This document is intended to serve as personal notes for self-study of *Monte Carlo Methods for Statistical Physics* by M.E.J. Newman and G.T. Barkema.

1 Introduction

1.1 Statistical Mechanics

Many systems of interest are composed of very many simple constituent systems. For example, a liter of oxygen is composed of 10^{22} oxygen molecules. The equations of motion for a single oxygen molecule are relatively simple but the sheer number of oxygen molecules in the larger system render the exact solutions infeasible. Instead, we approach the larger system via probabilistic methods. The goal is to express the state of the larger system as a set of probabilities of being in one state or another.

The typical systems studied in physics for which MCM are useful are systems described by a Hamiltonian H and an associated energy spectrum that is either continuous or given by a discrete set E_0, E_1, E_2, etc. We will also consider systems that interact with a thermal reservoir with which the system can exchange heat. Heat exchanges with the reservoir manifest as a small (negligible) perturbation in the Hamiltonian which pushes the system from one energy state to another.

Modeling the effect of the perturbation on the Hamiltonian is done by defining a *dynamics* for the system. The dynamics of the system may take the form of a set of transition rates. Suppose the system is in a state μ. Then, $R(\mu \to \nu)dt$ is defined to be the probability that the system transitions to a state ν in time dt, where $R(\mu \to \nu)$ is defined to be the *transition rate*. We may also define a set of weights $w_\mu(t)$ which represent that probability that the system will be found in state μ at time t. Using these elements, we can write down a *master equation*:

$$\frac{dw_\mu}{dt} = \sum_\nu [w_\nu(t)R(\nu \to \mu) - w_\mu(t)R(\mu \to \nu)] \tag{1}$$

The first term in the sum is the rate at which the system is transitioning into state μ. The second term is the rate at which the system is transitioning out of state μ.

We may define the expectation value of some observable Q, which takes the value Q_μ in state μ, as

$$\langle Q \rangle = \sum_\mu Q_\mu w_\mu(t)$$

The expectation value of Q can be interpreted two ways:

- It is the mean of the observed value of Q that would be obtained if we measured many identically prepared systems simultaneously.

- It is the time average of the quantity Q. That is, it might be the average observed value of Q after many measurements over time.

1.2 Equilibrium

An *equilibrium* state can be defined as a state in which all of the rates of change dw_μ/dt vanish, and therefore the transition rate into and out of each state μ must be equal. All systems goverened by the master equation must come to equilibrium after some time and so we can define the equilibrium values of the weights, the *equilibrium occupation probabilities*, as

$$p_\mu = \lim_{t \to \infty} w_\mu(t)$$

For a system in equilibrium with thermal reservoir at temperature T, the Boltzmann distribution is

$$p_\mu = \frac{1}{Z} e^{-E_\mu/kT} = \frac{1}{Z} e^{-\beta E_\mu}$$

where Z is the normalization constant, called the *partition function*, and defined as

$$Z = \sum_\mu e^{-\beta E_\mu}$$

Plugging this into our definition of expectation values gives us

$$\langle Q \rangle = \sum_\mu Q_\mu p_\mu(t) = \frac{1}{Z} \sum_\mu Q_\mu e^{-\beta E_\mu}$$

which is the expression for the expectation value of a system in equilibrium. For example, the internal energy U of a system is given by

$$U = \langle E \rangle = \frac{1}{Z} \sum_\mu E_\mu e^{-\beta E_\mu} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial logZ}{\partial \beta}$$

It is sometimes appropriate in MCM calculations to calculate the partition function, from which many other quantities can be derived. Some examples can be found in Section 1.2 of Newman and Barkema.

1.3 Fluctuations, Correlations, and Responses

It is useful to assess the fluctuations in observable quantities. Doing so allows us to determine how the quantity we are measuring varies over time and therefore how much of an approximation is being made by calculating the expectation value. As an example, the root-mean-square fluctuation in internal energy is given by

$$\sqrt{\langle E^2 \rangle - \langle E \rangle^2} = \sqrt{\frac{\partial^2 logZ}{\partial \beta^2}}$$

In the limit of a large system, called the *thermodynamic limit*, it is common for fluctuations to become negligible. For example, RMS fluctuations in internal energy scale as $\sqrt{V}$ in a system of volume V, but the internal energy itself scales

as V so as the system becomes very large, the fluctuations grow increasingly negligible.

Every parameter that can be fixed in a system (e.g. volume, external field) has a conjugate variable (e.g. pressure, magnetization) that is given by a derivative of the free energy. For example,

$$p = -\frac{\partial F}{\partial V}$$

$$M = \frac{\partial F}{\partial B}$$

The Hamiltonian contains terms of the form $-XY$ where Y is a "field" whose value is fixed, and X is the conjugate variable to which it couples. A common trick to calculate the thermal average of a quantity is the following:

- Make up a fictitious field that couples to the quantity of interest.

- Add an appropriate term to the Hamiltonian.

- Set the field to zero after performing the derivative on the free energy.

The *susceptibility* of X to Y is a measurement of the strength of the response of X to variation in Y. Typically, this quantity is denoted by χ and can be defined as

$$\chi = \frac{\partial \langle X \rangle}{\partial Y}$$

The *linear response theorem* tell us that the fluctuations in a variable are proportional to the susceptibility of the variable to its conjugate field.

1.3.1   Correlation Functions

The *disconnected correlation function* $G^{(2)}(i,j)$ is defined as

$$G^{(2)}(i,j) = \langle x_i x_j \rangle$$

This function gives us roughly an idea of how the variables x_i and x_j are correlated. Consider the following scenarios:

- If x_i and x_j move (roughly) together most of the time around the origin zero, then the average of their product will be a positive value.

- If x_i and x_j move (roughly) together but in opposite dirrections most of the time, then the average of their product will be a negative value.

- If x_i and x_j move with no discernible pattern relative to one another, then their product will be sometimes positive and sometimes negative, averaging to zero over a long time.