The canonical ensemble

Anders Malthe-Sørenssen

23. september 2013

1 Boltzmann statistics

We have so far studied systems with constant N, V, and E – systems in the microcanonical ensemble. Such systems are often used in dynamic simulations, such as in molecular dynamics simulations, and we can use the principle of conservation of energy to argue for how energy is redistributed in a system with constant energy.

However, studies of systems in the microcanonical ensemble requires the calculations of multiplicities – which often is not that simple in more complicated situations. Also, we may be interested in answering questions that are not directly answered in the microcanonical ensemble. For example, if we look at a single harmonic oscillator – a single atom in a lattice of atoms – we may be interested in the possible states this single oscillator may take: What is the probability for the oscillator to be in a particular state $|i\rangle$ when it is in contact with a large reservoir – a large system – kept at a constant temperature?

We can still use the methodology from the microcanonical ensemble to address such a problem: What is the probability for a single oscillator to be in a particular state $|i\rangle$ with energy ϵ_i?

We consider a big system, the reservoir or heat bath R – for exampe consisting of many oscillators – in contact with the system S – here consisting of a single oscillator. We will assume that the reservoir is so large that it is not affected by the system: If we place a hot cup of water in this auditorium, it will cool down to the temperature of the auditorium without increasing the temperature significantly. The air and other materials in the auditorium is an example of a large heat bath.

We assume that the composite system consisting of the system and reservoir is isolated. We can characterize the macrostate of this system by U, V, and N. The composite system is therefore a microcanonical system and all microstates in this system are equally probable. If the system S is in a state with energy ϵ_i the reservoir has energy $U_0 - \epsilon_i$.

What is the probability for the system to be in the state $|i\rangle$? The probability is given by the multiplicity of the whole system, consisting of both system S and reservoir R:

$$P(i) = \frac{\Omega_S \Omega_R}{\sum_i \Omega_S \Omega_R} \ ,$$

where the sum is for normalization of the probability (to ensure the sum of probabilities of all possible states is 1 – the system must be in one of the possible states).

Since the system S is in the state $|i\rangle$ the multiplicity of this state is 1: We have specified the state so it is only one way it can be in this state. (Note that state here is not the same as the energy, because several states can have the same energy and still be separate states. Here we consider states and not only the energy.)

What is the multiplicity of the reservoir? For a given microstate $|i\rangle$ of the system, the reservoir can be in many microstates. We call this $\Omega(U_0 - \epsilon_i)$, where U_0 is the total energy of the system.

The probability is therefore

$$P(i) = \frac{\Omega_S \Omega_R}{\sum_i \Omega_S \Omega_R} = C \cdot (1 \times \Omega_R) \ ,$$

Let us work with the logarithm of the probability:

$$\ln P(i) = \ln C + \ln \Omega_R(U_0 - \epsilon_i) \ .$$

We can expand $\ln \Omega_R$ around U_0 since $\epsilon_i \ll U_0$ (the heat bath is large):

$$\ln \Omega_R(U_0 - \epsilon_i) = \ln \Omega(U_0) + \left(\frac{\partial \ln \Omega_R(U)}{\partial U} \right)_{U_0} (-\epsilon_i) + \mathcal{O}(\epsilon_\rangle^\in) \ ,$$

Where we can now use that $S = \ln \Omega$ for the reservoir, and we then recognize that

$$\left(\frac{\partial \ln \Omega_R(U)}{\partial U} \right)_{U_0} (-\epsilon_i) = - \left(\frac{(1/k)\partial S}{\partial U} \right)_{U_0} \epsilon_i = -\frac{\epsilon_i}{kT} \ ,$$

where we have used that in equilibrium:

$$\left(\frac{\partial S}{\partial U}\right) = \frac{1}{T} \ .$$

We put this back into the probability and find that

$$P(i) = C \exp\left(-\frac{\epsilon_i}{kT}\right) \ .$$

What is the constant C? It is given by the summation rule:

$$\sum_i P(i) = 1 = C \sum_i \exp\left(-\frac{\epsilon_i}{kT}\right) \ .$$

and therefore

$$C = \frac{1}{Z}$$

where

$$Z = Z(N, V, T) = \sum_i \exp\left(-\frac{\epsilon_i}{kT}\right) \ ,$$

is called the partition function.

The system Z comes from the German "Zustandsumme" – sum of all *states*. The probability is therefore

$$P(i) = \frac{1}{Z} e^{-\frac{\epsilon_i}{kT}} \ ,$$

for the system to be in state $|i\rangle$ when it is in equilibrium with a heat bath with temperature T.

This discussion is completely general and can be applied to any system with a given state $|i\rangle$, not only an oscillator system.

2 Alternative derivation of Boltzmann statistics

We may derive the Boltzmann statistics in a different way that is instructive.

Variational method

3 Average values and fluctuations

Now, if we know the probability for a given state, how can we use this to calculate relevant macroscopic quantities?

We know the probability for a system to be in given state – to have a given energy. This means that in this case the energy of the system is not (strictly) a constant. But we can find a typical value for the energy.

For example, let us address a set of Einstein crystals. Each crystal is in a state $|i\rangle$ with energy ϵ_i. We count and find that of N systems, $N(i)$ are in state i. The total energy of this set of N systems is:

$$E_{TOT} = \sum_i \epsilon_i N(i) \ ,$$

which means that the energy per system, what we would call the average energy is:

$$\bar{E} = \frac{E_{TOT}}{N} = \sum_i \epsilon_i \frac{N(i)}{N} = \sum_i \epsilon_i P(i)$$

hence we can use the probabilities to find the average value of the energy – this is the internal energy U of the system. (Notice that U now refers to the system and not the system + reservoir).

Now we have the tools to connect the microscopic and the macroscopic also in systems where we have constant temperature T – where the system of interest is in contact with a large heat bath.

Let us look at some example applications.

4   Two-state system

Let us start with a very simple case – the two state system. We have an atom/molecule that may be in two states with energies ϵ_0 and ϵ_1 in contact with a thermal bath with temperature T. What is the internal energy and the heat capacity of this system?

