac{2k^2}{N} \right), \quad (25)$$

for large  $N$ . As expected, **this is a Gaussian with zero mean and a width of order  $\sqrt{N}$** . However, we should note that far in the tails of the distribution, where  $k \sim N$  (so  $M \sim N$ ), there are corrections to this Gaussian expression.

Summing Eq. (25) over all  $k$ , noting that  $k$  runs over *consecutive* integers, and replacing the sum by an integral<sup>2</sup>, one gets, from Eq. (25),

$$\int_{-\infty}^{\infty} g(N, k) dk = 2^N. \quad (26)$$

This correctly gives the total number of states as  $2^N$ . Note that we have extended the range of integration from  $\pm N/2$  to  $\pm\infty$ , which produces negligible error for large  $N$  because  $g(N, k)$  is tiny for  $|k| > N/2$ . (In fact it is tiny for  $k \gg \sqrt{N}/2$ , which is much less than  $N/2$ .)

Eq. (25) clearly gives  $\langle k \rangle = 0$  (since  $g(N, k)$  is an even function of  $k$ ) and we also find from standard properties of Gaussian integrals that

$$\langle k^2 \rangle = \frac{\int_{-\infty}^{\infty} g(N, k) k^2 dk}{\int_{-\infty}^{\infty} g(N, k) dk} = \frac{\int_{-\infty}^{\infty} k^2 e^{-2k^2/N} dk}{\int_{-\infty}^{\infty} e^{-2k^2/N} dk} = \frac{N}{4}, \quad (27)$$

which, from Eq. (20), implies  $\langle M^2 \rangle = N$ , in agreement with Eq. (19).

It is now useful to consider **the magnetization per spin**

$$m = \frac{M}{N} = \frac{2k}{N} = \frac{1}{N} \sum_{i=1}^N S_i. \quad (28)$$

The largest possible value of  $m$  is 1, and yet we see from Eqs. (25) and (28) that for large  $N$  the overwhelming majority of states have  $m$  very close to zero, more precisely within of order  $1/\sqrt{N}$  of

---

<sup>2</sup> This should be valid if  $N$  is large since  $g(N, k)$  then only changes by a small fraction of itself when  $k$  changes by 1.

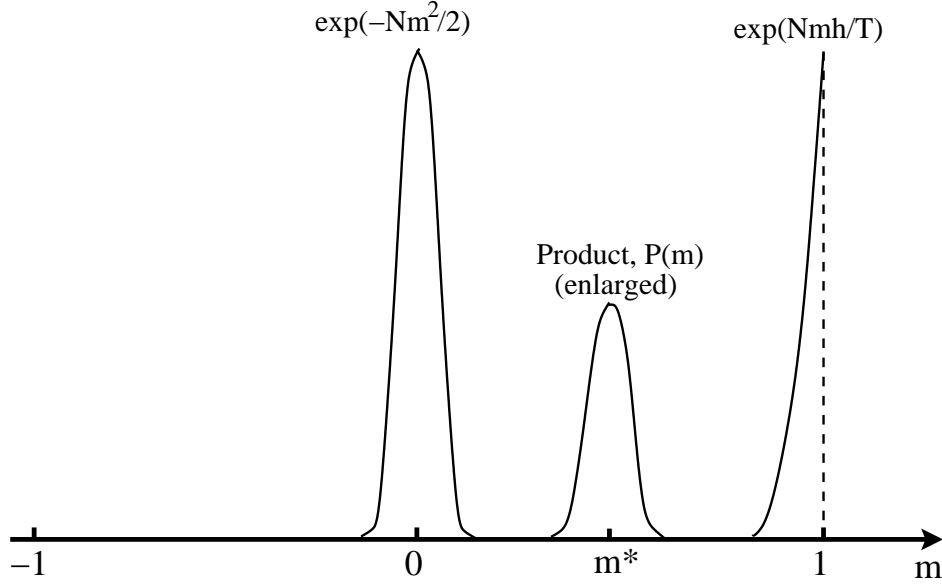


FIG. 1: The probability,  $P(m)$ , that the system has magnetization per spin  $m$  is a product of two factors: (i) the Boltzmann factor of a single state with magnetization  $m$ ,  $\exp(Nmh/k_B T)$ , which *increases* rapidly as  $m$  increases, and (ii) a multiplicity factor (the number of states with this value of  $m$ ),  $\propto \exp(-Nm^2/2)$ , which *decreases* rapidly as  $m$  increases from zero. The equilibrium value of  $m$ , labeled  $m^*$  here, is where the product has a maximum. **It is very important that you understand this figure.**

