

# Problems in Statistical Mechanics

## 1 Statistical Mechanics

Now we will apply the method of Monte Carlo simulation (any computer simulation that uses random numbers to simulate a random physical process) to statistical mechanics. Statistical mechanics applies statistical methods and probability theory to large ensembles of microscopic objects, and tries to explain macroscopic behavior from the behavior of the ensemble. The fundamental problem is to calculate the average value of a quantity in a system in thermal equilibrium at temperature  $T$ . The exact state of the system is unknown. The probability of state  $i$  with energy  $E_i$  is given by the Boltzmann formula

$$P_i = \frac{e^{-E_i/(k_B T)}}{Z} \quad (1)$$

$$Z = \sum_i e^{-E_i/(k_B T)} \quad (2)$$

where  $k_B$  is the Boltzmann constant and the sum in  $Z$  (the partition function) is over all possible states. Then the average value of some quantity  $X$  that takes the value  $X_i$  in state  $i$  is given by the weighted average over states (weighted by probability):

$$\langle X \rangle = \frac{\sum_i X_i P_i}{\sum_i P_i} = \sum_i X_i P_i \quad (3)$$

where the sum again is over all possible states. (The sum in the denominator is one because we are summing over all states, i.e. the system is guaranteed to be in some state.) In practice, summing over the states of a real physical system is not possible. Consider one mole of a gas, which contains  $10^{23}$  molecules. If each molecule had only two possible states, then the total number of states would be  $2^{10^{23}}$ , which is an enormous number! And there are many more than two states for each molecule.

You might think one solution is to randomly choose some number of states  $N$  and calculate the average over those  $N$  states to approximate the average:

$$\langle X \rangle \approx \frac{\sum_{k=1}^N X_k P_k}{\sum_{k=1}^N P_k} \quad (4)$$

However, this doesn't work; in reality, most states have a probability of near zero, and only a small fraction of states contribute significantly to the sum. If we choose  $N$  states completely

---

at random, we are unlikely to choose the states that contribute the most, and our estimate of the average given by Equation 4 will be bad. (In this case, the denominator does not sum to one since we are not summing over all states, i.e. the system is not guaranteed to be in one of the randomly chosen states.)

Instead of choosing states randomly from a uniform distribution (all states have equal probability of being chosen), we do something called “importance sampling,” and choose states based on their Boltzmann probability,  $P_i$ . I won’t go through the math (see the textbook), but it can be shown that if you choose the states based on this probability, your approximate average will be

$$\langle X \rangle \approx \frac{1}{N} \sum_{k=1}^N X_k \quad (5)$$

That is, if you choose the states appropriately (weighted by their Boltzmann probability), you can just take a straight average of your calculated  $X_k$ ’s and get something that is close to the true value given by Equation 3. The larger  $N$  is (i.e. the more states you choose), the better the approximation will be, but simply by using importance sampling you can get a good estimate with only a few states (because you are choosing the most important states).

So we need to choose states based on the Boltzmann probability of each state. The complication now is that in order to get the Boltzmann probability, we need to calculate the partition function  $Z$ , which is a sum over all states. If doing that sum were easy, we wouldn’t need the Monte Carlo simulation in the first place. Fortunately, we can get around this by using something called a “Markov chain.” Using the Markov chain, we don’t need the Boltzmann probability of each state, only the ratio of probabilities. In the ratio, the partition function cancels out, saving us from needing to calculate it.

We generate a series of states (the Markov chain). We pick some state as a starting point, and to get to the next state, we make some small change to the current state. The choice of the new state is determined by a transition probability that gives the probability of changing from state  $i$  to state  $j$ . If we choose the transition probabilities in just the right way, we can guarantee that the probability of choosing any one state is given by the Boltzmann probability (required in order to use Equation 5 to calculate  $\langle X \rangle$ ), without actually having to know the Boltzmann probability (which would require calculating  $Z$ ).

Let  $T_{ij}$  be the transition probability, the probability of changing from state  $i$  to state  $j$ . First we require that

$$\sum_j T_{ij} = 1 \quad (6)$$

where the sum  $j$  is over all the states. This just means that at every step of the Markov chain, we are guaranteed to start in state  $i$  and end up in any state  $j$ .

---

Second we require that

$$\frac{T_{ij}}{T_{ji}} = \frac{P_j}{P_i} = \frac{e^{-E_j/(k_B T)}/Z}{e^{-E_i/(k_B T)}/Z} = e^{-(E_j - E_i)/(k_B T)} \quad (7)$$

We are choosing a particular value for the ratio of the probability to go from  $i$  to  $j$  to the probability to go from  $j$  to  $i$ .

To understand why we choose the ratio above, let's consider the probability of choosing some state  $j$ . We know the probability of going to state  $j$  from state  $i$  is  $T_{ij}$ . If we assume that the probability of choosing any state  $i$  is the Boltzmann probability  $P_i$ , then the probability of choosing some state  $j$  is

$$\begin{aligned} \sum_i T_{ij} P_i &= \sum_i T_{ji} P_j \quad (\text{by Equation 7}) \\ &= P_j \sum_i T_{ji} \\ &= P_j (1) \quad (\text{by Equation 6}) \\ &= P_j \end{aligned} \quad (8)$$

This tells us that if the probability of choosing state  $i$  is the Boltzmann probability  $P_i$ , then the probability of choosing state  $j$  next using the Markov chain is the Boltzmann probability  $P_j$ . Once we are able to choose a starting state with the Boltzmann probability, the next states in the Markov chain are guaranteed to be chosen with the Boltzmann probability. So how do we choose the *first* state in the chain with the Boltzmann probability? We don't! It turns out that even if the starting point is randomly chosen (i.e. from a uniform distribution of states, all equally likely), as long as the transition probabilities obey the two rules, then after enough steps in the chain, the system will converge to a Boltzmann distribution. Again, the math is beyond us here, but a proof is shown in Appendix D of the textbook.

Bottom line: If we find a set of transition probabilities  $T_{ij}$  that satisfy the two conditions given by 6 and 7, then the probability of selecting state  $j$  will converge to the Boltzmann probability  $P_j$  after enough steps in the chain. When that's true, estimating  $\langle X \rangle$  can be done via Equation 5. The appropriate value for  $T_{ij}$  will be given later when we talk about the specific algorithm.

We are going to use a Monte Carlo simulation to model a ferromagnet in the Ising Model.

## 2 Ferromagnetism

All magnetic phenomenon are due to electric charges in motion: inside a piece of magnetic material are tiny currents caused by the spin and orbit of atomic electrons. Most of the time, all those tiny currents cancel each other out because the atoms are oriented randomly. But when a magnetic field is applied, the magnetic dipole moments become aligned and the material becomes **magnetized**. A few substances retain their magnetization even after the

---

external field has been removed. These materials are called ferromagnets.

Each electron is a magnetic dipole due to its quantum mechanical spin. An electron can be “spin up” with a magnetic dipole moment that points in the  $+z$  direction or “spin down” with a magnetic dipole moment that points in the  $-z$  direction.

In a ferromagnetic material, each dipole “likes” to point in the same direction as its neighbors, due to a quantum mechanical effect called the exchange interaction. The electrostatic energy of the electrons when their spins are aligned is less than the energy of the electrons when their spins are opposite; therefore having aligned spins is a more stable state. The alignment occurs in small patches, called domains. The domains themselves are randomly oriented. An object is made up of many domains, and the fields typically cancel.

A magnetic dipole experiences a torque in a magnetic field. This torque will cause the dipole to rotate until the dipole is aligned with the magnetic field. To produce a permanent magnet, you can put a piece of iron in a strong magnetic field. Since the dipoles tend to stay parallel to their neighbors, most of the dipoles resist the torque that would align them with the field. However, at the boundary between two domains, the dipoles have neighbors in each domain. The dipoles at the boundaries tend to align with the domain most nearly parallel to the field. The net effect of the magnetic field is to move the domain boundaries. Domains that are parallel to the field grow and the others shrink. If the field is strong enough, one domain takes over entirely, and the iron is **saturated**. When the field is turned off, most of the domains remain in this direction, creating a permanent magnet.