We start – and you should always to this – by finding the partition function, Z:

We can always choose our base energy level as we like, so let us choose $\epsilon_0 = 0$.

$$Z = \sum_i e^{-\epsilon_i} kT \; ,$$

where the sum is over all the states (not all the energies, generally, but all the states). Here, all the states is the same as all the energies.
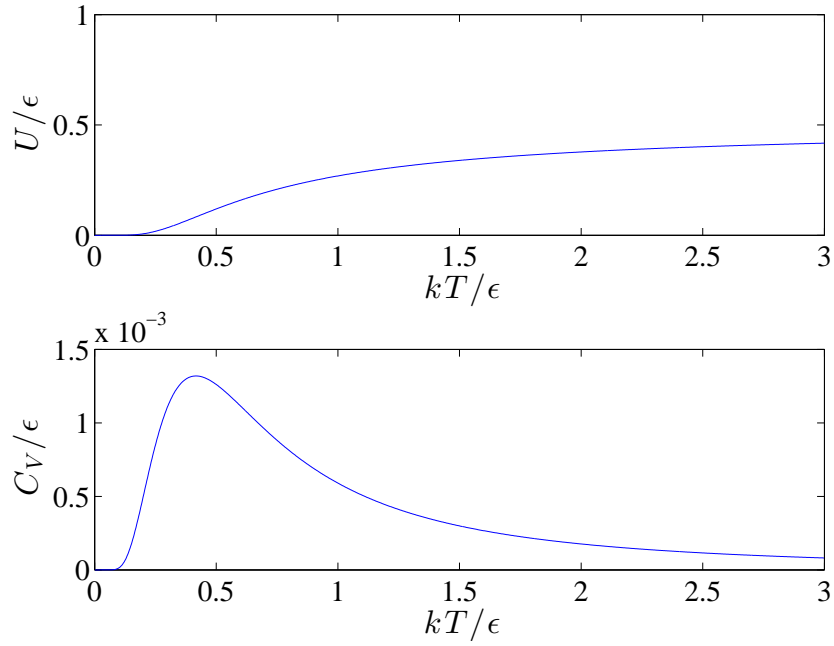
$$Z = \exp(0) + \exp(-\epsilon/kT) = 1 + \exp(-\epsilon/kT) \; ,$$

The average energy is

$$U = \bar{E} = \frac{1}{Z} \sum_i \exp(-\epsilon_i/kT)\epsilon_i = \frac{1}{Z} \left(0 \cdot \exp(0) + \epsilon \cdot \exp(-\epsilon/kT)\right) = \epsilon \frac{\exp(-\epsilon/kT)}{1 + \exp(-\epsilon/kT)}$$

We can sketch this function using the following program

```
% Two state system
clear all; clf
x = linspace(0,3,1000);
U = exp(-1./x)./(1+exp(-1./x));
subplot(2,1,1)
plot(x,U)
xlabel('kT/\epsilon');
ylabel('U/\epsilon');
subplot(2,1,2)
CV = diff(U)/(x(2)-x(1));
xx = 0.5*(x(2:end)+x(1:end-1));
plot(xx,CV);
xlabel('kT/\epsilon');
ylabel('C_V/\epsilon');
%
%mysavefig('tmp.pdf')
```

We find the heat capacity at constant V from

$$C_V = \left(\frac{\partial U}{\partial T}\right)_{N,V} \ ,$$

which is

$$C_V = \epsilon \frac{\partial}{\partial T} \frac{1}{\exp \epsilon/kT + 1} \ ,$$

here we can use a trick – which comes from a different derivation of Boltzmann's distribution: We introduce $\beta = 1/kT$, so that

$$U = \frac{1}{1 + \exp(\beta\epsilon)} \ ,$$

now we can write

$$\frac{\partial}{\partial T} = \frac{\partial}{\partial \beta} \frac{\partial \beta}{\partial T} \frac{\partial}{\partial \beta} \frac{-1}{kT^2} \ .$$

This makes it easier to do the derivative:

$$C_V = \epsilon \frac{\partial}{\partial T} \frac{1}{\exp \epsilon/kT + 1} = -\frac{\epsilon}{kT} \frac{\partial}{\partial \beta} \frac{1}{1 + \exp \beta\epsilon}$$

$$= k\epsilon\beta^2 \, \epsilon \, \exp \beta\epsilon \frac{1}{(1 + \exp \beta\epsilon)^2}$$

$$= k \left(\frac{\epsilon}{kT}\right)^2 \frac{\exp(\epsilon/kT)}{(\exp(\epsilon/kT) + 1)^2} \ ,$$

5

5 Relating fluctuations in energy to heat capacity

There is a general result, that is interesting from a computational perspective. First, let us find the average energy:

$$\bar{U} = \frac{1}{Z} \sum_s \epsilon_s \exp(-\beta\epsilon)$$

$$= \frac{1}{Z} \sum_s -\frac{\partial}{\partial\beta} \exp(-\beta\epsilon)$$

$$= -\frac{1}{Z} \frac{\partial}{\partial\beta} \sum_s \exp(-\beta\epsilon)$$

$$= -\frac{1}{Z} \frac{\partial Z}{\partial\beta}$$

$$= -\frac{\partial}{\partial\beta} \ln Z \;,$$

Similarly, we can express the heat capacity as the derivative of U:

$$C_V = \frac{\partial\bar{U}}{\partial T} = \frac{\partial\beta}{\partial T} \partial\bar{U}\partial\beta$$

$$= \frac{1}{kT^2} \left(\frac{1}{Z} \frac{\partial^2 Z}{\partial\beta^2} - \frac{1}{Z^2} \left(\frac{\partial Z}{\partial\beta}\right)^2 \right) \;.$$

Now, we see that

$$\bar{U^2} = \sum_s \epsilon_s^2 \exp(-\beta\epsilon_s) = \frac{1}{Z} \frac{\partial^2 Z}{\partial\beta^2} \;,$$

and we get the very nice relation:

$$C_V = \frac{1}{kT} \left(\bar{U^2} - \bar{U}^2 \right) \;.$$

This means that if we can measure the fluctuations in U we can use this to find the heat capacity – nice. In particular if we study a molecular/atomic system!

(Talk about how to measure things in an MD simulation).

6 Einstein crystal

It is now simple to characterize the energies of a single oscillator in the Einstein crystal. In this case, the states are numbered by n, and there is only one state per energy. We sum of the states to find the partition function:

$$Z = \sum_n \exp(-\beta\epsilon_n) = \sum_{n=0}^{\infty} \exp(-\beta\epsilon n) \;,$$

which is a geometric series, with sum:

$$Z = \frac{1}{1 - \exp(-\beta\epsilon)} \;.$$

We can then find the average energy either directly from the sum or from the expression we found above:

$$\bar{U} = -\frac{\partial}{\partial\beta} \ln Z = \epsilon \frac{1}{1 - \exp(-\beta\epsilon)} \exp(-\beta\epsilon)$$

$$= \epsilon \frac{1}{\exp(\beta\epsilon) - 1} \;.$$

We can rewrite this now, introducing $\beta = kT$ getting:

$$\bar{U} = \epsilon \frac{1}{\exp(\epsilon/kT) - 1} \ .$$

we can introduce the temperature $\theta_E = \epsilon/k$ – the Einstein temperature for the oscillator – getting:

