LECTURE 18
The Ising Model

(References: Kerson Huang, *Statistical Mechanics*, Wiley and Sons (1963) and Colin Thompson, *Mathematical Statistical Mechanics*, Princeton Univ. Press (1972)).

One of the simplest and most famous models of an interacting system is the Ising model. The Ising model was first proposed in Ising's Ph.D thesis and appears in his 1925 paper based on his thesis (E. Ising, *Z. Phys.* **31**, 253 (1925).) In his paper Ising gives credit to his advisor Wilhelm Lenz for inventing the model, but everyone calls it the Ising model. The model was originally proposed as a model for ferromagnetism. Ising was very disappointed that the model did not exhibit ferromagnetism in one dimension, and he gave arguments as to why the model would not exhibit ferromagnetism in two and three dimensions. We now know that the model does have a ferromagnetic transition in two and higher dimensions. A major breakthrough came in 1941 when Kramers and Wannier gave a matrix formulation of the problem. In 1944 Lars Onsager gave a complete solution of the problem in zero magnetic field. This was the first nontrivial demonstration of the existence of a phase transition from the partition function alone.

Consider a lattice of N sites with a spin S on each site. Each spin can take one of two possible values: $+1$ for spin up and -1 for spin down. There are a total of 2^N possible configurations of the system. A configuration is specified by the orientations of the spins on all N sites: $\{S_i\}$. S_i is the spin on the ith lattice site. The interaction energy is defined to be

$$E_I\{S_i\} = -\sum_{\langle i,j\rangle} J_{ij} S_i S_j - \sum_{i=1}^{N} B_i S_i \tag{1}$$

where the subscript I represents the Ising model. A factor of 2 has been absorbed into J_{ij} and we set $g\mu_B = 1$ in the last term. $\langle i,j\rangle$ means nearest neighbor pairs of spins. So $\langle i,j\rangle$ is the same as $\langle j,i\rangle$. J_{ij} is the exchange constant; it sets the energy scale. For simplicity, one sets J_{ij} equal to a constant J. If $J > 0$, then the spins want to be aligned parallel to one another, and we say that the interaction is ferromagnetic. If $J < 0$, then the spins want to be antiparallel to one another, and we say that the interaction is antiferromagnetic. If J_{ij} is a random number and can either be positive or negative, then we have what is called a spin glass. For simplicity we will set $J_{ij} = J > 0$ and study the ferromagnetic Ising model. The last term represents the coupling of the spins to an external magnetic field B. The spins are assumed to lie along the z-axis as is the magnetic field $\vec{B} = B\hat{z}$. The spins lower their energy by aligning parallel to the field. I put B_i to indicate the possibility that the field could vary from spin to spin. If the field B_i is random, this is called the random field Ising model. We will assume a constant uniform magnetic field so that $B_i = B > 0$. So the interaction energy becomes

$$E_I\{S_i\} = -J\sum_{\langle i,j\rangle} S_i S_j - B\sum_{i=1}^{N} S_i \tag{2}$$

The partition function is given by

$$Z = \sum_{s_1=-1}^{+1} \sum_{s_2=-1}^{+1} ... \sum_{s_N=-1}^{+1} e^{-\beta E_I\{S_i\}} \tag{3}$$

One Dimensional Ising Model and Transfer Matrices

Let us consider the one-dimensional Ising model where N spins are on a chain. We will impose periodic boundary conditions so the spins are on a ring. Each spin only interacts with its neighbors on either side and with the external magnetic field B. Then we can write

$$E_I\{S_i\} = -J \sum_{i=1}^{N} S_i S_{i+1} - B \sum_{i=1}^{N} S_i \tag{4}$$

The periodic boundary condition means that

$$S_{N+1} = S_1 \tag{5}$$

The partition function is

$$Z = \sum_{s_1=-1}^{+1} \sum_{s_2=-1}^{+1} ... \sum_{s_N=-1}^{+1} \exp\left[\beta \sum_{i=1}^{N} (JS_i S_{i+1} + BS_i)\right] \tag{6}$$

Kramers and Wannier (*Phys. Rev.* **60**, 252 (1941)) showed that the partition function can be expressed in terms of matrices:

$$Z = \sum_{s_1=-1}^{+1} \sum_{s_2=-1}^{+1} ... \sum_{s_N=-1}^{+1} \exp\left[\beta \sum_{i=1}^{N} \left(JS_i S_{i+1} + \frac{1}{2} B\left(S_i + S_{i+1}\right)\right)\right] \tag{7}$$