Very high temperatures destroy the alignment of the domains. This occurs at a specific temperature called the **Curie point**, which is  $770^\circ\text{C}$  for iron.

### 3 The Ising Model

We will model a ferromagnet as a collection of magnetic moments which we can think of as being atoms with magnetic moments due to quantum mechanical spin. We will assume each spin is located on a regularly-spaced 2D lattice. Each spin points in the  $+z$  or  $-z$  direction,  $s_i = \pm 1$ . Each spin can interact with other spins in the lattice; in this model we will assume that only the nearest neighbors interact with each other. (In a real magnetic material, the interaction is the largest with nearest neighbors, and then falls off rapidly with increasing separation.) The energy of this system is given by

$$E = -J \sum_{(ij)} s_i s_j \quad (9)$$

where the sum is over pairs of nearest neighbors only and  $J$  is called the exchange constant and is assumed to be positive. The nearest neighbors are one unit up, down, left, and right. (Diagonal neighbors are not included, since the distance is more than one unit.)

---

According to this function, two neighboring spins will have a energy of interaction  $-J$  if they are parallel and  $+J$  if they are antiparallel. Since  $J$  is positive, this means that parallel alignment is favored because it's a lower energy state. If each spin is parallel to its neighbors, then every spin on the lattice will be parallel. This would lead to a nonzero (macroscopic) magnetic moment and yield a ferromagnet. A system that has a magnetic moment in the absence of a magnetic field is said to have spontaneous magnetization.

If we add an external magnetic field to the system, the energy becomes

$$E = -J \sum_{(ij)} s_i s_j - \mu H \sum_i s_i \quad (10)$$

where  $H$  is the magnetic field (there are reasons for calling it  $H$  instead of  $B$  which you will understand once you take upper-level E&M; for our purposes here it's not important), and  $\mu$  is the magnetic moment associated with each spin.

We will assume our spin system is in equilibrium with a heat bath at temperature  $T$ . (You can think of the “heat bath” as just the environment surrounding the system. The environment is large enough that taking energy or giving energy to our spin system has a negligible effect on the temperature of the environment.) For a system in equilibrium with a heat bath, the probability of finding the system in any particular state is proportional to the Boltzmann probability:

$$P_i = \frac{e^{-E_i/(k_B T)}}{Z} \quad (11)$$

Each state  $i$  is a particular configuration of spins, which we will call a “microstate” of the system.

The interaction of the spin system with the heat bath is what causes the system to transition from one microstate to another. Individual spins can flip from  $+1$  to  $-1$  and vice versa as they gain energy from or lose energy to the heat bath. This interaction is what allows us to move along our Markov chain from one microstate to another.

Let's consider two different microstates. Let's call the energies of these two microstates  $E_1$  and  $E_2$  and assume that  $E_1 > E_2$ . According to Equation 11, the ratio of the probabilities to find the system in microstate 1 to microstate 2 is

$$\begin{aligned} \frac{P_1}{P_2} &= \frac{e^{-E_1/(k_B T)}/Z}{e^{-E_2/(k_B T)}/Z} \\ &= \frac{1}{e^{E_1/(k_B T)} e^{-E_2/(k_B T)}} \\ &= \frac{1}{e^{(E_1 - E_2)/(k_B T)}} \\ &= e^{-(E_1 - E_2)/(k_B T)} \end{aligned} \quad (12)$$

This is the relative probability of finding the system in microstate 1 compared to microstate 2. In our Monte Carlo simulation, we will use this probability to determine if the system changes from one microstate to another.

## 4 Simulation: Calculating the energy of a microstate

### Section 6, Exercise 1

For now, let's assume  $H = 0$  and just worry about the energy due to the interaction between pairs of spins. We need to know how to calculate

$$E/J = - \sum_{(ij)} s_i s_j \quad (13)$$

where pairs  $(ij)$  are neighboring spins on the lattice in the horizontal and vertical directions (not diagonal). So how many nearest neighbor pairs are there in an  $n \times n$  lattice?

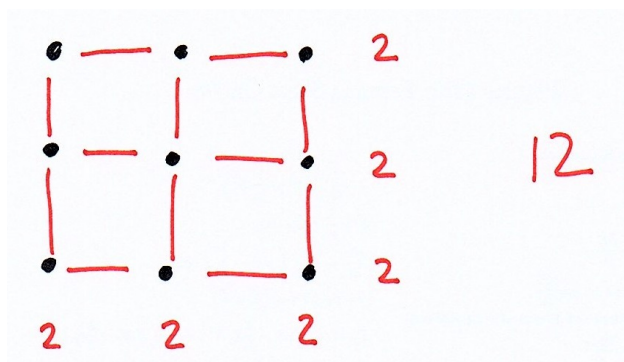


Figure 1:

See Figure 1. For a  $3 \times 3$  lattice, there are two neighboring pairs in each row and two neighboring pairs in each column for  $2 \times 3 + 2 \times 3 = 12$  total pairs. If you can imagine extending this to an  $n \times n$  lattice, there would be  $n - 1$  neighboring pairs in each of  $n$  rows plus  $n - 1$  neighboring pairs in each of  $n$  columns for  $(n - 1)n + (n - 1)n = 2(n - 1)n$  total neighboring pairs. (For the  $3 \times 3$  case, this works out:  $2(3 - 1)(3) = 12$ .)

To perform the sum in Equation 13, let the lattice be represented by a matrix  $L_{ij}$ , where  $i$  represents the row and  $j$  represents the column.

- For each row  $i$ , loop through the columns and sum the product of the entry for the column and the entry for the column one unit to the right (except for the last column)
- For each column  $j$ , loop through the rows and sum the product of the entry for the row and the entry for the row one unit below (except for the last row)

We can write it this way:

$$E/J = - \left( \sum_{i=1}^n \sum_{j=1}^{n-1} L_{ij} L_{i,j+1} + \sum_{j=1}^n \sum_{i=1}^{n-1} L_{ij} L_{i+1,j} \right) \quad (14)$$

### Example: Sum Over Pairs

What about the edges? With a 2D lattice, each interior point on the lattice has 4 nearest neighbors. An edge point has only 3 neighbors, while a corner point has only 2 neighbors. Since it is the interactions with neighbors that make the spins want to flip, and edge/corner points have fewer neighbors, the edge/corner points will be more resistant to spin flips. This is true in real systems as well as our simulated one. However, the *fraction* of edge/corner points gets smaller as the number of lattice points gets larger. (For 9 lattice points,  $8/9 = 89\%$  are edges/corners. For 16 lattice points,  $12/16 = 75\%$  are edges/corners. For 100 lattice points,  $36/100 = 36\%$  are edges/corners.) So in real systems ( $\sim 10^{23}$  lattice points), the fraction of edge/corner points is extremely small, and the fact that those spins are more resistant to change is a negligible effect. But in our simulation, the fraction of edge/corner points is significant, and the fact that those spins are more resistant to change creates noticeable effects. To make our simulations more realistic, but avoid using  $\sim 10^{23}$  lattice points, we can eliminate these effects by this using “periodic boundary conditions.” For a given row on the lattice, we assume the right-most point is connected to the left-most point. (This is equivalent to taking our flat lattice and rolling it up to make a cylinder.) But also for a given column on the lattice, we assume the top point is connected to the bottom point. (This is equivalent to taking your cylinder and connecting the end points to make a torus.) Now we have a continuous lattice, which allows us to simulate a real system without using an extremely large number of points.

Now how many nearest neighbor pairs are there in an  $n \times n$  lattice?