$$\bar{U} = \epsilon \frac{1}{\exp(\theta_E/T) - 1} \ .$$

It is illustrative to plot this expression as a function of T/θ_E using:

```
% Einstein crystal energy and heat capacity
% U(T/T_E) - variables U/epsilon and x = T/T_E
clear all; clf
x = linspace(0,1.5,1000);
U = 1./(exp(1./x)-1);
subplot(2,1,1)
plot(x,U)
xlabel('$T/\theta_E$');
ylabel('$U/\epsilon$');
subplot(2,1,2)
CV = diff(U)/(x(2)-x(1));
xx = 0.5*(x(2:end)+x(1:end-1));
plot(xx,CV);
xlabel('$T/\theta_E$');
ylabel('$C_V/\epsilon$');
%
%mysavefig('tmp.pdf')
```
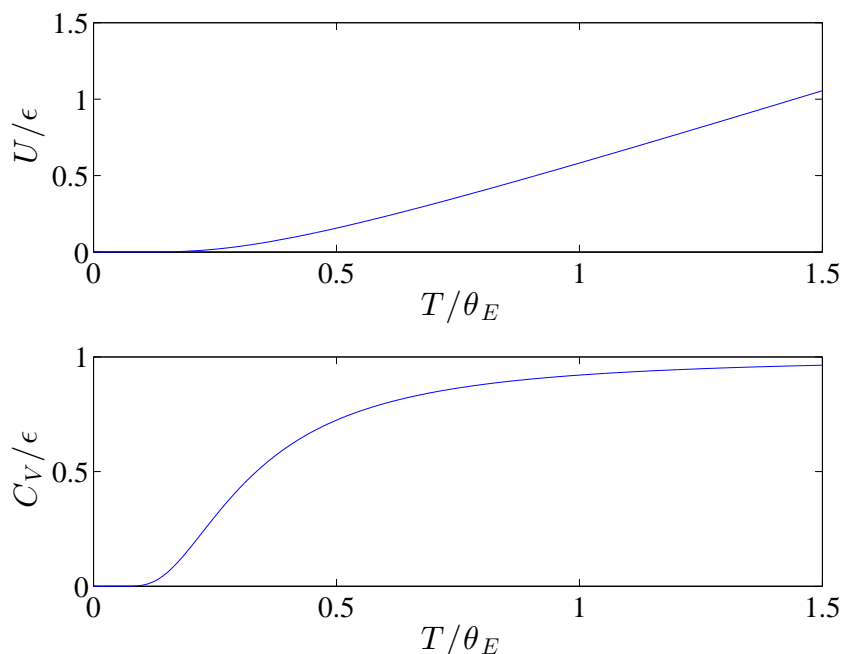

(I leave it for you to find an analytic expression for the heat capacity.)

It is interesting to notice the behavior of the heat capacity for a single harmonic oscillator now – this would be the contribution to the heat capacity for the vibration of a diatomic molecule. Notice that when $T \ll T_E$ there is no contribution from the oscillations, whereas when $T \gg T_E$ (or about at T_E), there is a contibution from the oscillations. This allows us to sketch the expected effects of oscillations on the behavior of a diatomic ideal gas – where the behavior goes from $3/2NkT$ to $5/2NkT$ at around T_E for the corresponding oscillation. A nice prediction.

7   Einstein crystal – numerical example

We can also use the Einstein crystal to demonstrate what happens with a single oscillator in a system with many oscillators. Let us assume that we have a system with $N = 1000$ oscillators, and $q = 500$ energy units. Let us then look at the probability for a given oscillator to be in a given energy state.

We divide the system into two parts, part A consists of a single oscillator, and part B consists of the remaining $N - 1$ oscillators.

How can we find the probability for q_A ($q_B = q - q_A$)? We count the number of states with q_A.

We already have developed programs to do this. For example the following which calculates the probability P_A for the system to be in a state q_A:

```
% Einstein solid NA - NB probability
% Random sampling compared with exact values
clear all;clf;
NA = 1;
NB = 199;
qq = [150 200 250 300 400 450 500];
nqq = length(qq);
for iqq = 1:nqq
    q = qq(iqq);
%q = 200;
N = NA + NB;
% Find multiplicity of all macrostates
nstate = q+1;
omegaA = zeros(nstate,1);
qAvalue = omegaA;
omegaB = omegaA;
omegaTOT = omegaA;
% Loop through all macrostates and find multiplicity
for istate = 1:nstate
    qA = istate-1;
    qAvalue(istate) = qA;
    omegaA(istate) = nchoosek(qA+NA-1,qA);
    qB = q - qA;
    omegaB(istate) = nchoosek(qB+NB-1,qB);
    omegaTOT(istate) = omegaA(istate)*omegaB(istate);
end
PA = omegaTOT/sum(omegaTOT);
j = find(qAvalue<40);
subplot(2,1,1)
hold all
plot(qAvalue(j),log10(PA(j)),'-o');
xlabel('$q_A$')
ylabel('$log_{10}[P(q_A)]$')
subplot(2,1,2)
hold all
plot(qAvalue(j)/q,log10(PA(j)),'-o');
xlabel('$q_A/q$')
ylabel('$log_{10}[P(q_A)]$')
end

%
%matrix = [omegaA omegaB omegaTOT]
```
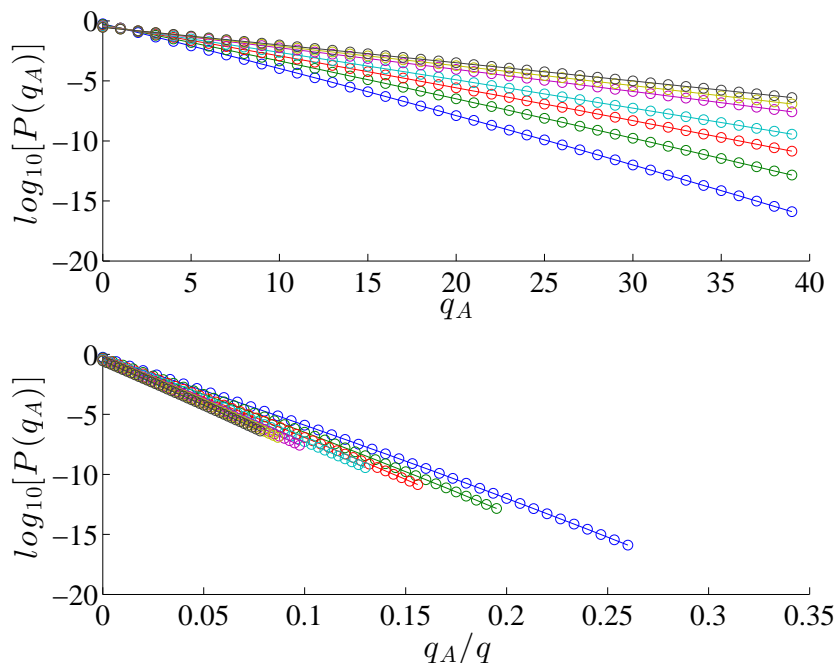

This indeed shows that the probability has an exponential form.