zero. This result will be important in developing a Monte Carlo method, since it shows that states generated randomly must have  $m$  very close to zero. For large  $N$  one would have to generate a huge number of configurations at random to find even one which had  $m$  significantly different from zero.

**Next** we make the problem a bit less trivial by giving the states some energy. The simplest case, for which the model is still easily exactly solvable, is to include a magnetic field,  $h$ . The Hamiltonian (energy) is then given by

$$\mathcal{H} = -h \sum_i S_i = -hM. \quad (29)$$

The Boltzmann weight of a state of magnetization  $M$  is therefore proportional to  $\exp(\beta h M) = \exp(N\beta h m)$ , where  $\beta$  is defined in Eq. (8). Hence *individual* states with a larger  $m$  have a *larger* Boltzmann weight than those with smaller  $m$ . Because the energy is proportional to  $N$ , the Boltzmann factor varies *hugely* depending on the value of  $m$ . However, the *number* of states is sharply peaked near  $m = 0$  and so *decreases* with increasing  $m$  (for  $m > 0$ ).

The probability that the system has a particular value of the magnetization per spin  $m$  is the

*total* Boltzmann weight of *all* states with that value of  $m$ . This is a product of two factors: (a) the *number* of states with that value of  $m$ , Eq.(25) with  $k = Nm/2$ , and (b) the Boltzmann factor of *one* such state. Hence the probability of the system having a particular value of  $m$  is given by

$$P(m) \propto \exp \left[ -N \left( \frac{m^2}{2} - \beta h m \right) \right]. \quad (30)$$

This has a maximum at  $m = m^*$  where<sup>3</sup>

$$m^* = \beta h. \quad (31)$$

Note that the maximum arises because  $P(m)$  is a product of  $\exp(-Nm^2/2)$ , which *decreases* rapidly as  $m$  increases, and  $\exp(N\beta h m)$  which *increases* rapidly as  $m$  increases, see Fig. 1.  $P(m)$  is *sharply peaked* at  $m^*$  as we can see by expanding Eq. (30) up to second order in  $\delta m = m - m^*$ :

$$P(m) = P(m^*) \exp \left[ -N(\delta m)^2/2 \right], \quad (32)$$

This shows that the width of  $P(m)$  is of order  $1/\sqrt{N}$ , and so all the states that a large system visits (in equilibrium) in a field  $h$  have magnetization very close to  $m^*$ .

To summarize we have seen for our simple model, the Ising model in which the energy comes entirely from the external field, that

1. At  $T = \infty$ , which corresponds to equal weighting of all states, a large system only visits states with  $m$  close to zero, the left hand peak in the above figure.
2. At finite  $T$ , a large system only visits states which have  $m$  close to a particular value  $m^* = \beta h$ , the middle peak in the above figure.

It turns out that these results are not special to the toy model and analogous results hold quite generally. In particular:

1. At *infinite temperature* the states visited by the system will have energy close to a *certain value* (corresponding to the energy where most of the states lie, often taken to be zero).
2. At a *finite temperature* the system will visit states with energy in the vicinity of a *different value*.

---

<sup>3</sup> If we go beyond the Gaussian expression for  $g(N, k)$  in Eq. (25), we find that  $m^*$  is given by  $m^* = \tanh(\beta h)$  rather than Eq. (31). Hence there are corrections to Eq. (31) unless  $\beta h \ll 1$ , i.e. unless  $M \ll N$ . Nonetheless the structure of Fig. 1 is still valid.