This is a product of 2×2 matrices. To see this, let the matrix P be defined such that its matrix elements are given by

$$\langle S|P|S'\rangle = \exp\left\{\beta\left[JSS' + \frac{1}{2}B(S+S')\right]\right\} \tag{8}$$

where S and S' may independently take on the values ± 1. Here is a list of all the matrix elements:

$$\begin{aligned} \langle +1|P|+1\rangle &= \exp\left[\beta(J+B)\right] \\ \langle -1|P|-1\rangle &= \exp\left[\beta(J-B)\right] \\ \langle +1|P|-1\rangle &= \langle +1|P|-1\rangle = \exp[-\beta J] \end{aligned} \tag{9}$$

Thus an explicit representation for P is

$$P = \begin{pmatrix} e^{\beta(J+B)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-B)} \end{pmatrix} \tag{10}$$

With these definitions, we can write the partition function in the form

$$\begin{aligned}
Z &= \sum_{s_1=-1}^{+1}\sum_{s_2=-1}^{+1}...\sum_{s_N=-1}^{+1}\langle S_1|P|S_2\rangle\langle S_2|P|S_3\rangle...\langle S_N|P|S_1\rangle \\
&= \sum_{s_1=-1}^{+1}\langle S_1|P^N|S_1\rangle \\
&= TrP^N \\
&= \lambda_+^N + \lambda_-^N
\end{aligned} \tag{11}$$

where λ_+ and λ_- are the two eigenvalues of P with $\lambda_+ \geq \lambda_-$. The fact that Z is the trace of the Nth power of a matrix is a consequence of the periodic boundary condition Eq. (5). The eigenvalue equation is

$$det\begin{vmatrix} e^{\beta(J+B)}-\lambda & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-B)}-\lambda \end{vmatrix} = \lambda^2 - 2\lambda e^{\beta J}\cosh(\beta B) + 2\sinh(2\beta J) = 0 \tag{12}$$

Solving this quadratic equation for λ gives

$$\lambda_\pm = e^{\beta J}\left[\cosh(\beta B) \pm \sqrt{\cosh^2(\beta B) - 2e^{-2\beta J}\sinh(2\beta J)}\right] \tag{13}$$

When $B = 0$,

$$\lambda_+ = 2\cosh(\beta J) \tag{14}$$

$$\lambda_- = 2\sinh(\beta J) \tag{15}$$

Now back to the general case with $B \neq 0$. Notice that $\lambda_-/\lambda_+ \leq 1$ where equality is in the case of $J = B = 0$. In the thermodynamic limit ($N \to \infty$), only the larger eigenvalue λ_+ is relevant. To see this, we use $(\lambda_-/\lambda_+) < 1$ and write the Helmholtz free energy per spin:

$$\begin{aligned}
-\frac{F}{Nk_BT} &= \lim_{N\to\infty}\frac{1}{N}\ln Z \\
&= \lim_{N\to\infty}\frac{1}{N}\ln\left\{\lambda_+^N\left[1+\left(\frac{\lambda_-}{\lambda_+}\right)^N\right]\right\} \\
&= \ln\lambda_+ + \lim_{N\to\infty}\frac{1}{N}\ln\left[1+\left(\frac{\lambda_-}{\lambda_+}\right)^N\right] \\
&= \ln\lambda_+
\end{aligned} \tag{16}$$

So the Helmholtz free energy per spin is

$$\begin{aligned}
\frac{F}{N} &= -\frac{k_BT}{N}\ln Z = -k_BT\ln\lambda_+ \\
&= -J - k_BT\ln\left[\cosh(\beta B) + \sqrt{\cosh^2(\beta B) - 2e^{-2\beta J}\sinh(2\beta J)}\right]
\end{aligned} \tag{17}$$

The magnetization per spin is

$$\begin{aligned} m &= \frac{M}{N} \\ &= \frac{1}{\beta N}\frac{\partial \ln Z}{\partial B} \\ &= -\frac{1}{N}\frac{\partial F}{\partial B} \\ &= \frac{\sinh(\beta B)}{\sqrt{\cosh^2(\beta B) - 2e^{-2\beta J}\sinh(2\beta J)]}} \end{aligned} \tag{18}$$

At zero field ($B = 0$), the magnetization is zero for all temperatures. This means that there is no spontaneous magnetization and the one-dimensional Ising model never exhibits ferromagnetism. The reason is that at any temperature the average configuration is determined by two opposite and competing tendencies: The tendency towards a complete alignment of spins to minimize the energy, and the tendency towards randomization to maximize the entropy. The over-all tendency is to minimize the free energy $F = E - TS$. For the one-dimensional model the tendency for alignment always loses out, because there are not enough nearest neigbors. However, in higher dimensions, there are enough nearest neighbors and a ferromagnetic transition can occur.