How would it depend on temperature? How can we change the temperature of the system? By changing the internal energy, q. Let us therefore look at different versions with different temperatures by changing q. We can also see how the rescaled plot would look like by plotting $P(q_A)$ as a function of $q_A/kT \sim q_A/q$, since $q = NkT/\epsilon$, and N is not changed in these simulations.

8   Spin system example

We can now apply these methods to address the behavior of a paramagnetic system. For a single dipole in a magnetic field, there are two possible states, one state with spin up and enegy $-\mu B$ and one state with spin down and energy μB.

The partition function for a single spin is therefore:

$$Z = \sum_s \exp(-\beta \epsilon_s) = \exp(-\beta \mu B) + \exp(\beta \mu B) = 2 \cosh(\beta \mu B) \ .$$

That was easy!

We can find the average energy, $\bar{U}$, from:

$$\bar{U} = -\frac{\partial}{\partial \beta} \ln Z = -\frac{1}{\cosh(\beta \mu B)}(\sinh(\beta \mu B)\mu B) = \mu B \tanh(\beta \mu B) \ ,$$

and to get the result for N independent spins we simply multiply by N.

This was ver simple compared with the long calculation you had to do in the microcanonical ensemble!

9   Ising spin system

However, this does become more complicated when we introduce interactions among the spins. For example, we may introduce the Ising model, where the interactions between two neighboring spins are given by the interaction energy:

$$-J s_i s_j \ ,$$

where $s_i = \pm 1$ and $s_j = \pm 1$ describes the states of spin i and spin j – where i and j are nearest neighbors. This means that the energy contribution from this interaction is low $(-J)$ is the two spins are pointing in the same direction, and it is relatively hight $(+J)$ is the two spins are pointing in the opposite direction. This spin interaction will therefore lead to a situation where all the spins point in the same direction – given by the extenal field – in the lowest energy configuration. However, as temperature increases fluctuations will arise, and it is not so clear what happens.

In addition, each spin interacts with the external magnetic field, B. The total energy of the system is therefore

$$\epsilon_s = -\sum_{i,j} J s_i s_j - \mu B \sum_i s_i \ ,$$

where the sum is over all the neighbors i, j in the system – this depends on the dimensionality of the system.

How would we find the partition function now?

1. You may solve the 1d model – Ising did this.

2. You may be *very* smart. Onsager solved the 2d Ising model. Nobody has solved the 3d Ising model. The Nobel prize was awarded for solving the 2d Ising model (but not to Onsager for this).

3. You may have a computer!

Let us pursue the last option – what if we have a computer? We may then use a simple method to sample the space of lattice configuration. Let us sketch a method here, called heat bath Monte Carlo, and we will discuss this method more further on.

The heat-bath Monte Carlo method for the Ising model consists of the following steps:

- Choose a random site i

- Find the energy of this site if $s_i = 1$ and if $s_i = -1$, E_+ and E_-.

- Set the spin up with probability $\exp(-\beta E_+)/C$ and down with probability $\exp(-\beta E_-)/C$ with $C = \exp(-\beta E_+) + \exp(-\beta E_-)$.

- Repeat.

How do we find the energy of the site? We calculate it from the energy function. We therefore need to know $m_i = \sum_j s_j$ for the neighbors of size i and calculate $E_+ = -Jm_i - H$ and $E_- = Jm_i + H$, where $H = \mu B$.

We can do this for one site i at a time – and this is done in the following script.

This is implemented in the following simple program for heat bath Ising systems.

```
% Heat bath Monte Carlo Ising model one spin at a tim
clear all; clf;
nstep = 100; % Number of lattice changes
N = 100; % System size in x and y direction
Jdivk = 2.0/3.0; % Interaction
Hdivk = 0.0; % External field
T = 0.1; % Dimensionless temperature
JdivkT = Jdivk/T;
HdivkT = Hdivk/T;
% Make initial random spin configuration
spins = randi(2,N,N)*2-3;
sizespins = size(spins);
N2 = N*N;
% Evolve system
for istep = 1:nstep
    for jstep = 1:N2
        posx = randi(N,1);
        posy = randi(N,1);
        neighbors = [sub2ind(sizespins,mod(posx+1-1,N)+1,posy) ...
            sub2ind(sizespins,mod(posx-1-1,N)+1,posy) ...
            sub2ind(sizespins,posx,mod(posy+1-1,N)+1) ...
            sub2ind(sizespins,posx,mod(posy-1-1,N)+1)];
        sumneigh = sum(spins(neighbors));
        thisspin = spins(posx,posy);
        DeltaE = -JdivkT*thisspin*sumneigh - HdivkT*thisspin;
        change = (rand(1,1)<exp(DeltaE))*(-2)+1;
        spins(posx,posy) = spins(posx,posy)*change;
    end
    imagesc(spins); axis square; drawnow;
end
```

However, this method is very slow. We can speed it up by doing it for all states each step. Now, we cannot really do it for all sites at the same time – because if we change both i and one of its neighbors j at the same time, we do not really know how the energy changed. Instead, what we can do is to change.

```
% Heat bath Monte Carlo Ising model half lattice at a time
clear all; clf;
nstep = 1000; % Number of lattice changes
N = 200; % System size in x and y direction
T = 0.8; % Dimensionless temperature
JdivkT = 1.0/T; % Interaction
HdivkT = 0.0/T; % External field
% Make initial random spin configuration
spins = randi(2,N,N)*2-3;
% Find half-matrix of sites for spin change
halflattice = zeros(N,N);
halflattice(1:2:N,2:2:N)=1;
halflattice(2:2:N,1:2:N)=1;
% Evolve system
for istep = 1:nstep
    sumneighbors = circshift(spins,[0 1]) + ...
        circshift(spins,[0 -1]) + ...
        circshift(spins,[1 0]) + ...
        circshift(spins,[-1 0]);
    DeltaEdivkT = -spins.*(JdivkT*sumneighbors+HdivkT);
    pboltzmann = exp(DeltaEdivkT);
    changespin = -2*(rand(N,N)<pboltzmann).*halflattice+1;
    spins = spins.*changespin; % Flip spins
    halflattice = 1-halflattice; % Other half is next
    imagesc(spins); axis square; drawnow;
end
```

How can we know that this method produces states that are consistent with a particular temperature? We cannot from the argument we have provided so far, but if you are interested in pursuing a good argument for this you can look up Markov chains. This will be taught in more advanced courses in statistical mechanics. However, we do see that the method ensures that all transitions are according to Boltzmann statistics for the relative probabilities of two states, so that we may expect the state also to reflect Boltzmann statistics.