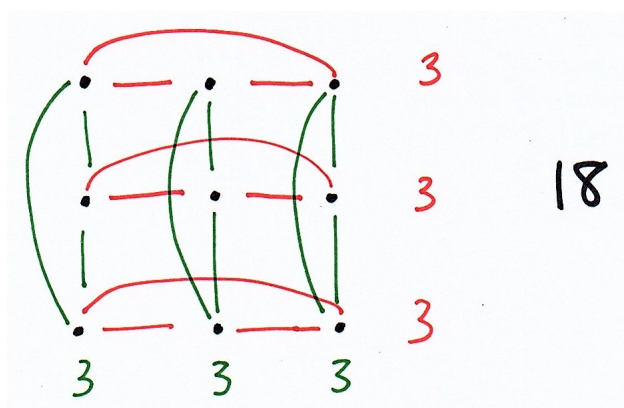


Figure 2:

See Figure 2. For a  $3 \times 3$  lattice, adding a connection between the first and last spin means there are now three neighboring pairs in each row and three neighboring pairs in each column

---

for  $3 \times 3 + 3 \times 3 = 18$  total pairs. If you can imagine extending this to an  $n \times n$  lattice, there would be  $n$  neighboring pairs in each of  $n$  rows plus  $n$  neighboring pairs in each of  $n$  columns for  $n^2 + n^2 = 2n^2$  total neighboring pairs. (For the  $3 \times 3$  case,  $2(3^2) = 18$ .)

With periodic boundary conditions, we can write the sum in Equation 13 this way:

$$E/J = - \left( \sum_{i=1}^n \sum_{j=1}^n L_{ij} L_{i,j+1} + \sum_{j=1}^n \sum_{i=1}^n L_{ij} L_{i+1,j} \right) \quad (15)$$

if  $j + 1 > n$  set  $j + 1 = 0$   
if  $i + 1 > n$  set  $i + 1 = 0$

where we've added a term to each sum (by summing to  $n$  instead of  $n - 1$ ) and specified that if we reach the end of a row/column, we loop back around to the first entry.

### Example: Sum Over Pairs with Periodic Boundary Conditions

## 5 Simulation: Calculating the change in energy

In our simulation, we're going to be using the probability given in Equation 12, which requires us to know the difference in energy between two microstates  $E_1 - E_2$ . We're only going to consider changing one spin at a time, so the difference between microstate 1 and microstate 2 will always be just one spin. Instead of summing over all the pairs twice to calculate both  $E_1$  and  $E_2$  and finding the difference, it will be faster to just calculate the difference directly, which we can do by considering only the terms in the sum that will change. If  $L_{ij}$  is the location of the one spin that's different between the two microstates, then  $L_{ij}$  is the value of the spin before the flip, and  $-L_{ij}$  is the value after. Let  $S_{\text{other pairs}}$  be the sum for the all the pairs that do not include this particular spin. The energy difference between the states before and after the one spin flip is then

$$\begin{aligned} \Delta(E/J) &= (E/J)_{\text{after}} - (E/J)_{\text{before}} \\ &= -[S_{\text{other pairs}} - L_{ij}L_{i+1,j} - L_{ij}L_{i-1,j} - L_{ij}L_{i,j+1} - L_{ij}L_{i,j-1}] \\ &\quad - (-[S_{\text{other pairs}} + L_{ij}L_{i+1,j} + L_{ij}L_{i-1,j} + L_{ij}L_{i,j+1} + L_{ij}L_{i,j-1}]) \\ &= -S_{\text{other pairs}} + L_{ij}L_{i+1,j} + L_{ij}L_{i-1,j} + L_{ij}L_{i,j+1} + L_{ij}L_{i,j-1} \\ &\quad + S_{\text{other pairs}} + L_{ij}L_{i+1,j} + L_{ij}L_{i-1,j} + L_{ij}L_{i,j+1} + L_{ij}L_{i,j-1} \\ &= 2L_{ij} (L_{i+1,j} + L_{i-1,j} + L_{i,j+1} + L_{i,j-1}) \end{aligned} \quad (16)$$

Keep in mind we are using periodic boundary conditions.

### Example: Energy Change



---

## 6 Implementing the simulation

For now, we will assume  $H = 0$  (no magnetic field) and simulate the exchange of energy between the spin system and the heat bath. First of all, regardless of the heat bath, if a spin can lower the total energy of the microstate by flipping (to match its neighbors), then it will do that spontaneously. A spin can also flip in such a way that increases the total energy of the microstate, due to the interaction with the heat bath.

We'll use what is called the *Metropolis algorithm*. In our Markov chain, we will start with some initial state (an initial set of spins). To get to the next step in the chain, we choose one spin on the lattice, and calculate what the energy of the system would be if that spin flipped. We decide whether or not to flip the spin based on the energy difference between the current state  $i$  and the state after flipping. We flip the spin with probability

$$P_{\text{flip}} = \begin{cases} 1 & E_{\text{flipped}} \leq E_{\text{current}} \\ e^{-(E_{\text{flipped}} - E_{\text{current}})/(k_B T)} & E_{\text{flipped}} > E_{\text{current}} \end{cases} \quad (17)$$

If the energy of the state with the spin flip is less than the current energy, we definitely flip the spin. (This spin would flip spontaneously to reach a more energetically favorable state). If the energy of the state with the spin flip is greater than the current energy, we flip based on the relative probability of the two states, Equation 12. This is due to the interaction with the heat bath.

Based on the probability above, we can now define the transition probabilities  $T_{ij}$  that we mentioned earlier and verify that they meet the condition we specified about the ratio earlier.

If  $E_j > E_i$ ,

$$T_{ij} = \frac{1}{M} e^{-(E_j - E_i)/(k_B T)} \quad (18)$$

$$T_{ji} = \frac{1}{M} \quad (19)$$

where  $M$  is the number of spins on the lattice. The factor  $1/M$  comes from the probability of picking that particular spin to flip. The probability of going from state  $i$  to state  $j$  (increasing the energy) is given by the ratio of the Boltzmann probabilities. The probability of going from state  $j$  to state  $i$  is one, because state  $i$  is at a lower energy than state  $j$  and the transition is guaranteed. The ratio is given by

$$\frac{T_{ij}}{T_{ji}} = \frac{(1/M) e^{-(E_j - E_i)/(k_B T)}}{1/M} = e^{-(E_j - E_i)/(k_B T)} \quad (20)$$

as required. (You can show the equivalent thing for the case of  $E_j < E_i$ .)

So now we can be confident that after enough steps in the chain, each new state is selected according to the Boltzmann probability, and that macroscopic quantities can be calculated as straight averages over the steps of the Markov chain.

Here is the algorithm we will use to simulate this spin system:

- 
- Set the value of  $T$ .
  - Begin with a particular microstate (for example, all  $s_i = 1$ ). Calculate the starting energy.
  - Choose a particular spin on the lattice. What would be the change in energy of the system if this spin was flipped? Calculate  $\Delta E = E_{\text{flipped}} - E_{\text{current}}$ .
  - If the energy of the system would be lower if this spin flipped ( $\Delta E < 0$ ), then flip the spin.
  - If the energy of the system would be higher if this spin flipped ( $\Delta E > 0$ ), then we need to make a choice based on probabilities. According to Equation 12, the probability that the system will be in the flipped state compared to the current state is given by  $\exp(-\Delta E/(k_B T))$ . Generate a random number between 0 and 1. If the random number is less than this probability, the the spin is flipped.
  - Record the new energy, which will be  $E_{\text{current}}$  if the spin did not flip and  $E_{\text{current}} + \Delta E$  if it did.
  - Repeat for all spins on the lattice.

This is one loop over all the spins in the lattice. Repeat this many times.

To simplify the interpretation, we're going to let  $J = 1$ , and give  $T$  in units of  $1/k_B$ . We will use a  $10 \times 10$  lattice, and start with all the spins up. (With all spins up, the system is already in its lowest energy state, so the changes are dominated by the interaction with the heat bath.) In this first example, we'll only loop through the lattice one time.

### Example: Simulation with no magnetic field

So we see that the number of flipped spins increases as  $T$  increases. As  $T$  gets larger,  $\Delta E/(k_B T)$  gets smaller, so  $\exp(\Delta E/(k_B T))$  gets smaller, so  $\exp(-\Delta E/(k_B T))$  gets larger. When that factor gets larger, the probability that a random number between 0 and 1 is less than that number also gets larger. Hence, more flipped spins.