### III. THE ISING MODEL WITH INTERACTIONS

The example in the previous section is very simple. Each spin is independent of other spins. One can make the model a much richer model for magnetism by including interactions between nearby sites. We suppose that the spins are on a lattice, which could be simple cubic lattice in three-dimensions, a square lattice in two-dimensions, or a chain in one-dimension, and the energy depends on the relative orientation of neighboring pairs, i.e.

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i S_j - h \sum_i S_i, \quad (33)$$

where  $J$  is the interaction strength and the sum  $\langle i,j \rangle$  is over nearest neighbor pairs. This model is known as the Ising model. If  $J > 0$  the spin prefer to align parallel at low temperature, and such a system is said to be “ferromagnetic”. If  $J < 0$  neighboring spins prefer to point in opposite directions at low temperature, and such systems are said to be “antiferromagnetic”. In the rest of this handout we will consider only the case of zero field,  $h = 0$ .

In general, the problem is now highly non-trivial. Only in one-dimension is there a fairly simple exact solution, but unfortunately this model has no transition. In two dimensions, some quantities like the energy and magnetization in zero field can be calculated exactly, but only with monumental effort, while in in three dimensions there are no exact solutions.

In two and three dimensions, and higher, there is a transition at a finite temperature  $T_c$  where the spins align spontaneously even in the absence of an external field  $h$ . If  $J > 0$  the spins align parallel and give rise to a state with a macroscopic magnetization. This is then a very simple model for the magnetism of iron. The magnetization per spin  $m$  increases from zero as  $T$  is reduced below  $T_c$  and, in this model, tends to unity as  $T \rightarrow 0$ . This is shown in Fig. 2.

In one dimension there is no transition at a finite  $T$ , though at  $T = 0$  all the spins are aligned so one often talks about a transition at zero temperature.

### IV. MONTE CARLO TECHNIQUES

In my opinion, the best book on Monte Carlo simulations in statistical physics is *Monte Carlo Simulations in Statistical Physics* by Newman and Barkema. Chapters 2 and 3 cover the basic material. Another good book is *A Guide to Monte Carlo Simulations in Statistical Physics* by David Landau and Kurt Binder. See also the first 3-1/2 pages of the article on Monte Carlo on my web site [http://physics.ucsc.edu/~peter/converge\\_new.pdf](http://physics.ucsc.edu/~peter/converge_new.pdf).



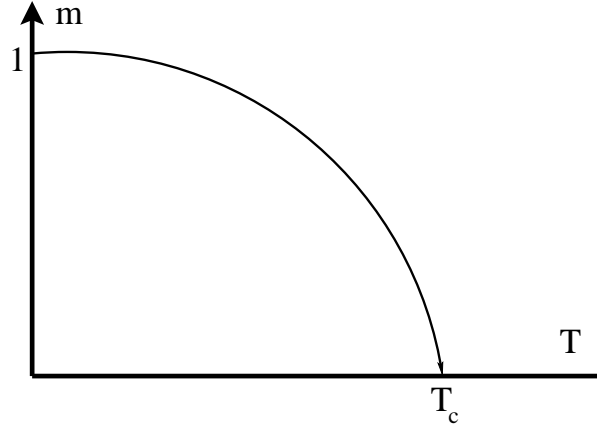


FIG. 2: A sketch of the magnetization per spin as a function of temperature for the Ising model in dimension  $d$  greater than or equal to 2.

Since there are no exact solutions in three dimensions, it is useful to be able to simulate models like the Ising model numerically. In particular we would like to calculate averages as in Eq. (5). As we’ve said, except for very small  $N$ , there are too many states ( $2^N$ ) to enumerate them all, so we need to do some sort of “random” sampling<sup>4</sup>. However,

we can’t do *truly* random sampling, which we used earlier in the class for Monte Carlo integration, because, as discussed in the previous section, the vast majority of states generated *do not have the correct energy* for the (finite) temperature being considered, see Fig. 1.