10   States and energy levels

Notice – and this is important – that the sum in the partition function is over all the states of a system. Not over all the possible energies, but over all the states. What is the difference? Several states may have the same energy!

We can group microstates with the same energies together. Each such state will contribute a term $\exp(-\beta\epsilon_i)$ to the sum. Instead of summing over all the states i, we can sum over all the energies, but then we need to include how many states have the energy ϵ:

$$Z = \sum_{\text{states } i} \exp(-\beta\epsilon_i) = \sum_{\text{energy levels } n} \Omega(E_n)\exp(-\beta E_n) \ .$$

We will use this distinction many times, also when the sum is difficult to calculate.

10.1   Example

For example, we could have studied a system where the number of states with the same energy increases with the energy:

$$\epsilon_i = \epsilon(1, 2, 2, 3, 3, 3, 4, 4, 4, 4)$$

In this case the multiplicity of energy ϵ is 1, of 2ϵ is 2, of 3ϵ is 3 and so on.

The partition function in this case is

$$Z = \sum_i \exp(-\beta\epsilon) = \sum_n \Omega(\epsilon_n)\exp(-\beta\epsilon_n) \ ,$$

where $\Omega(1) = 1$, $\Omega(2) = 2$, $\Omega(3) = 3$ and $\Omega(4) = 4$. The sum is therefore:

$$Z = 1 \cdot \exp(-\beta\epsilon) + 2 \cdot \exp(-2\beta\epsilon) + 3 \cdot \exp(-3\beta\epsilon) + 4 \cdot \exp(-4\beta\epsilon) \ ,$$

And the mean energy, we find in the same way:

$$\bar{U} = \sum_i \epsilon_i \exp(-\beta\epsilon_i) = \sum_n \Omega(n)\epsilon_n \exp(-\beta\epsilon_n) \ ,$$

where the sum over i is over all the states, and the sum over n is over all the energies.

The average energy is therefore:

$$\bar{U} = \epsilon \frac{1 \cdot \exp(-\beta\epsilon) + 2 \cdot 2\exp(-2\beta\epsilon) + 3 \cdot 3\exp(-3\beta\epsilon) + 4 \cdot 4\exp(-4\exp\epsilon)}{Z} \ .$$

11   Sharpness of the canonical ensemble

You may be worried since in the microcanonical ensemble the energy is given – it is constant – but in the canonical ensemble the energy may take on any value. Its average is given, but it may in principle vary from the average value. But how large are the fluctuations in energy?

We can estimate this because we know that the energy typically will be an extensive property – it will be proportional to N. Similarly with C_V: Both are proportional to N. We can therefore use the result from above to discuss the relative fluctuations of the energy:

$$\frac{\sqrt{\bar{U^2} - \bar{U}^2}}{\bar{U}} = \frac{\sqrt{kT^2 C_V}}{\bar{U}} \sim \frac{N^{1/2}}{N} \sim N^{-1/2} \ ,$$

where we have used the symbol $\sim$ to represent "scales as", meaning that it has the same dependence on N. In the limit of large N the fluctuations in the energy $\bar{U}$ is therefore small – this suggests that we can use the results form either example when we address thermodynamic systems.

12 Pressure

What happens if we change the volume of a system in a microstate ϵ_s? Generally, we expect the energy to be a function of the volume, but we assume that the volume change is so small that the system remains in the same quantum state, only the energy of the states. If we now change the volume slowly (reversibly) from V to $V + \Delta V$ by an external force. The energy of state s after the reversible volume change is

$$\epsilon_s(V + \Delta V) = \epsilon_s(V) + \frac{d\epsilon_s}{dV}\Delta V + \dots .$$

Now, if the change in volume corresponds to changing all the walls with Δx, Δy, and Δz, then the volume change is $\Delta V = A(\Delta x + \Delta y + \Delta z)$. The work done by this volume change is

$$W = -p_s A(\Delta x + \Delta y + \Delta z) = -p_s \Delta V ,$$

Since we have assumed that the change in volume is without any changes to the energy states of the system, the change is without any change in entropy and hence also without any heat (when the system is changed reversibly, quasistatically). The work therefore corresponds to the change in energy:

$$dU = Q + W = -p_s \Delta V ,$$

and wer can now recognize that

$$p_s = -\frac{d\epsilon_s}{dV} .$$

Now, if we introduce the average pressure of the system as:

$$\bar{P} = \sum_s p_s P(s) = -\sum_s \frac{d\epsilon_s}{dV} P(s) = -\frac{d}{dV}\sum_s \epsilon_s P(s) = -\left(\frac{\partial U}{\partial V}\right)_S ,$$

where we have written that S must be constant. It may be tempting to also write that T, N is constant, and thus refer it to the Canonical ensemble that this calculation was done in. This is correct, while not properly proven here.

We could also express the sum in eq.12 directly from the partition function, and in that case we find that:

$$\bar{P} = \sum_s \frac{d\epsilon_s}{dV}\frac{1}{Z}\exp(-\beta\epsilon_s) = kT\left(\frac{\partial \ln Z}{\partial V}\right)_{T,N} ,$$

since for the derivation of the partition function and $P(s) = (1/Z)\exp(-\beta\epsilon_s)$ we assumed constant N, V, T, and now we are allowing V to change.

12.1 Thermodynamic identity

We can use these expressions now to again reproduce the thermodynamic identity. For example, we can directly expand the differential for S:

$$dS(U, V, N) = \left(\frac{\partial S}{\partial U}\right)_{N,V} dU + \left(\frac{\partial S}{\partial V}\right)_{U,N} dV ,$$

which now again gives us the theromdynamic identity.

12.2 Statistical entropy

This expression now allows us to derive an interesting result for the entropy.

From the average energy we get:

$$\bar{U} = \sum_s \epsilon_s P(s) \ ,$$

Now, the total change in energy is therefore:

$$dU = \sum_s \epsilon_s dP(s) + \sum_s P(s) d\epsilon_s \ ,$$

Where we can write the second term as:

$$\sum_s P(s) d\epsilon_s = \sum_s P(s) \frac{d\epsilon_s}{dV} dV = -\bar{P} dV \ ,$$

and therefore the change in energy is:

$$d\bar{U} = \sum_s \epsilon_s dP(s) - \bar{P} dV \ .$$

If we compare with the thermodynamic identiy, we see that we can interpret:

$$TdS = \sum_s \epsilon_s dP(s) \ ,$$

This expressions can be (with some tricks) be related to another expression for the entropy – which we could instead have started from:

$$TdS = \sum_s \epsilon_s dP(s) = -kT \sum_s \ln Z dP(s) - kT \sum_s \ln P(s) dP(s)$$

where we have used that

$$P(s) = (1/Z) \exp(-\epsilon_s/kT) \ ,$$

and therefore that

$$\epsilon_s = -kT(\ln Z + \ln P(s)) \ .$$

Notice that $\sum_s dP(s) = d \sum_s P(s) = d(1) = 0$. We therefore get:

$$dS = -k \sum_s \ln P(s) dP(s) = -k \sum_s d(P(s) \ln P(s)) \ ,$$

and

$$S = -k \sum_s P(s) \ln P(s) \ ,$$

which is a general expression for the entropy. This relation also holds in the microcanonical ensemble, since then $P_s = 1/\Omega$ for all states and therefore we find $S = k \ln \Omega$ after summing.