The method of transfer matrices can be generalized to two and higher dimensions, though the matrices become much larger. For example, in two dimensions on an $m \times m$ square lattice, the matrices are $2^m \times 2^m$. In 1944, Onsager solved the two dimensional Ising model exactly for the zero field case, and found a finite temperature ferromagnetic phase transition. This is famous and is known as the Onsager solution of the 2D Ising model. No one has found an exact solution for the three dimensional Ising model.

Applications of the Ising Model

The Ising model can be mapped into a number of other models. Two of the better known applications are the lattice gas and the binary alloy.

Lattice Gas

The term lattice gas was first coined by Yang and Lee in 1952, though the interpretation of the model as a gas was known earlier. A lattice gas is defined as follows. Consider a lattice of V sites (V = volume) and a collection of N particles, where $N < V$. The particles are placed on the vertices of the lattice such that not more than one particle can occupy a given site, and only particles on nearest-neighbor lattice sites interact. The interaciton potential between two lattice sites i and j is given by $V(|\vec{r}_i - \vec{r}_j|)$ with

$$V(r) = \begin{cases} \infty & (r = 0) \\ -\varepsilon_o & (r = a) \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where a is the lattice spacing. The occupation n_i of a lattice site i is given by

$$n_i = \begin{cases} 1 & \text{if site i is occupied} \\ 0 & \text{if site i is unoccupied} \end{cases} \tag{20}$$

The interaction energy is

$$E_G\{n\} = -\varepsilon_o \sum_{\langle i,j \rangle} n_i n_j \tag{21}$$

We can map this into the Ising model by letting spin-up denote an occupied site and letting spin-down denote an unoccupied site. Mathematically, we write

$$S_i = 2n_i - 1 \tag{22}$$

So $S_i = 1$ means site i is occupied and $S_i = -1$ means site i is unoccupied. One can then map the lattice gas model into Ising model. For example, by comparing the partition functions, it turns out that $\epsilon_o = 4J$.

Binary Alloy

A binary alloy is a solid consisting of 2 different types of atoms. For example, $\beta-$brass is a body-centered cubic lattice made up of Zn and Cu atoms. At $T = 0$, the lattice is completely ordered and a copper atom is surrounded by zinc atoms and vice-versa. However, at non-zero temperatures the zinc and copper atoms can exchange places. Above a critical temperature of $T = 742$ K, the Zn and Cu atoms are thoroughly mixed so that the probability of finding a Zn atom on given site is 1/2. Similarly the probability of finding a Cu atom on given site is 1/2. To model a binary alloy, one starts with a lattice of N sites, and two different types of atoms, A and B. Each site has only one atom so that $N_A + N_B = N$. The occupation of each site is

$$n_i = \begin{cases} 1 & \text{if site i is occupied by atom A} \\ 0 & \text{if site i is occupied by atom B} \end{cases} \tag{23}$$

There are interaction energies between nearest neighbor sites: ε_{AA}, ε_{BB} and ε_{AB}. One can map the binary alloy model into the lattice gas model and the Ising model.

Generalizations to Other Spin Models

One can generalize the Ising model in a number of ways, though the Ising spins refer to entities with two possible values. Classically other types of spins are xy spins which can rotate in the $x - y$ plane and have a fixed length (usually $|S| = 1$). They have two components: S_x and S_y. Heisenberg spins are fixed length spins that can point anywhere on a unit sphere. Heisenberg spins have three components: S_x, S_y, and S_z.

Quantum mechanically, the values of S_z are discretized. So Ising spins correspond to $\langle S_z \rangle = \pm 1/2$. $\vec{S} = \vec{\sigma}/2$ where $\vec{\sigma}$ are the Pauli spin matrices.

We have already mentioned that the Ising model can be considered in higher dimensions. Another variation of the Ising model is to consider other types of lattices. For example in two dimensions, we could have a triangular lattice. In three dimensions there are a wide variety of lattice structures that could be considered. Or one could throw away the lattice and have randomly placed sites to make a type of spin glass.