The procedure is therefore to generate states *with the Boltzmann probability distribution itself*. How to do this will occupy us for the rest of this handout, but first we assume that we *can* do it and discuss how to calculate averages like  $\langle A \rangle$  in Eq. (5). The correct Monte Carlo average is to give *equal weight* to the generated configurations,

$$\langle A \rangle \simeq \frac{1}{M_{\text{meas}}} \sum_{\ell=1}^{M_{\text{meas}}} A_{\ell}, \quad (34)$$

where  $M_{\text{meas}}$  is the number of “measurements”. Although the Boltzmann factor in Eq. (5) does not appear *explicitly* in Eq. (34), it *is* present *implicitly* because the states are generated with

<sup>4</sup> This is known as Monte Carlo sampling by analogy with the randomness of the roulette wheels in the casinos in Monte Carlo.

the Boltzmann probability. Note that the average in Eq. (34) is like a “time average” in real experiments.

Now we discuss how to generate states with the Boltzmann probability. An iterative procedure is required. Starting from an initial state “ $l_0$ ”, which may *not* be a state with the correct energy for the temperature being considered, we generate states  $l_1, l_2, l_3$  etc. such that eventually, after an “equilibration period”, states *are* generated with the Boltzmann distribution for the specified temperature. It is convenient to think of this sequence as representing the state of the system at successive “times”  $t$ .

Suppose the system is in state  $\ell$  with energy  $E_\ell$  at time  $t$ . What will the state be at time  $t + 1$ ? The procedure is to choose a *trial state*  $m$ , with energy  $E_m$ , which typically differs from  $\ell$  by flipping a single spin. The state at time  $t + 1$  is then taken to be  $m$  with a certain probability  $w_{\ell \rightarrow m}$  (discussed below), and otherwise the system stays in the old state  $\ell$ . We start the system off in some initial state  $\ell_0$  say, so

$$P_\ell(0) = \delta_{\ell, \ell_0} , \quad (35)$$

and require that, at long times, the  $P_\ell(t)$  equal the equilibrium, Boltzmann, distribution  $P_\ell^{\text{eq}}$ :

$$\lim_{t \rightarrow \infty} P_\ell(t) = P_\ell^{\text{eq}}. \quad (36)$$

Let’s see how this can be achieved with an appropriate choice of the  $w_{\ell \rightarrow m}$ . Consider how the probability for the system to be in a given state changes with time. If  $P_\ell(t)$  is the probability that the system is in state  $\ell$  at time  $t$ , then

$$\boxed{P_\ell(t + 1) - P_\ell(t) = \sum_{m \neq \ell} [P_m(t)w_{m \rightarrow \ell} - P_\ell(t)w_{\ell \rightarrow m}]}, \quad (37)$$

which is known as the *master equation*. The first term on the RHS describes transitions at time  $t + 1$  into state  $l$  from state  $m$ , which increases  $P_l$ , and the second describes transitions out of  $l$  into  $m$ . Note that the probabilities at time  $t + 1$  only depend on the probabilities at time  $t$ , and not on the probabilities at earlier times. This is called a “Markov Process” and the sequence of configurations thus generated is called a “Markov Chain”.

Clearly a *necessary* condition for Eq. (36) to be valid is that if the distribution is the equilibrium one at time  $t$  then it remains the equilibrium distribution at time  $t + 1$ . From Eq. (37), this requires

$$\sum_{m \neq \ell} [P_m^{\text{eq}} w_{m \rightarrow \ell} - P_\ell^{\text{eq}} w_{\ell \rightarrow m}] = 0. \quad (38)$$

For convenience, we usually require that *each* term in the sum vanishes separately, which leads to the important *detailed balance condition*

$$\boxed{\frac{w_{\ell \rightarrow m}}{w_{m \rightarrow \ell}} = \frac{P_m^{\text{eq}}}{P_\ell^{\text{eq}}} = \exp(-\beta(E_m - E_\ell))}. \quad (39)$$