## 7 Second-Order Phase Transitions

Next we will look at properties of the system averaged over many loops over the lattice.

Suppose we want to measure the magnetization of the system. Magnetization is defined as the total magnetic dipole moment per unit volume. So up to some constant factors (magnitude of individual magnetic moments and volume of the system), the magnetization of a particular microstate is given by

$$M_\alpha = \sum_i s_i \quad (21)$$

---

## Section 6, Exercise 2

Using the same algorithm and initial condition as before (all spins up), we're going to make 1000 sweeps over the lattice (every spin has 1000 opportunities to flip) and see what the magnetization looks like as a function of time for different temperatures. You can think of each sweep through the lattice as a unit of time. (Relating this unit of time to a quantitative measure of real time is complicated, and not really necessary for our purposes here.)

To calculate the magnetization after each sweep of an  $n \times n$  lattice, we will use

$$M_{\text{lattice}} = \frac{1}{n^2} \sum_i s_i \quad (22)$$

We do this so the magnetization can be compared between different simulations. For example, for a  $10 \times 10$  lattice, the sum of the spins will range between 100 (all spins up) and -100 (all spins down). But for a  $12 \times 12$  lattice, the sum of spins will range from 144 (all spins up) to -144 (all spins down). By normalizing by the total number of lattice points ( $n \times n$ ), the magnetization will range from 1 (all spins up) to -1 (all spins down) regardless of how many lattice points we use.

### Example: Magnetization vs Time with No External Magnetic Field

We see at a low temperatures ( $T = 1$ ), the effect of the heat bath is minimal, and the magnetization stays close to its original value with some small fluctuations. As we increase the temperature to  $T = 2$ , the magnetization is on average less than 1 with larger fluctuations. At  $T = 2.25$ , spins are more likely to flip, and the magnetization might approach zero or the material could even acquire a magnetization in the opposite direction from the initial magnetization. These large fluctuations indicate the presence of a *critical point* near  $T = 2.25$ . At a temperature  $T = 4$  greater than the critical point the magnetization of the material stays near zero. (This is one of the results we wanted our model to have, that high temperatures destroy the magnetization.)

Now let's take a look at the average magnetization vs temperature, where the average is computed over 1000 sweeps of the lattice.

### Example: Magnetization vs Temperature with No External Magnetic Field

We see that the magnetization drops to zero right around  $T = 2.25$ , as shown in Figure 3. The known value of the critical point for this system is  $T_C = 2.27/k_B$ , consistent with our results.

Let's talk about how our lattice size affects our calculations. The periodic boundary conditions we are using help alleviate some but not all of the problem.

First of all, with our  $10 \times 10$  lattice, we see that the average magnetization over 1000 sweeps has a high probability of being negative with  $T = 2.2$ , see Figure 4. If we use a  $20 \times 20$

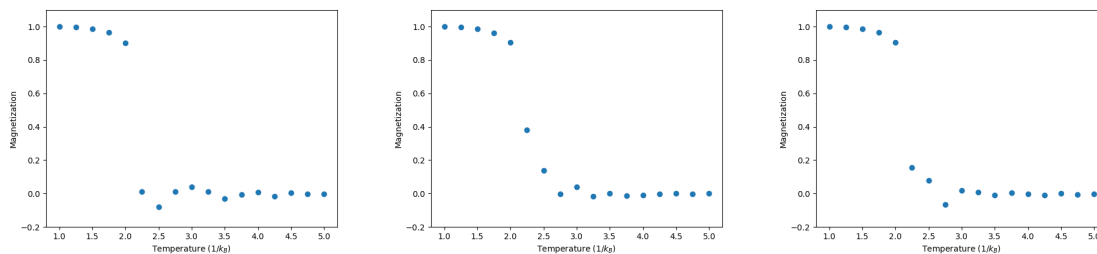


Figure 3:

lattice or  $50 \times 50$  lattice instead (see Figures 5 and 6), this is less likely to happen. (In all cases, these were the first two runs of the program, so not cherry-picking these examples.)

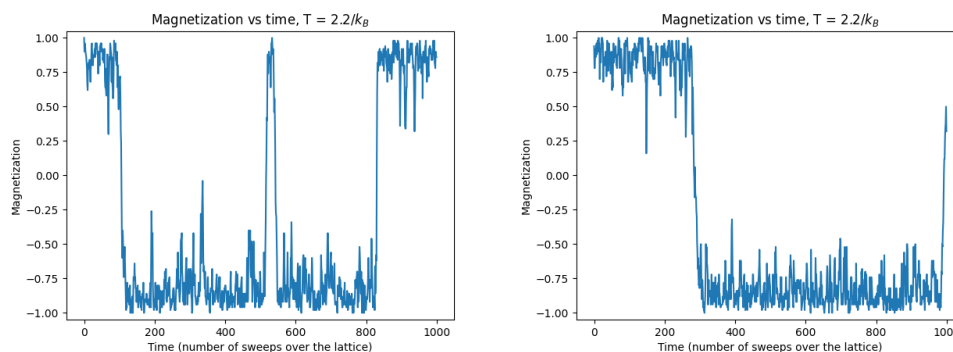


Figure 4: Two examples with the 10x10 lattice with  $T = 2.2$ .

As we've already noted, the fluctuations are large near the critical temperature. With a small lattice, all the spins are "closer" to each other than they are with a large lattice, which makes the probability of a complete flip of the magnetization larger.

A theoretical examination of this problem (which we won't get into) shows that the variation of the magnetization of this system near the critical point is given by

$$M \sim (T_C - T)^\beta \quad (23)$$

where the parameter  $\beta$  is a *critical exponent*. "Near the critical point" in this case means roughly in the range of  $T = 2$  to  $T = T_C$ . To calculate  $\beta$  based on our simulation, let

$$M = h(T_C - T)^\beta \quad (24)$$

where  $h$  is some constant (that we don't need to determine). Then

$$\begin{aligned} \ln M &= \beta \ln(h(T_C - T)) \\ \ln M &= \beta \ln(T_C - T) + \beta \ln h \\ \ln M &= \beta \ln(T_C - T) + \kappa \end{aligned} \quad (25)$$

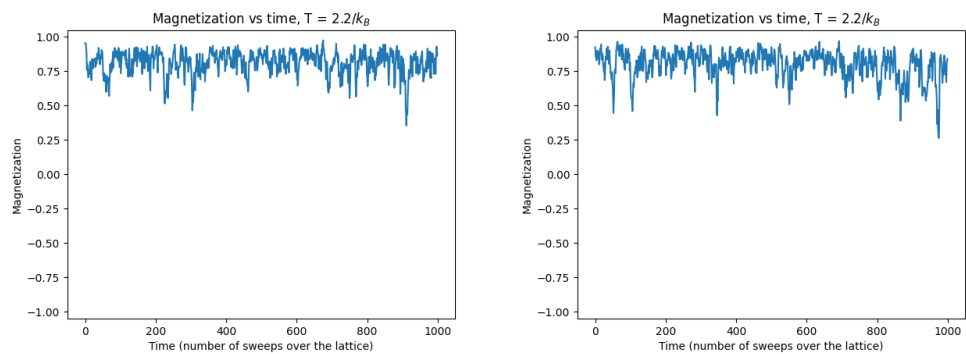


Figure 5: Two examples with the 20x20 lattice with  $T = 2.2$ .