13   Equipartition theorem

Let us look at a general result for systems with an energy that has a quadratic shape:

$$\epsilon(q) = cq^2 \ ,$$

where q is a degree of freedom, such as the position, the velocity, the angular velocity, or any other.

We will assume that the states are densely, but discretely placed in q-space (but that they are not degenerate), so that there are states every Δq. The partition function is therefore

$$Z = \sum_q \exp(-\beta cq^2) \ ,$$

which we will approximate by the integral over all q-values:

$$Z = \frac{1}{\Delta q} \sum_q \exp(-\beta c q^2) \Delta q = \frac{1}{\Delta q} \int \exp(-\beta c q^2) dq ,$$

we change variable to $x = \sqrt{\beta c} q$, so that $dx = \sqrt{\beta c} dq$, and

$$Z = \frac{1}{\Delta q} \frac{1}{\sqrt{\beta c}} \int_{-\infty}^{\infty} e^{-x^2} dx ,$$

where the integral is simply a numerical value (which we do not need to know) – we call it b.

$$Z = \frac{1}{\Delta q \sqrt{\beta c}} b = C \beta^{-1/2} .$$

(Notice that we are here only conserned with the scaling with β of this function – not the exact functional form.)

The energy is then found from:

$$\bar{U} = -\frac{1}{Z} \left(\frac{\partial Z}{\partial \beta} \right) = \frac{1}{C \beta^{-1/2}} (-1/2) C \beta^{-3/2} = \frac{1}{2} \beta = \frac{1}{2} kT ,$$

and we have proved the equipartition theorem – which is true as long as the states are spaced in q-space with no degeneracy.

This is ok as long as the spacing between levels is much smaller than kT – which is not true for quantum systems at low energies.

14 Helmholz free energy

The function

$$F = U - TS ,$$

is called Helmholtz free energy.

This function plays the same role in systems with constant T as the energy U plays in ordinary mechanical processes (constant entropy).

Helmholtz Free energy is minimum in a system S in thermal contact with a reservoir R, if the volume of the system is constant.

We call such a principle a "Free energy minium principle". There will be more such principles later – for other ensembles.

First, we show that F is an extremum in equilibrium at constant T and V. For an infinitesimal, reversible transfer from R to S

$$dF_S = dU_s - T dS_s ,$$

at constant T. Here, we know that $1/T = (\partial S/\partial U)_V$ so that $dU_S = T dS_S$ when the volume is constant. Therefore $dF_S = 0$ and F_S is an extremum.

We can also show that it is a minimum:

The total energy is $U = U_S + U_R$. The total entropy is

$$S = S_R + S_S = S_R (U - U_S) + S_S(U_S) \simeq S_R(U) - U_S \cdot \left(\frac{\partial S_R}{\partial U_R} \right)_{V,N} + S_S(U_S) .$$

Where we also know that

$$\left(\frac{\partial S_R}{\partial U_R}\right)_{V,N} = \frac{1}{T} \ ,$$

and therefore we get

$$S = S_R(U) - \frac{1}{T}F_S \ ,$$

where $F_S = U_S - TS_S$ is the free energy of the system. Now $S_R(U)$ is a constant. We also know that $S = S_R + S_S$ is maximum with respect to U_S in equilibrium. There F_S must be a minimum.

The equilibrium state of a system with constant V and T is characterize by the Helmholtz free energy being minimal.

In a process going toward equilibrium, we will find that $\Delta F \leq= 0$.

Now, F is a state function, a function of T, V, N since E and S are state functions.

If we change a bit the external values V, N and T, the state function $F(T, V, N)$ will get a new value:

$$dF = d(E - TS) = dE - SdT - TdS$$

but we also know that

$$dE = TdS - pdV + \mu dN \ ,$$

and therefore

$$dF = -S\,dT - p\,dV + \mu\,dN \ ,$$

Now, since F is a state function, $F = F(V, T, N)$, we also know that

$$dF = \left(\frac{\partial F}{\partial T}\right)_{N,V} dT + \left(\frac{\partial F}{\partial V}\right)_{T,N} dV + \left(\frac{\partial F}{\partial N}\right)_{T,V} dN \ ,$$

We therefore find that

$$S = -\left(\frac{\partial F}{\partial T}\right)_{N,V} \ ,$$

$$p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \ ,$$

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \ .$$

In addition, we notice that

$$(\Delta F)_T = -p\Delta V = -\Delta W \ ,$$

so that the work ΔW done on the environment by the system at constant temperature (and N) is equal the loss in free energy for the system: Hence the name *free energy*. (This is only true for reversible processes).

15   Maxwell relations

We can now find a new group of relations – thermodynamic relations – based on the second derivative of Helmholtz free energy: If the energy is continuous in its variables, we know that the order of derivatives for the second derivative is not important, therefore:

$$\frac{\partial^2 F}{\partial V \partial T} = \frac{\partial^2 F}{\partial T \partial V} \ ,$$

we can use this to find that

$$\frac{\partial p}{\partial T}_V = frac\partial S\partial V_T \ .$$

We will find more such relations later as we introduce more variables. (There are three possible permutations in this case, but the other two involve the chemical potential, which we are not familiar with yet).

16 Connection between micro- and macro in the canonical ensemble

Why do we care about F – It is very simple to calculate F from the partition function – hence it is a method to find all the other thermodynamic properties directly.