Other variations center around the form of the interaction. For example, we can allow the nearest neighbor interactions to be antiferromagnetic. Or we can allow the interactions to extend over a longer range to include next nearest-neighbors or even farther, e.g., infinite range. The interaction is contained in the exchange constant J_{ij}. So one could have something like

$$J_{ij} = \frac{A}{|\vec{r}_i - \vec{r}_j|^n} \tag{24}$$

where $n = 1$ is the Coulomb interaction and $n = 3$ is similar to a dipolar interaction. Another interaction is the RKKY interaction. RKKY stands for Ruderman-Kittel-Kasuya-Yosida. The RKKY interaction has the form

$$\begin{aligned} J(r) &\sim \frac{\sin(2k_F r) - 2k_F r\cos(2k_F r)}{(k_F r)^4} \\ &\sim \frac{\cos(2k_F r)}{(k_F r)^3} \end{aligned} \tag{25}$$

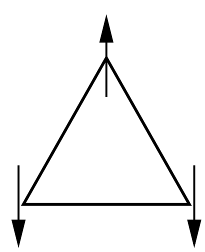
where k_F is the Fermi wavevector. This interaction is found in metals with magnetic atoms. The interaction is mediated by the conduction electrons. Notice that the interaction oscillates and decays as a power law.

Frustration and Spin Glasses

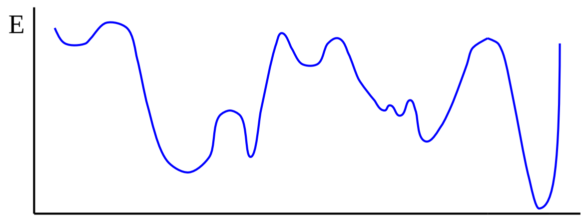
Magnetic impurities randomly placed in a metal, e.g., Mn impurities in copper, will interact with one another via the RKKY interaction. Because of the oscillations, the interactions will be random. This spin system is called a spin glass. For simplicity, the RKKY interaction is replaced by a random J_{ij} in the spin Hamiltonian:

$$\mathcal{H} = -\sum_{i>j} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j \tag{26}$$

Typically, J_{ij} is chosen from a distribution $P(J)$ centered at $J = 0$. For example, $J_{ij} = \pm J$ where J is a positive constant, and there is an equal probability of choosing the plus or minus sign. Another possibility is to have $P(J)$ be a Gaussian distribution centered at $J = 0$. Obviously the ground state of a such a system will be disordered, but a phase transition from a paramagnetic phase at high temperatures to a frozen spin configuration is possible.



One concept that is associated with spin glasses is "frustration." The idea is best illustrated by considering an equilateral triangle with a spin on each vertex. Suppose the interaction between nearest neighbor spins is antiferromagnetic so that a spin wants to point opposite from its neighbors. There is no way all the spins on the triangle can be satisfied. This is an example of frustration. Frustration occurs when there is no spin configuration where all the spins have their lowest possible interaction energy. In a spin glass, there is a great deal of frustration. As a result there is no clear ground state configuration. One can associate an energy landscape with the energies of different spin configurations. Valleys correspond to low energy configurations and mountains to high energy configurations. The landscape exists in the space of spin configurations. So to go from one valley to another, the system must climb out of the first valley by going through some high energy configurations and then descend into the second valley by passing through configurations with decreasing energy.



Spin glasses are often used to model interacting systems with randomness. They were originally proposed to explain metals with magnetic impurities and were thought to be a simple model of a glass. Spin glass models have a wide range of applications, e.g., they have been used to model the brain and were the basis of neural networks and models of memory.

Monte Carlo Simulations

Reference: D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge Univ. Press (2000).

As one can see, analytic solutions to spin systems can be difficult to obtain. So one often resorts to computer simulations, of which Monte Carlo is one of the most popular. Monte Carlo simulations are used widely in physics, e.g., condensed matter physics, astrophysics, high energy physics, etc. Typically in Monte Carlo simulations, one evolves the system in time. The idea is to visit a large number of configurations in order to do statistical sampling. For a spin system we would want to obtain average values of thermodynamic quantities such as magnetization, energy, etc. More generally, Monte Carlo is an approach to computer simulations in which an event A occurs with a certain probability P_A where $0 \leq P_A \leq 1$. In practice, during each time step, a random number x is generated with uniform probability between 0 and 1. If $x \leq P_A$, event A occurs; if $x > P_A$, event A does not occur. Monte Carlo is also able to handle cases where multiple outcomes are possible. For example, suppose there are three possibilities so that either event A can occur with probability P_A or event B can occur with probability P_B or neither occurs. Then if $x \leq P_A$, A occurs; if $P_A < x \leq (P_A + P_B)$, B occurs; and if $(P_A + P_B) < x \leq 1$, neither occurs.