However, the detailed balance condition does not determine the  $w$  uniquely, only the ratio of  $w$  for a transition to that for the inverse transition. Naively, one might think that

$$w_{\ell \rightarrow m} = A \exp(-\beta(E_m - E_\ell)/2) \quad (\text{wrong!}) \quad (40)$$

would work, where  $A$  is a constant. However, this is not correct, because the probabilities  $w$  cannot be greater than one, and this would be violated if  $\beta(E_\ell - E_m)/2 > \ln(1/A)$ . A common choice, which keeps the probabilities no greater than one and satisfies detailed balance, is the *Metropolis* updating probability:

$$\boxed{w_{\ell \rightarrow m} = \begin{cases} \exp(-\beta(E_m - E_\ell)), & \text{if } E_m - E_\ell > 0 \\ 1, & \text{otherwise,} \end{cases} \quad (\text{Metropolis probability})} \quad (41)$$

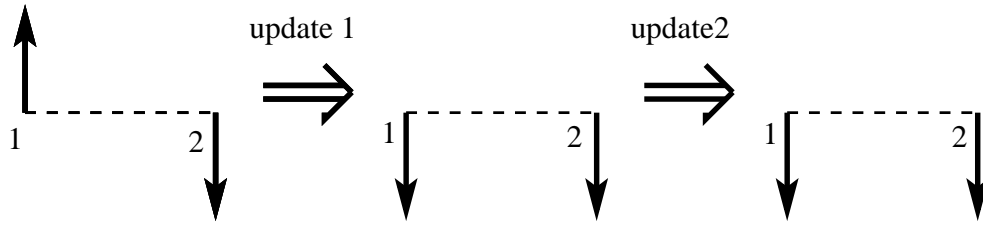
It is easy to see that Eq. (39) is satisfied because, if  $E_m > E_\ell$  for example,  $w_{\ell \rightarrow m} = \exp(-\beta(E_m - E_\ell))$  and  $w_{m \rightarrow \ell} = 1$ , while if  $E_m < E_\ell$  then  $w_{\ell \rightarrow m} = 1$  and  $w_{m \rightarrow \ell} = \exp(-\beta(E_\ell - E_m))$ . In the Metropolis updating rule we always accept the move if the system gains energy but, accept the move with a probability less than unity if it costs energy. Intuitively this means there is a greater likelihood of the system being in states of lower energy, as required by the Boltzmann distribution.

Another common choice is the “heatbath” probability, where, irrespective of the sign of  $E_m - E_\ell$ , the spin flip occurs with probability

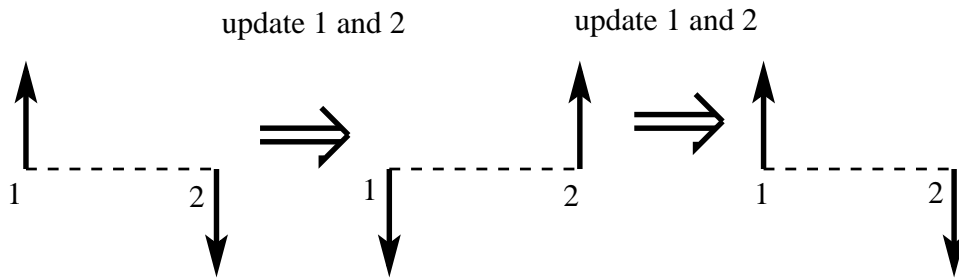
$$\boxed{w_{\ell \rightarrow m} = \frac{1}{\exp[\beta(E_m - E_\ell)] + 1} \quad (\text{Heat bath probability})}. \quad (42)$$

This satisfies the detailed balance because, writing  $f(x) = (\exp(x) + 1)^{-1}$ , we have  $f(x) = \exp(-x)f(-x)$ .

Note that that if two spins have an interaction between them one can not update both of them simultaneously (in parallel). One must first test one of them, and flip it if necessary which changes the probability of flipping the second. Only then can the second spin be updated. Flipping both in parallel is wrong as can be seen from the example of two spins with a strong ferromagnetic interaction between them at low temperature (so the spins should be parallel most of the time). Let’s suppose we start the spins antiparallel, see the figure below.



Sequential (correct). Spins become parallel.



Parallel (incorrect). Spins stay antiparallel and keep flipping together.