To connect the microscopic and the macroscopic pictures in the canonical ensemble, we will show that Helmholtz free energy is related to the partition function through the beautifully simple relation

$$F = -kT \ln Z \ .$$

We have introduced $F = U - TS$ and we know that $S = (\partial F/\partial T)_V$:

$$F = U + T \left(\frac{\partial F}{\partial T} \right)_V \ ,$$

This differential equation is satisfied by

$$F = -kT \ln Z \ ,$$

as we see by substitution:

$$\left(\frac{\partial F}{\partial T} \right)_V = \frac{\partial}{\partial T} - kT \ln Z = -k \ln Z - kT \frac{\partial \ln Z}{\partial T} = -k \ln Z - kT \frac{U}{kT^2} \ ,$$

we insert this into the differential equation above, and find that

$$U + T \left(-k \ln Z - (U/T) \right) = -k \ln Z \ ,$$

which means that this is one possible solution. But there may be a constant term $F = -kT \ln Z + aT$. However, the entropy must reduce to $\ln \Omega_0$ when T goes to zero, which only can happen if $a = 0$. Hence, we have found that fundamental relationship between thermodynamics and the microscopic models for the canonical ensemble:

$$F = -kT \ln Z \ .$$

I find the equation beautiful. It is simple. It connects the micro and the macro. It allows all types of calculations of S, p and μ from the partition function. (Ok - we have found some of these relations before too, but the derivation of the expression for p was somewhat weak. This expression is very strong!)

We can now apply this to find all thermodynamic properties of systems described in the canonical ensemble.

Let us use this to address the ideal gas – again.

17 Composite systems

Before we do the ideal gas of many, identical particles, let us starty by studying systems of identical and non-identical particles.

Let us start by a system consisting of two different particles, each described independently by the quantum states $\epsilon_{1,i}$ and $\epsilon_{2,j}$.

The partition function is

$$Z = \sum_i \sum_j \exp(-\beta\epsilon_{1,i}) \exp(-\beta\epsilon_{2,j}) = \left(\sum_i \exp(-\beta\epsilon_{1,i}) \right) \left(\sum_j \exp(-\beta\epsilon_{2,j}) \right) = Z_1 \cdot Z_2 \ .$$

For two non-interacting systems, with distinguishable particles, the partionfunction is the product of the partition functions!

Now, what if the two particles are not distinguishable?

In that case, we have included too many elements in the sum, because we have included both

$$\exp(-\beta(\epsilon_{2,j} + \epsilon_{1,i})$$

and

$$\exp(-\beta(\epsilon_{2,i} + \epsilon_{1,j}) \ ,$$

but these two states are the same, and should only be counted once.

We should therefore divide by 2:

$$Z_{TOT} = \frac{1}{2} Z_1 Z_2 \ ,$$

and this is almost correct, but I have then also divided the terms

$$\exp(-\beta(\epsilon_{2,i} + \epsilon_{1,i}) \ ,$$

by two – and these terms have not been double counted. However, the mistake we make by this will be small as long as there are few of such states.

Generalized to N particles we get:

$$Z = Z_1 Z_2 \ldots Z_N \ ,$$

for distinguishable particles, and

$$Z = \frac{Z_1 Z_2 \ldots Z_N}{N!} = \frac{Z_1^N}{N!} \ ,$$

for identical particles.

18   Ideal gas in the canonical ensemble

We would therefore like to find Z_1 for one particle in a box to find the partition function for an ideal gas.

But even for a single particle, there can be several, independent degrees of freedom, and each of these would be considered independent states, and therefore independent partition function. For a diatomic gas we may have translation, vibration and rotation, and the corresponding partition function for a single particle is then:

$$Z_1 = Z_{\text{trans}} \cdot Z_{\text{rot}} \cdot Z_{\text{vib}} \ ,$$

and if there are further internal degrees of freedom we should include these as well.

Let us now address the partition function for a single atom (particle) in a box of size L (and volume V). Agagin, we know that the states of the system are given by n_x, n_y, n_z for the particle, with

$$\epsilon_n = \frac{h^2}{8mL^2} \left(n_x^2 + n_y^2 + n_z^2 \right) \ .$$

The partition function is given by the sum over all the states n_x, n_y, n_z:

$$Z = \sum_{n_x} \sum_{n_y} \sum_{n_z} \exp\left(-\beta \frac{h^2}{8mL^2}\left(n_x^2 + n_y^2 + n_z^2\right)\right) \,,$$

which is the product of three identical sums, ξ, so that:

$$Z = (\xi)^3 \,,$$

where

$$\xi = \sum_{n=0}^{\infty} \exp(-\lambda^2 n^2) \,,$$

where we have introduced λ as

$$\lambda^2 = \beta \frac{h^2}{8mL^2} \,.$$

For small values of λ we can approximate the sum by the integral:

$$\xi \simeq \int_0^{\infty} \exp(-\lambda^2 n^2) \,, dn = \frac{\sqrt{\pi}}{2\lambda} \,.$$

The partition function is therefore given by:

$$Z = \frac{\pi^{3/2}}{8}\left(\frac{2mL^2}{\beta\hbar^2\pi^2}\right)^{3/2} = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} V \,.$$

It is usual to introduce the *quantum density* n_Q through:

$$n_Q = \left(\frac{m}{2\pi\hbar^2\beta}\right)^{3/2} \,,$$

and the partition function is then

$$Z = V n_Q = \frac{n_Q}{n} \,,$$

where $n = 1/V$ is the density of particles.

The average particle energy is given by:

$$\bar{U} = \langle\epsilon\rangle = -\frac{\partial \ln Z}{\partial \beta} = \frac{3}{2}kT \,.$$

This is in correspondence with the equipartition principle.

Notice the effect of adding in addition internal degrees of freedom, this would correspond to adding in addition products in Z, which would give additional addends in the sums in $\bar{U}$ and in F. We can therefore indeed add in the effect of rotations and vibrations in both the energy and in the free energy – as we did to address the heat capacity of a vibrating atom in an ideal gas.

19 Numerical modeling in the canonical ensemble

The canonical ensemble lends itself to modeling. Indeed, the Monte Carlo method is widely used as a numerical/computational method to gain insight into real systems – biological, technological, geological.

We saw a first version of the Monte Carlo method above – the heat bath Monte Carlo. For the Ising system with energy:

$$\epsilon = \sum_i \left(S_i \sum_{j:i} -JS_j - H\right) \,,$$

we picked a spin at random and replaced it with a spin S_i' with probabilities $P(E_1) \propto \exp(-\beta E_1)$ and $P(E_2) \propto \exp(-\beta E_2)$ for the two possible energies E_1 and E_2 of the spin in the system – thus trying to ensure that the system is consistent with Boltzmann statistics.

19.1 Monte Carlo estimates of integrals

This method is an example of a Monte Carlo algorithm. Monte Carlo methods can be used for many things – not only physics, but also to solve mathematical problems. For example, we could use Monte Carlo methods to find an area or to solve an integral.

We can use the "hit or miss" algorithm to calculate the definite integral of a function:

$$y = \int_{x_0}^{x_1} f(x)\, dx \ .$$

This integral can be interpreted as the integral under the curve $f(x)$ from x_0 to x_1. To estimate this we choose a box from $y = 0$ to some $y_1 > f(x)$ for all x on the interval. We then select N points randomly in the interval x_0 to x_1 and $y_0 = 0$ to y_1. We can then estimate the integral by the fraction of points falling below the curve $f(x)$ multiplied by the area of the box:

$$y \simeq \frac{N_{\text{below}}}{N} (y_1 - y_0)(x_1 - x_0) \ .$$

We expect this to become a better and better estimate as N goes to infinity.