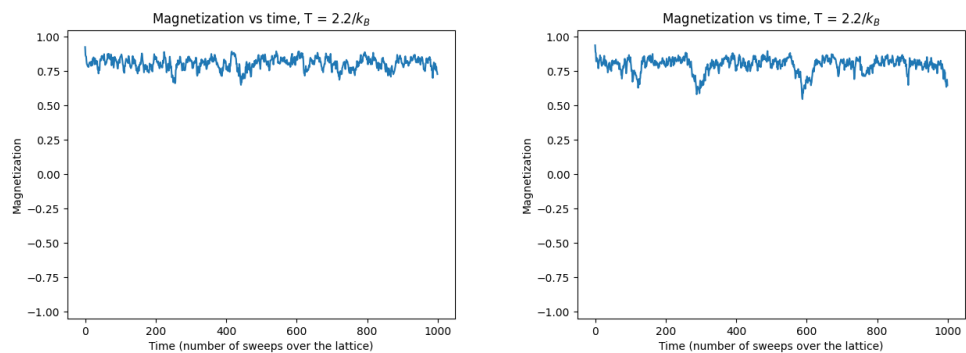


Figure 6: Two examples with the 50x50 lattice with  $T = 2.2$ .

A plot of  $\ln M$  vs  $\ln(T_C - T)$  should have a slope of  $\beta$ . Let's look at  $\ln M$  vs  $\ln(T_C - T)$  and calculate the slope between two temperatures just below the critical point,  $T = 2.2$  and  $T = 2.24$ . Figure 7 shows the results for three different lattice sizes. The exact analytical value of  $\beta$  for a two dimensional lattice is  $\beta = 1/8 = 0.125$ . These values approach that, but it's hard to show definitively due to statistical fluctuations (increasing the number of sweeps though the lattice would help with this), expected fluctuations due to being close to the critical point, and finite lattice size.

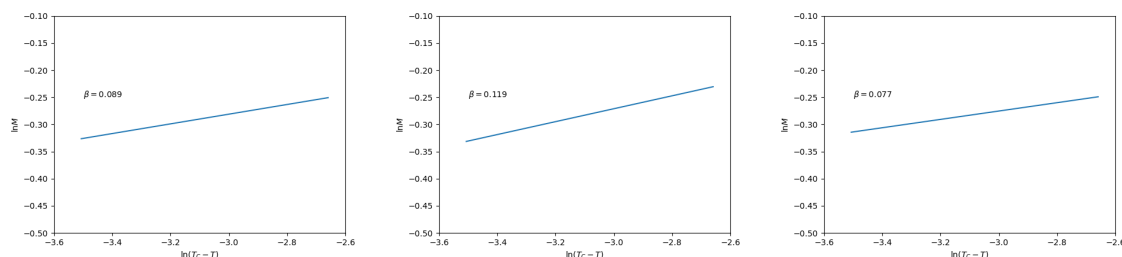


Figure 7:  $\ln M$  vs  $\ln(T_C - T)$  for two temperatures ( $T = 2.2$  and  $T = 2.24$ ) just below the critical point, where  $\beta$  is estimated from the slope between these two temperatures. Left:  $100 \times 100$  lattice, middle:  $110 \times 110$  lattice, right:  $120 \times 120$  lattice.

Now we'll consider some other measurements. First let's consider energy, remembering Equation 15. If all the entries in the lattice are  $+1$ , or all the entries are  $-1$ , the sum becomes

$$\begin{aligned}
 (E/J)_{\text{all aligned}} &= - \left( \sum_{i=1}^n \sum_{j=1}^n L_{ij} L_{i,j+1} + \sum_{j=1}^n \sum_{i=1}^n L_{ij} L_{i+1,j} \right) \\
 &= - \left( \sum_{i=1}^n \sum_{j=1}^n (\pm 1)^2 + \sum_{j=1}^n \sum_{i=1}^n (\pm 1)^2 \right) \\
 &= - \left( \sum_{i=1}^n n + \sum_{j=1}^n n \right) \\
 &= - (n^2 + n^2) \\
 &= -2n^2
 \end{aligned} \tag{26}$$

Since energy depends on  $n$ , the results can't be compared between simulations with different lattice dimensions. Therefore, we will normalize the energy as we did with the magnetization, by dividing by  $n^2$ , which gives the energy per spin:  $E_{\text{per spin}} = E/n^2$ . So the energy per spin for the case of all spins aligned is

$$(E/J)_{\text{all aligned}} = -\frac{1}{n^2}(2n^2) = -2 \tag{27}$$

Important: This is really just for visualization purposes. The  $\Delta E$  we use to make a decision about spin flips in the simulation will still be the total energy, not normalized per spin.

---

Using the same algorithm and initial condition as before (all spins up), we're going to make 1000 sweeps over the lattice and see what the normalized energy looks like as a function of time for different temperatures.

### Example: Energy vs Time with No External Magnetic Field

We see features similar to the magnetization at different temperatures. At low temperature, the energy stays close to the starting value. As  $T$  approaches the critical temperature, there are more fluctuations. Above the critical temperature, the energy hovers around a particular value different from the starting value. (Note this value is not zero.) As we did with magnetization, let's look at energy vs temperature.

### Example: Energy vs Temperature with No External Magnetic Field

Figure 8 shows the results. At low temperatures, the spins mostly remain aligned, and the energy stays close to the minimum.

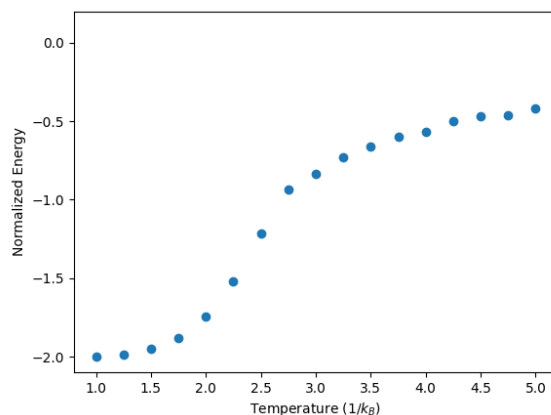


Figure 8:

Let's consider the energy at high temperatures. Above the critical point, the energy levels off somewhat at a value of around -0.5. Remember that at temperatures above the critical point, we found that the average magnetization is zero, indicating that about half the spins are up and half are down. You might assume that such a state should have exactly zero energy as well, but that's not necessarily true. However, the energy does tend to be *close to zero* if these spins are randomly arranged; this is because a typical spin on the lattice would tend to have 2 spin up neighbors and 2 spin down neighbors, leading to a near-zero energy.

### Example: Zero Magnetization Microstates

So we see that the average energy for a random arrangement of half up and half down spins

---

is close to zero. The fact that we see an average near -0.5 in our simulation, even at twice the critical temperature, indicates that the spins are not randomly arranged, and that there is *correlation* among neighboring spins.

Consider a particular spin on the lattice and assume it points up. The system will have a lower energy if the four neighbors of this spin also point up. Even if the temperature is above the critical point, and the average magnetization is zero, these neighboring spins will still have a higher probability of being aligned. The same argument applies to any of these spins and their neighboring spins. This tendency for being aligned propagates through the lattice via common neighbors.

This correlation can be measured using the *correlation function*:

$$f(i) = \langle s_0 s_i \rangle \quad (28)$$

where  $s_i$  is a spin that is located  $i$  lattice sites away from spin  $s_0$ . To compute this, we choose some spin to be the central spin  $s_0$ , then compute the product  $s_0 s_i$  for all spins a distance  $i$  from  $s_0$ . The average is computed over many choices for  $s_0$  and many microstates. For our 2D lattice, we will calculate the correlations for spins in the same row or column.

### Section 6, Exercise 3

Now let's use our simulation to plot the correlation function vs distance. We will conduct the simulation as usual, with 1000 sweeps over the lattice. After each sweep through the lattice to allow spins to flip, we will then loop over the lattice again, setting each spin as  $s_0$  and calculating the quantity  $s_i s_0$  and distance  $i$  between spin  $s_0$  and spin  $s_i$ , where  $s_i$  is every spin in the same row and column as  $s_0$ . After all sweeps over the lattice, we find the average correlation  $s_i s_0$  for each distance  $i$ .