If we update sequentially, then the first update (of spin 1) puts the spins parallel. Subsequently updating spin 2, this spin does not flip with high probability, and so the spins stay parallel as desired. This is shown in the first row of the figure. However, the second row of the figure shows what happens if we update in parallel. Both spins calculate that they will both gain energy by flipping and so they both flip, which leaves them antiparallel. Continuing, both spins flip every move while maintaining an antiparallel orientation. This is wrong.

However, spins that do not have an interaction *can* be flipped in parallel, because the probability to flip one is independent of the orientate of the other. This is exploited in Monte Carlo algorithms designed to run on parallel computers.

If the detailed balance condition is satisfied then, once the system has reached equilibrium, it will stay in equilibrium. However, it is also necessary to show that the system will *converge* to equilibrium starting from some arbitrary initial distribution like Eq. (35). Most of the proofs of this are quite mathematical. However, a simple proof is given in the article on my web site quoted above; see also Narayan and Young, Phys. Rev. E **64**, 021104 (2001), (also available at <http://arxiv.org/abs/cond-mat/0008046>). Convergence occurs for an algorithm satisfying detailed balance provided, in addition, it is “ergodic”, which means that starting from any state one can get to any other state in a finite number of moves. Physically ergodicity means that the system won’t get trapped in just part of the configuration space.

Here, for simplicity, we won't go through the proof and will just *assume* that convergence occurs if the algorithm is ergodic and satisfies the detailed balance condition.

We can now describe how to implement a Monte Carlo simulation for the Ising model:

1. We denote the testing of a spin, and flipping it with the probability in Eq. (41) or Eq (42), as a single “**update**”. The basic unit of Monte Carlo time, however, is the “**sweep**”, which consists of  $N$  updates, where  $N$  is the number of spins. Usually one performs a sweep by going *sequentially* through the lattice in which case each spin is updated once. (Note “updated” doesn't necessarily mean flipped). Alternatively, the sites can be chosen at random (in which case, during one sweep, most spins will be updated once but some spins will be updated twice, or more, and some not at all, according to a Poisson distribution with mean one.) Both approaches work fine, but sequential updating is simpler so I recommend it.

To update the spin at site  $i$ , we first compute the energy  $\Delta E$  to flip it. This is given in terms of  $S_i$  and its neighbors  $S_j$ , by

$$\Delta E = 2JS_i \sum_{j \text{ neighbor of } i} S_j, \quad (43)$$

where the sum is over the neighbors of  $i$ . We then generate a random number  $r$  with a uniform distribution between 0 and 1, and flip  $S_i$  if

$$r < \exp(-\beta\Delta E), \quad (\text{Metropolis}), \quad (44)$$

$$r < \frac{1}{\exp(\beta\Delta E) + 1}, \quad (\text{Heatbath}). \quad (45)$$

Note that if  $\Delta E < 0$  the condition in Eq. (44), the Metropolis approach, is automatically satisfied (since the random number  $r$  can not be bigger than 1) so it is not necessary to do an `if` statement to test for the sign of  $\Delta E$ . However, if one does this test and  $\Delta E$  is found to be negative, then it is not necessary to generate the random number or compute the exponential. However, the extra CPU time in doing the `if` statement probably means there is little benefit in including it.

2. Since the system is not initially in equilibrium, perform  $M_{\text{drop}}$  initial sweeps during which no measurements are made, in order to equilibrate it.
3. Then perform  $M_{\text{run}}$  sweeps during which measurements are carried out. One might do a measurement after every sweep, in which case the number of measurements  $M_{\text{meas}}$  is equal

to  $M_{\text{run}}$ , or possibly only after  $n (> 1)$  sweeps, in which case  $M_{\text{meas}} = M_{\text{run}}/n$ . Averages are then given by Eq. (34).

This type of sampling, in which we don't choose points completely at random but choose them predominantly in the important region of configuration space, is called *importance sampling*. It was introduced into statistical physics by Metropolis et al. (1953). In this approach we generate a sequence of states in a probabilistic manner, in which the probability of a state occurring at some point in the sequence only depends on the *previous* state, but not on the earlier history of the system. Such a sequence of states is called a “Markov Chain” in the mathematics literature.