We can use this method to estimate the area of a circle and thereby estimate the value of π. The method is as follows: We select N points on the interval 0 to 1 for both x and y and count how many times the points are close to the origin than 1, that is, we count when $x^2 + y^2 \le 1$. (Notice that we compare the square – we do not take the square root since this is computationally inefficient!). The number of point inside is N_{below} and the estimate for π is:

$$\frac{1}{4}\pi \simeq \frac{N_{\text{below}}}{N} \ ,$$

and therefore

$$\pi \simeq 4\,\frac{N_{\text{below}}}{N} \ .$$

This is done in the following program:

```
% Monte Carlo estimate of pi
N = 100000;
N0 = 0;
piest = zeros(N,1);
for i = 1:N
    x = rand(1,1);
    y = rand(1,1);
    r2 = x*x+y*y;
    N0 = N0 + (r2<=1);
    piest(i) = 4*N0/i;
    if (mod(i,100)==0)
    plot((1:i),piest(1:i),'-r',[i i],[pi pi],'o')
    drawnow
    end
end
piest = 4*N0/N
```

19.2 Metropolis Algorithm

This method is not that efficient. It corresponds to the most direct method for simulating the canonical ensemble. If we simulate an Einstein crystal, we could simulate the model by choosing an oscillator at random and randomly changing the energy by ± 1. We will then generate a sequence of microstates, and from this sequence we could estimate for example the average energy by:

$$\bar{U} = \frac{\sum_i \epsilon_i \exp(-\beta \epsilon_i)}{\sum_i \exp(-\beta \epsilon_i)} \ ,$$

where ϵ_i is the energy of state i of the Einstein crystal, and the sum is over all the states we have generated – not all possible states.

Unfortunately, this method, just like the integration method above is not very efficient. We may end up with many microstates that are unlikely, and therefore do not contribute much the sum. Instead,

we could choose microstates in a way that ensures that we have many states that contributed to the sum – we could choose microstates according to their importance. Such as method is called importance sampling, and one of the simples importance sampling methods is called the *Metropolis Algorithm*.

In the Metropolis Algorithm we generate a sequence of states, but the transition between the states depends on the transition probability from one state to another. This generates a sequence of states which we call a Monte Carlo path.

How do we select the transition probabilities between the states ϵ_i of the system?

Since the probability for a state with energy ϵ_i is

$$P(\epsilon_i) = \frac{1}{Z} \exp(-\beta \epsilon_i) \ ,$$

we could use this – but we do not know Z!

However, we can avoid this problem if we are only interested in the probability for a transition from one state ϵ_i to another state with energy ϵ_j – for example by changing a single spin at a time in the Ising system or if we move energy from one oscillator to another in the Einstein crystal. In this case, it is only the relative probabilities that matter (this will become clearer when you are introduced to Markov chains in a later course).

The change in energy when we go from state i to state j is ΔE, and the relative probability for the states are:

$$P(i)/P(j) = \exp(-\beta \Delta E) \ .$$

We may interpret this ratio as the transition probability from state i to state j, and this is indeed what is done in the first method developed – the Metropolis algorithm. In this case, the transition probability (rate) is chosen to be

$$W = \left(\begin{array}{ll} \exp(-\Delta E/kT) \ , & \Delta E > 0 \\ 1 \ , & \Delta E < 0 \end{array} \right.$$

We can therefore implement the algorithm using the following steps:

1. Choose an initial microstate – for example random spins in the Ising magnet or random energies for each oscillator in the Einstein crystal.

2. Make a random change in the microstate – a trial change. Calculate the change of energy going from the old state to the new state, ΔE. For example, for the Ising system we may choose a spin at random and see what happens if we flip it; for the Einstein crystal choose an oscillator at random and change its energy by ± 1.

3. Generate a random number, r, between 0 and 1

4. if $r \leq \exp(-\Delta E/kT)$ accept the change. (This formulation means that we will always accept a change where $\Delta E > 0$.)

5. Repeat

This generates a sequence of states that you can now make measurements on.

We can use this method to study the Ising system as well. This is implemented in the following program. It is not that different from the heat bath Monte Carlo method – but there are small differences.

19.3   Detailed balance

Are there any limits to what kinds of rules we can choose for the transitions? Could we invent any rule we like?

No! The transition rates must satisfy a principle called detailed balance. What is this.

What do we want to model? We want to model a sequence of microstates for the system – like a time development of microstates. Let us write the states as $|j\rangle$: The possible states of the system is $|1\rangle$, $|2\rangle$, $|3\rangle$, etc. If the system is in state $|i\rangle$ at time t we can write this as $|i\rangle(t)$.

We want to find the next state, $|j\rangle(t + \Delta t)$. Then, we need the probability to get to state $|j\rangle$ for all possible configurations at time t. And not only that – in general we may consider that case that the probability to be in state $|j\rangle$ at time $t + \Delta t$ depends not only on the state the system was in at a time t, but also all the previous states at earlier time. Hmmm. That was a lot of information. We can simplify the problem if we remove the history dependence – if we assume that the probabilty to be in a state $|j\rangle(t + \Delta t)$ only depends on the state $|i\rangle(t)$ it was in at time t and not on what happened before that. We can think of this as implying that the state $|i\rangle(t)$ contains all the relevant information about the system and determines what new states can occur. If the state only depends on the previous state and not on the history we call the sequence of states a Markov chain.

What we need is the probability $P(j, t + \Delta t)$ to be in state $|j\rangle$ for each possible state $|i\rangle$ at t. And we need this probability for all states $|j\rangle$ and $|i\rangle$. We call the probablity to go from state $|i\rangle$ at t to state $|j\rangle$ at $t + \Delta t$ the transition probability:

$$W_{ij} = W\left(i \rightarrow j\right) = P(j(t + \Delta t)|i(t)) \ ,$$

where the symbol | means "given than" – it is a conditional probability.

We know that $W_{ij} \geq 0$ and that $\sum_j W_{ij} = 1$: Starting from state $|i\rangle$ at time t we know that it must end up in some state $|j\rangle$ at time $t + \Delta t$.

The probability $P(j, t + \Delta t)$ is therefore:

$$P(j, t + \Delta t) = \sum_i P(j, t + \Delta t|i, t) = \sum_i W_{ij} P(i, t) \ .$$

What determines how this system develops? The master equation – a cool name – which is given as

$$\frac{dP(j, t)}{dt} = -\sum_i W_{ji} P(j, t) + \sum_i W_{ij} P(i, t) \ .$$

(Not finished)