#### Example: Correlations

Figure 9 shows the results for a  $20 \times 20$  lattice for different temperatures. At  $T = 1.5$ , the correlation function is large and almost independent of distance. This is because the system has a large nonzero value of magnetization, so even spins that are far apart will tend to point in the same direction. There is an ever-so-slight increase of  $f$  at the smallest distance. This feature is slightly more pronounced at  $T = 2$ . The notable feature here is that the *extra* alignment due to being in close proximity is weak and limited to a very short range. (Most of the alignment is due to the overall magnetization.)

At  $T = 2.25$ , near the critical temperature, there is noticeably more alignment at short distances compared to the furthest distances, and it drops off more slowly with distance (i.e. it's not just the first point). As the temperature is increased above the critical temperature, the correlations are smaller in magnitude and extend only over the first few lattice spacings. But there are non-zero correlations, which is why the average energy does not go to zero when the average magnetization does at high temperatures.



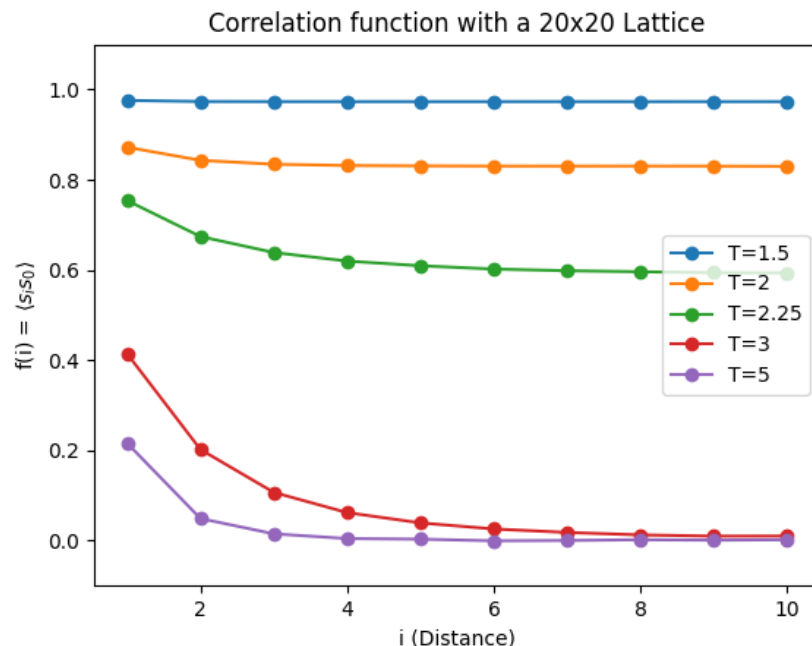


Figure 9:

At the critical point, the range over which the correlations matter becomes essentially infinite, so that the fluctuations of every spin in the lattice are sensitive to those of every other spin. This is the reason for the behavior of quantities such as the magnetization and specific heat at the critical point.

Returning to our results in Figure 8, let's examine the behavior of the average energy near the critical point. Near the critical point, the  $\langle E \rangle$  vs  $T$  curve has an inflection point with a large slope. In fact, for infinitely large systems, the slope is infinite at  $T_C$ . Recall from thermodynamics that the specific heat is proportional to the change in energy with respect to temperature. For our purposes we will define:

$$C \equiv \frac{d\langle E \rangle}{dT} \quad (29)$$

(Quick note: this is usually what we call heat capacity, with specific heat being this quantity per unit mass. The distinction is not important here.) Looking at our plot of  $\langle E \rangle$  vs  $T$ , we see that  $C$  will diverge at  $T = T_C$ . As with the magnetization, this can be described by a power law:

$$C \sim \frac{1}{|T - T_C|^\alpha} \quad (30)$$

where  $\alpha$  is the critical exponent.

---

Given that we know  $\langle E \rangle$  vs  $T$  from our simulations, we could calculate this derivative numerically. However, the statistical fluctuations will lead to large errors. There's a general theorem in statistical mechanics called the "fluctuation dissipation theorem." The proof is beyond our scope, but it can be used to show that the variance in the energy is related to the derivative of the energy with respect to temperature (specific heat) in this way:

$$\begin{aligned}\sigma_E^2 &= k_B T^2 \frac{d\langle E \rangle}{dT} \\ \sigma_E^2 &= k_B T^2 C \\ C &= \frac{\sigma_E^2}{k_B T^2}\end{aligned}\tag{31}$$

Therefore, we can study the behavior of the specific heat  $C$  by considering the energy fluctuations.

We computed the average energy as

$$\langle E \rangle = \frac{1}{N_{\text{sweeps}}} \sum_{\alpha} E_{\alpha}\tag{32}$$

The variance of the energy is

$$\sigma_E^2 = \langle E^2 \rangle - \langle E \rangle^2\tag{33}$$

where

$$\langle E^2 \rangle = \frac{1}{N_{\text{sweeps}}} \sum_{\alpha} E_{\alpha}^2\tag{34}$$

Another note: we have been measuring temperature in units of  $1/k_B$  to avoid actually plugging in the constant. Given  $T = x/k_B$ , the factor in the denominator is

$$k_B T^2 = k_B (x/k_B)^2 = k_B x^2 / k_B^2 = x^2 / k_B\tag{35}$$

So we still have a factor of  $k_B$ . To continue ignoring factors of  $k_B$  in our simulation, we will technically look at the quantity

$$C k_B = \frac{\sigma_E^2}{T^2}\tag{36}$$

where  $T$  is measured in units of  $1/k_B$ .

Let's calculate the specific heat vs temperature. Again, we will normalize the result by  $n^2$ , so that we are plotting the specific heat *per spin*.

### Example: Specific Heat vs Temperature

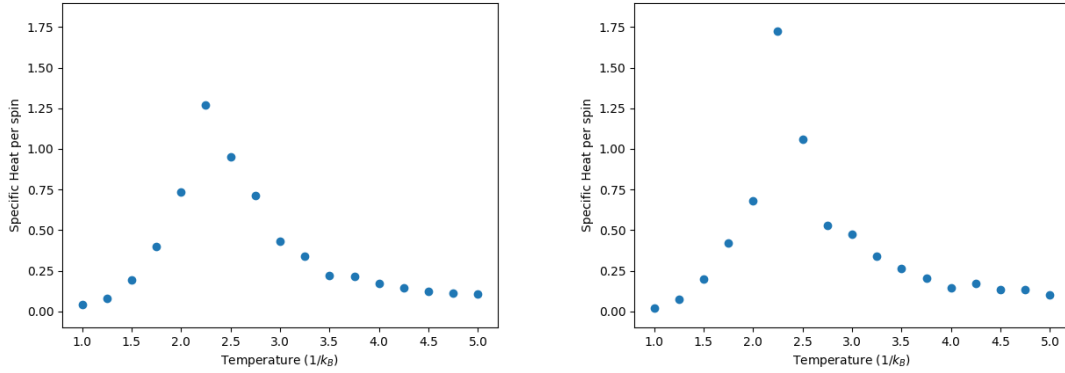


Figure 10: Left:  $10 \times 10$  lattice, right:  $20 \times 20$  lattice.

In Figure 10, we see the specific heat is sharply peaked at the critical temperature as expected. Because the lattice is finite, the specific heat does not go to infinity at this point. With a larger lattice size, the peak value gets greater, and would be infinite with an infinitely large lattice. Figure 10 shows the results for both a  $10 \times 10$  lattice and a  $20 \times 20$  lattice; the peak value is indeed larger for the larger lattice.

The fluctuation dissipation theorem applies to other quantities, including the magnetic susceptibility

$$\chi = \frac{dM}{dH} \quad (37)$$

In this case, applying the fluctuation dissipation theorem gives

$$\chi = \frac{\sigma_M^2}{k_B T} \quad (38)$$

where  $\sigma_M^2$  is the variance of the magnetization, which can also be studied using our simulation.

The transition at the critical point from a ferromagnetic phase with non-zero magnetization to a phase (called the paramagnetic phase) with magnetization of zero is an example of a *second-order phase transition*, also called *continuous phase transitions*. The divergent behavior of the specific heat and the infinite range of correlations at the critical point are characteristic of second-order phase transitions.