As usual we need to estimate error bars. This is not as simple in our earlier class discussion of error bars data with noise, see the handout [http://physics.ucsc.edu/~peter/115/error\\_bar.pdf](http://physics.ucsc.edu/~peter/115/error_bar.pdf). There we used the fact that the different data was *statistically independent*. (Remember that two random variables  $X$  and  $Y$  are statistically independent if the value of one does not affect the value of the other so, for example,  $\langle X Y \rangle = \langle X \rangle \langle Y \rangle$ .) In that case we showed that, if we have  $M_{\text{meas}}$  estimates  $X_\alpha$ ,  $\alpha = 1, 2, \dots, M_{\text{meas}}$ , the estimate for  $\mu$ , the average of  $X$ , is just the average over the measurements,

$$\mu \simeq \bar{X} = \frac{1}{M_{\text{meas}}} \sum_{\alpha=1}^{M_{\text{meas}}} X_\alpha, \quad (46)$$

and the error bar in the average is given by

$$\sigma_\mu = \frac{s}{\sqrt{M_{\text{meas}} - 1}}, \quad (47)$$

where

$$s^2 = \frac{1}{M_{\text{meas}}} \sum_{\alpha=1}^{M_{\text{meas}}} (X_\alpha - \bar{X})^2, \quad (48)$$

is the variance of the  $M_{\text{meas}}$  values  $X_\alpha$ . However, the relationship (47) between  $s$  and the error  $\sigma_\mu$  is only true for statistically independent data. For our Monte Carlo simulations there will be a correlation between  $X(t_0)$  and  $X(t_0 + t)$  up some time  $\tau$  called the *relaxation time*. (We measure time in units of a sweep from now on.) Roughly speaking, we should get  $\sigma_\mu$  from  $s$  in Eq. (47) by dividing by  $\sqrt{M_{\text{meas}}/\tau}$  rather than by  $\sqrt{M_{\text{meas}}}$  (forgetting about the usually unimportant factor of  $-1$ ).

In order to estimate the error bar from the simulation we either need to estimate  $\tau$  precisely (see below) or just use statistically independent data. The latter approach, which is the simpler, is

usually obtained by “binning”. We divide our  $M_{\text{meas}}$  measurements into  $p$  groups or “bins”. The number of sweeps corresponding to each bin is  $M_{\text{run}}/p$ . If this is much larger than  $\tau$  there won’t be much correlation between averages obtained from different bins. (Only a small amount due to the data at the end of one bin being correlated with that at the beginning of the next bin.) Hence we estimate the error bar from Eqs. (46)–(48) in which  $X_\alpha$  refers to the average of the data from a single bin, and  $M_{\text{meas}} = p$ , the number of bins.

Knowing how many sweeps to include in a bin requires a rough estimate of  $\tau$ . For the purposes of the Monte Carlo project this is not essential and any plausible assumption will do. However, for professional work one could compute the “correlation function”

$$C(t_0, t_0 + t) = \langle X(t_0)X(t_0 + t) \rangle - (\overline{X})^2, \quad (49)$$

where the average here is taken over different values of the starting time,  $t_0$ , in a given run. The “time-dependent correlation function”  $C(t_0, t_0 + t)$  will tend to zero for  $t > \tau$ .

A simpler approach to getting the error bar, **which I recommend for the Monte Carlo project** is to do  $n$  *completely independent* runs, starting from different initial states  $\ell_0$ , and allowing each run to equilibrate. The data from the different runs will certainly be uncorrelated. A small disadvantage with this approach is that the initial  $M_{\text{drop}}$  sweeps for equilibration, have to be done separately for each run, which is a bit wasteful. However, for the purposes of this class, the extra CPU time is more than compensated for by the simplicity of this approach to get statistically uncorrelated data.

Note that in practice one can do this in one long run. One initializes the random number once, and then repeats the following procedure  $n$  times (where  $n$  is the number of statistically independent results that will be obtained): (i) initialize the spins, (ii) equilibrate for  $M_{\text{drop}}$  sweeps, and (iii) do  $M_{\text{run}}$  sweeps during which measurements are carried out.