The Metropolis algorithm (Metropolis *et al.*, *J. Chem. Phys.* **21**, 1087 (1953)) is a classic Monte Carlo method. Typically, configurations are generated from a previous state using a transition probability which depends on the energy difference ΔE between the initial and final states. For relaxational models, such as the (stochastic) Ising model, the probability obeys a master equation of the form:

$$\frac{\partial P_n(t)}{\partial t} = \sum_{n \neq m} \left[-P_n(t) W_{n \to m} + P_m(t) W_{m \to n} \right] \tag{27}$$

where $P_n(t)$ is the probability of the system being in state n at time t, and $W_{n \to m}$ is the transition rate from state n to state m. This is an example of a master equation. Master equations are found in a wide variety of contexts, including physics, biology (signaling networks), economics, etc.

In equilibrium $\partial P_n(t)/\partial t = 0$ and the two terms on the right hand side must be equal. The result is known as 'detailed balance':

$$P_n(t) W_{n \to m} = P_m(t) W_{m \to n} \tag{28}$$

Loosely speaking, this says that the flux going one way has to equal the flux going the other way. Classically, the probability is the Boltzmann probability:

$$P_n(t) = \frac{e^{-E_n/k_B T}}{Z} \tag{29}$$

The problem with this is that we do not know what the denominator Z is. We can get around this by generating a Markov chain of states, i.e., generate each new state directly from the preceding state. If we produce the nth state from the mth state, the relative probability is given by the ratio

$$\begin{aligned} \frac{P_n(t)}{P_m(t)} &= \frac{e^{-E_n/k_B T}}{e^{-E_m/k_B T}} = \frac{W_{m \to n}}{W_{n \to m}} \\ &= e^{-(E_n - E_m)/k_B T} \\ &= e^{-\Delta E/k_B T} \end{aligned} \tag{30}$$

where $\Delta E = (E_n - E_m)$.

Any transition rate which satisfies detailed balance is acceptable. Historically the first choice of a transition rate used in statistical physics was the Metropolis form:

$$W_{n \to m} = \begin{cases} \exp\left(-\Delta E/k_B T\right) & \Delta E > 0 \\ 1 & \Delta E < 0 \end{cases} \tag{31}$$

Here is a recipe on how to implement the Metropolis algorithm on a spin system:

1. Choose an initial state.

2. Choose a site i.

3. Calculate the energy change ΔE which results if the spin at site i is overturned.

4. If $\Delta E < 0$, flip the spin. If $\Delta E > 0$, then
    (a) Generate a random number r such that $0 < r < 1$.
    (b) If $r < \exp(-\Delta E/k_B T)$, flip the spin.
5. Go to the next site and go to (3).

The random number r is chosen from a uniform distribution. The states are generated with a Boltzmann probability. The desired average of some quantity A is given by $\langle A \rangle = \sum_n P_n A_n$. In the simulation, this just becomes the arithmetic average over the entire sample of states visited. If a spin flip is rejected, the old state is counted again for the sake of averaging. Every spin in the system is given a chance to flip. One pass through the lattice is called a "Monte Carlo step/site" (MCS). This is the unit of time in the simulation.

For purposes of a spin model, it is easier to calculate ΔE if we write

$$\begin{aligned}\mathcal{H} &= -\sum_{i>j} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j \\ &= -\frac{1}{2}\sum_{i,j} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j \\ &= -\sum_i \mathbf{S}_i \cdot \left(\sum_j \frac{1}{2} J_{ij} S_j\right) \\ &= -\sum_i \mathbf{S}_i \cdot \mathbf{h}_i \end{aligned} \tag{32}$$

where the local magnetic field h due to the other spins is

$$\mathbf{h}_i = \frac{1}{2}\sum_j J_{ij}\mathbf{S}_j \tag{33}$$

To find the change in energy, we would do a trial flip of $\mathbf{S}_i$ and easily calculate the new energy if we know the local field $\mathbf{h}_i$. If we accept the flip, then we have to update the local fields of the neighboring spins.

As we said before, Monte Carlo simulations are used widely in physics as well as other fields such as chemistry, biology, engineering, finance, etc.