---

## 8 First-Order Phase Transitions

Now we will consider the effect of an external magnetic field on our simulation. Recall the energy with a non-zero field:

$$E = -J \sum_{(ij)} s_i s_j - \mu H \sum_i s_i \quad (39)$$

where the first sum is over nearest neighbor pairs, and the second sum is over all spins in the lattice. Note that  $\sum_i s_i$  is what we have defined as the magnetization,  $M$ . In the equation above for the total energy, we can therefore write the energy as

$$E = -J \sum_{(ij)} s_i s_j - \mu H M \quad (40)$$

In our simulation, we will assume the magnitude of the dipole moment  $\mu$  is 1.

For our simulation, we need to calculate the change in energy if one spin flips. As we saw in one of the exercises, the magnetization (sum of spins) changes by  $\pm 2$  if one spin flips. If  $L_{ij}$  is the location of the one spin that's different between the two microstates, then  $L_{ij}$  is the value of the spin before the flip and  $-L_{ij}$  is the value after. Let  $S_{\text{other spins}}$  be the sum for all the other spins on the lattice (that don't change). The difference in magnetization is then:

$$\begin{aligned} \Delta M &= M_{\text{after}} - M_{\text{before}} \\ &= (S_{\text{other spins}} + (-L_{ij})) - (S_{\text{other spins}} + L_{ij}) \\ &= -2L_{ij} \end{aligned} \quad (41)$$

Using Equation 16 for the energy difference coming from the first term, the difference in energy between the two microstates is given by:

$$\begin{aligned} \Delta E &= 2JL_{ij} (L_{i+1,j} + L_{i-1,j} + L_{i,j+1} + L_{i,j-1}) - \mu H \Delta M \\ &= 2JL_{ij} (L_{i+1,j} + L_{i-1,j} + L_{i,j+1} + L_{i,j-1}) - \mu H (-2L_{ij}) \\ &= 2JL_{ij} (L_{i+1,j} + L_{i-1,j} + L_{i,j+1} + L_{i,j-1}) + 2\mu H L_{ij} \end{aligned} \quad (42)$$

Now we add the effect of a magnetic field into our simulation. First, let's assume the magnetic field is pointing up, with  $H = 2$ , and start with all the spins pointing down, with a (normalized) magnetization of -1. We'll assume  $J = 1, \mu = 1$ .

### Example: Magnetization vs Time with an External Magnetic Field

The magnetization basically flips immediately to align with the external field. We see more fluctuations as the temperature is increased, but still an overall large positive magnetization.

Now let's consider the behavior as a function of magnetic field at fixed temperature.

### Example: Magnetization vs Field

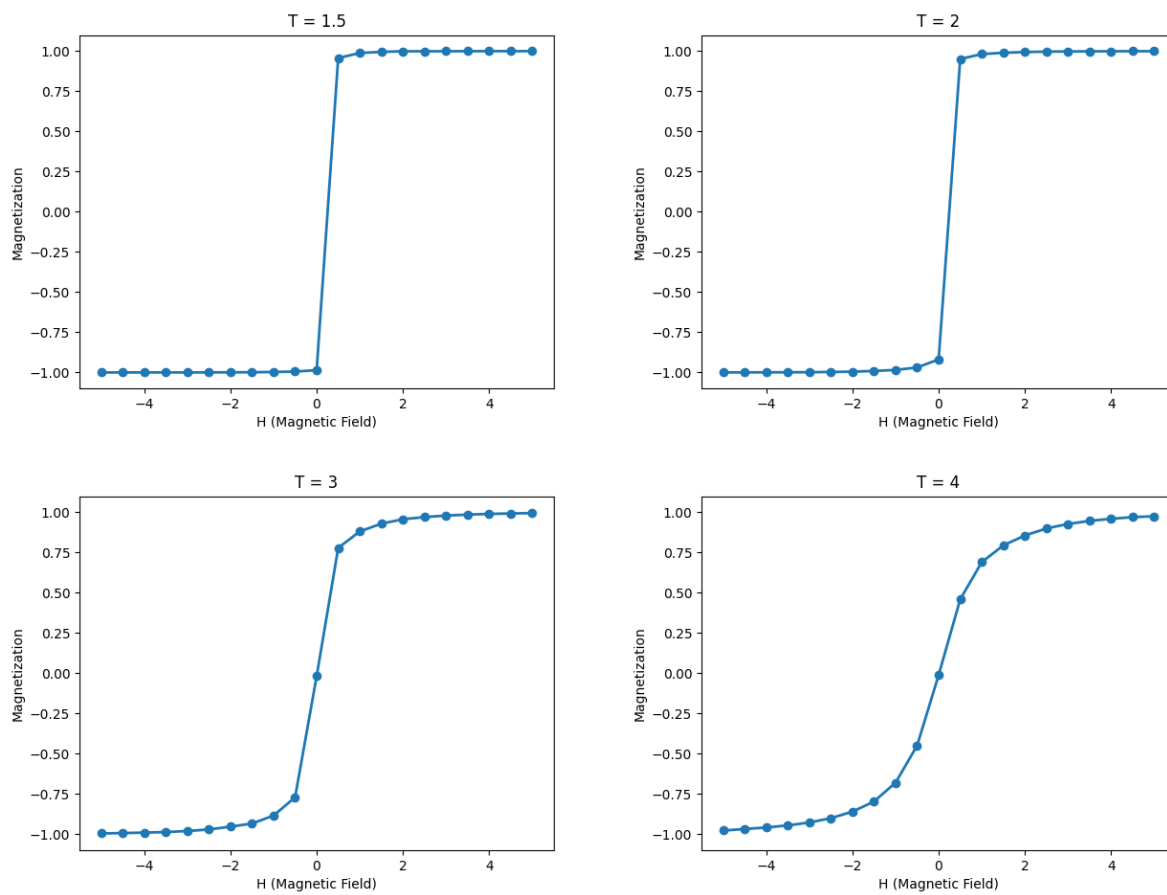


Figure 11:

---

The results are shown in Figure 11. In general, when  $H < 0$ , the magnetization is large and negative, because the field points in the same direction as the initial magnetization. When  $H > 0$ , spins flip to align with the field, creating a large positive magnetization. At low temperatures, the magnetization stays very close to  $-1$  when the field is negative. This is because the probability of a spin flip due to the interaction with the heat bath is low at low temperatures. At low temperatures, the magnetization changes signs abruptly when  $H$  is increased through zero. This discontinuous change of  $M$  indicates the presence of a *first-order phase transition*; it happens at all temperatures below the critical temperature.

At temperatures above the critical point, the transition from negative magnetization to positive magnetization happens gradually. While flipping a given spin to align with the field will lower the energy, the probability of a spin flip to be opposite the field is higher at higher temperatures.

## Section 6, Exercise 4

Let's think about phase transitions we are more familiar with. Figure 12 shows the pressure-temperature phase diagram of a typical pure substance which can be in a solid, liquid, or gas phase. The curves represent the first-order phase transitions that separate the various phases. The vapor-pressure curve (transition between liquid and gas) terminates at a critical point.

Figure 12 also shows the phase diagram of a ferromagnet in the  $H - T$  plane. The line along the temperature axis represents the first-order phase transition between the  $\pm 1$  states of magnetization. At low temperatures, the system has two distinct states of  $\pm 1$  magnetization (represented by the arrows). The material passes from one state to the other abruptly by crossing the temperature axis (going from negative  $H$  to positive  $H$ ). The first-order transition ends at the critical temperature. Beyond the critical point, there is no distinct difference between the states; values of magnetization other than  $\pm 1$  are possible. Similarly, in the pressure-temperature phase diagram, beyond the critical point, there is no distinction between the liquid and the gas state.

In the second-order phase transition, fluctuations become very large near the critical point, leading up to singularities which occur at the critical point, such as the divergence in the specific heat. There are indications in the system that indicate the critical point is near. In first-order phase transition, the discontinuity occurs abruptly; there is no "warning" as there is with a second-order phase transition. For the system to transition from one state  $M = -1$  to the other state  $M = 1$  requires that essentially all the spins flip spontaneously, and the probability for this to happen is extremely small. Because of this, there are no clues that the transition point is close.

Let's consider some lower temperatures than we have looked at so far. Our previous simulations showed a jump in the magnetization that looked like it occur exactly at  $H = 0$ . However, that's not always true; this is shown in Figure 13. At these lower temperatures, the magnetization remains in its initial state even after  $H$  changes sign. The state of the system depends on the past history of the system, which is known as *hysteresis*. This is a

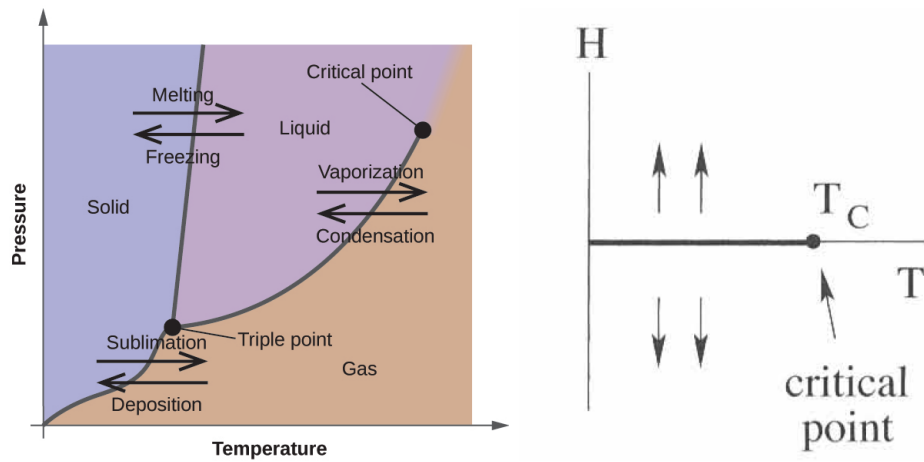


Figure 12: Left figure from <https://cnx.org/contents/oFo044pW05/Phase-Diagrams>, right figure from Giordano and Nakanishi, 2006.

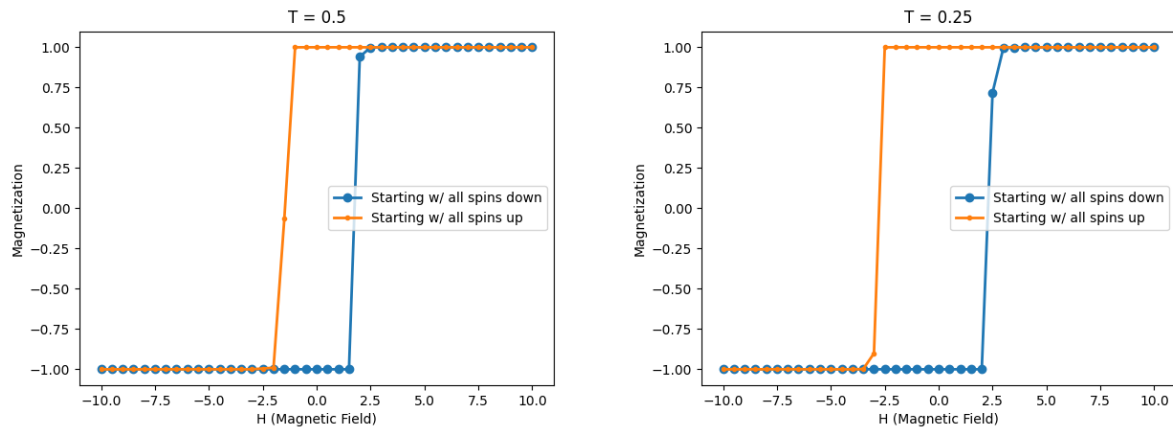


Figure 13: Hysteresis at two different temperatures. The blue curves show a state starting with all spins down ( $M = -1$ ) progressing through field values from left to right. The orange curves show a state starting with all spins up ( $M = 1$ ) progressing through field values from right to left.

consequence of the extremely low probability that a single spin makes the flip, which leaves the system stuck in a *metastable* (nonequilibrium) state. But once one spin flips, the rest immediately follow. In principal, this effect should be reduced if we use a longer time scale in our simulation (more sweeps over the lattice), as it gives the system a longer time to reach equilibrium. But this behavior can persist in macroscopically long times.

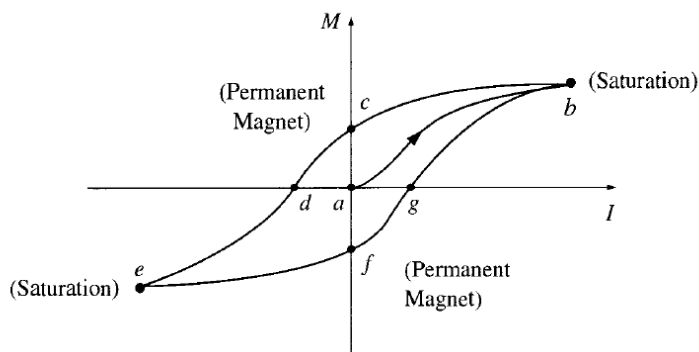


Figure 14:

Suppose we try to magnetize a piece of iron by wrapping a coil of wire around the object, illustrated in Figure 14. We run a current  $I$  through the coil, increasing the current, which increases the magnetization  $M$  until reaching saturation (going from point a to point b on the graph). Then we reduce the current. Because of the tendency of the individual dipoles to stay aligned with their neighbors, the magnetization does not head back towards zero - there's only a slight decrease in  $M$  when  $I$  is zero (point c). Now if we continue changing the current, making it negative, we have flipped the direction of the applied field.  $M$  does go to zero (point d), because at that point, half the domains are aligned with the new field and half are still aligned with the old field. We continue decreasing the current (making it more negative) until we hit saturation again, where all the domains are aligned with the new field, opposite from what it was before (point e). Now if we start increasing the current again, the process will repeat in reverse, until we reach the other saturation point again. This is called a hysteresis loop. The magnetization depends not only on the applied field (proportional to  $I$ ), but also on its previous magnetic history. The current was zero at three different points (a, c, and f), but the magnetization was different each time.

Hysteresis loops are usually plots of  $B$  vs  $H$ , as shown in Figure 15. In our example, the current  $I$ , determined the applied external field  $H$ , and the magnetic field  $B$  is determined by the magnetization of the material  $M$ .  $\mu_0 H$  is the field that the coil would have produced in the absence of any iron,  $B$  is the field we got including the effect of the ferromagnetic material.  $B$  is much greater than  $\mu_0 H$ ! (Notice when  $\mu_0 H = 5 \times 10^{-4}$  T,  $B = 1$  T. It doesn't take a very large external field to move the domain boundaries, and then the magnetized material produces a much larger field than you would have gotten with the coil alone.



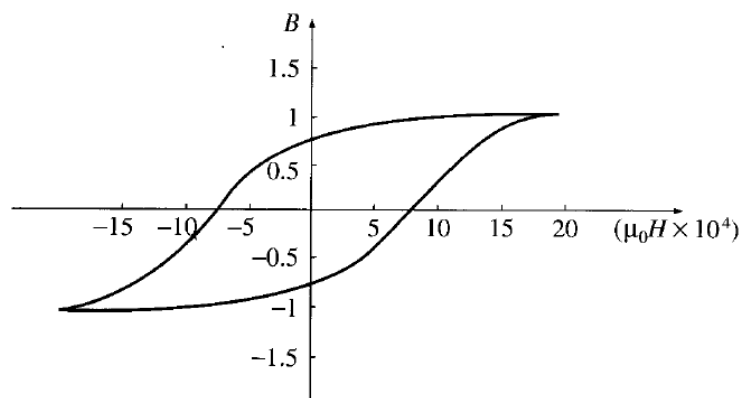


Figure 15: